NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Contents:

The following documents are made available to consultees and commentators:

The **final scope and final stakeholder list** are available on the NICE website.

1. **Company submission** from **Roche Products**

2. **Clarification questions and company responses**
   a. Initial questions
   b. Follow up questions

3. **Patient group, professional group and NHS organisation submission** from:
   a. GIST Support UK
   b. Roy Castle Lung Cancer Foundation
   c. Sarcoma UK - *endorsed by patient expert Bradley Price*
   d. Royal College of Physicians - *endorsed by clinical expert Matthew Krebs*

4. **Expert personal perspectives** from:
   a. Debashis Sarker– clinical expert, nominated by Roche Products
   b. Peter Clark – CDF clinical lead, NHS England

5. **Evidence Review Group report** prepared by Centre for Reviews and Dissemination and Centre for Health Economics – York

6. **Evidence Review Group – factual accuracy check**

7. **Evidence Review Group - addendum**

8. **Technical engagement response** from **Roche Products**
   a. Additional information request

9. **Technical engagement response from consultees and commentators:**
   a. GIST Support UK
   b. Royal College of Physicians

10. **Evidence Review Group critique of company response to technical engagement** prepared by Centre for Reviews and Dissemination and Centre for Health Economics – York

11. **Final Technical Report**

12. Additional Submissions from NHS England:

a. Entrectinib and larotrectinib diagnostic costs

b. Options for phasing introduction of detection of NTRK gene fusions

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

ID1512: Entrectinib for treating *NTRK* fusion-positive solid tumours

Document B

Company evidence submission

May 2019

File name	Version	Contains confidential information	Date
ID1512_Entrectinib for treating NTRK fusion-positive solid tumours_DocB _20May19	V1.0	Yes	20th May 2019

Contents

Tables

Figures

Abbreviations

AE	adverse event
AIC	Akaike Information Criterion
ALK	anaplastic lymphoma kinase
ALT	alanine transaminase
AR	adverse reaction
ASCO	American Society of Clinical Oncology
AST	aspartate transaminase
BBB	blood brain barrier
BIC	Bayesian Information Criterion
BICR	blinded independent central review
BMI	body mass index
BNF	British National Formulary
BOR	best overall response
BSA	body surface area
BSC	best supportive care
CAP	College of American Pathologists
CBR	clinical benefit rate
CCOD	clinical cutoff date
CDF	Cancer Drugs Fund
CE	cost-effectiveness
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CLIA	Clinical Laboratory Improvement Amendments
CNS	central nervous system
CR	complete response
CRC	colorectal cancer
CSR	clinical study report
CT	computed tomography
DLT	dose limiting toxicity
DNA	Deoxyribonucleic
DOR	duration of response
DSU	decision support unit
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor

EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
ESMO	European Society for Medical Oncology
FDA	Food and Drugs Administration
FISH	fluorescence in situ hybridization
FM	Foundation Medicine
FMI	Foundation Medicine, Inc.
GP	general practitioner
HCRU	health care resource utilisation
HR	hazard ratio
HRG	Healthcare Resource Group
HS	health state
HTA	Health Technology Assessment
IC	intracranial
ICER	incremental cost-effectiveness ratio
IHC	immunohistochemistry
IRC	independent review committee
IV	intravenous
KM	Kaplan-Meier
LYG	life years gained
MASC	mammary analogue secretory carcinoma
MHRA	Medicines & Healthcare products Regulatory Agency
MIBG	metaiodobenzylguanidine
MSI-H	microsatellite instability-high
MTD	maximum tolerated dose
NE	not estimable
NGS	next generation sequencing
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NNS	number needed-to-screen
NSCLC	non-small cell lung cancer
NTRK	neurotrophic tropomyosin receptor kinase
ORR	overall response rate
OS	overall survival
PAS	Patient Access Scheme
PD	progressive disease

PFS	progression-free survival
PIM	Promising Innovative Medicine
PK	pharmacokinetic
PPS	post-progression survival
PR	partial response
PRO	patient reported outcome
PS	performance status
PSA	probabilistic sensitivity analysis
PSS	personal social services
PSSRU	Personal Social Services Research Unit
PT	preferred term
QALY	quality adjusted life years
QD	Once a day
QLQ	Quality of Life Questionnaire
RANO	Radiographic Assessment in Neurooncology
RCT	randomised controlled trial
RECIST	Response Evaluation Criteria for Solid Tumors
RNA	ribonucleic acid
RP2D	recommended phase 2 dose
RT-PCR	Reverse transcription polymerase chain reaction
SACT	Systemic Anti-Cancer Therapy
SAE	serious adverse event
SD	standard deviation
SE	standard error
SLD	sum of longest diameters
SLR	systematic literature review
SOC	standard of care
SSAR	Scottish Science Advisory Council
STS	soft tissue sarcoma
TA	Technology Appraisal
TCGA	The Cancer Genome Atlas
TKI	tyrosine kinase inhibitor
TTOT	time to off treatment
TTR	time in therapeutic range
UK	United Kingdom
USA	United States of America

WBRT	whole-brain radiotherapy

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

This submission covers the technology's full marketing authorisation for the proposed indication (Table 1).

Entrectinib will be one of the first tumour-agnostic therapies to be appraised in a Health Technology Assessment (HTA) by the National Institute of Health and Care Excellence (NICE). Given the differences to standard appraisals, close collaboration and flexibility will be required between the National Health Service (NHS), academia, industry and the public to enable incorporation into clinical practice (1). In light of these challenges, Roche Products Ltd. has made a number of assumptions for this appraisal.

There are currently no treatment options available in the NHS that specifically target solid tumours with neurotrophic tyrosine receptor kinase (NTRK)-fusions, therefore the costs and outcomes for an 'average' chemotherapy comparator is estimated, with scenario analyses to test changes in the costs or outcomes of this comparator on the cost-effectiveness of entrectinib. The 'average' chemotherapy is defined as an equally weighted mix of costs and outcomes of treatment options for a given tumour type at a given line of therapy.

Although Roche's preference is to work with the NHS to support the implementation of genomic screening, guidance regarding the process of introducing and reimbursing a new target within the NHS Genomic Testing Directory has not yet been established. For the purposes of including screening costs within the model, a hierarchical approach is therefore proposed where immunohistochemistry testing is conducted followed by confirmatory screening with a next-generation sequencing panel. However, Roche wishes to draw the attention of the committee to the following:

- The utility of next-generation sequencing (NGS) spans far beyond the identification of a single rare genomic aberration, e.g. patients may be identified for a clinical trial or alternative medicine or information may be obtained that may lead to further health benefits or cost efficiencies for the health care system (2-4). With current methodologies, it is not possible to capture this extra benefit to the healthcare system in this assessment.

- The assessment of cost-effectiveness of NGS for cancer is complex (4); the costs and outcomes of genomic screening are likely to skew the drivers of this analysis towards those related to the screening process, effectively moving the focus away from the assessment of entrectinib and the benefit of identifying the entrectinib-eligible population
- Screening costs will vary by tumour and may disadvantage patients with tumours where fewer/no reflex tests are used in clinical practice, due to significantly higher incremental screening costs in these tumours
  - Given the commitment from the NHS to implement genomic testing, *NTRK* is likely to be routinely screened and it would therefore be simple to include this as a routine part of the test directory; this commitment also means that it is likely that screening costs will decrease significantly in the near future, potentially disadvantaging first-in-class molecules like entrectinib
  - Roche is keen to work in partnership with the NHS and other relevant stakeholders to support the implementation of screening and the introduction of new biomarkers within the Genomic Test Directory

The clinical efficacy and safety evidence will be presented from a pooled analysis of four studies in the clinical development programme (n=54 and n=355, respectively). Roche acknowledges the uncertainty of the current clinical efficacy data, compared to traditional submissions (with randomised controlled trial data and for a single tumour) and wish to collaborate to address the complexities of assessing the integrated analysis data.

Taken together, Roche is therefore proactively positioning entrectinib for funding *via* the Cancer Drugs Fund (CDF), with discussions currently ongoing to establish a data collection programme that will build on the clinical efficacy data seen to date, alongside a commercial access agreement.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with *NTRK* fusion-positive locally advanced or metastatic solid tumours who: • have progressed following prior therapies • have no acceptable standard therapies	People with *NTRK* fusion-positive locally advanced or metastatic solid tumours who: • have progressed following prior therapies • have no acceptable standard therapies The population addressed in the submission is limited to the tumour histologies represented in the entrectinib clinical trials.	Economic analysis may only be conducted on the entrectinib trial population which is limited to 10 tumour types. In clinical practice, *NTRK* gene fusions may be present in additional tumour types and histologies.
Intervention	Entrectinib	Entrectinib	As per final NICE scope
Comparator(s)	Established management without entrectinib.	Established management without entrectinib, as defined by NICE-recommended therapies for the tumour-types represented in the trial population, at the position in the treatment pathway that entrectinib is anticipated to occupy in accordance with its anticipated licence.	Comparators for a tumour-agnostic indication for a product that may be used in multiple different lines of therapy are difficult to define. A pragmatic approach has been taken to decide on a line or lines of therapy by tumour histology, and resultant comparators have been selected in accordance with current NICE recommendations. Where possible, the choice of comparators has been validated by consultation with clinicians specialising in the given tumour histology.
Outcomes	The outcome measures to be considered include: • overall survival • progression free survival • response rate • duration of response • adverse effects of treatment	The outcome measures to be considered include: • overall survival • progression free survival • response rate • duration of response • adverse effects of treatment	As per NICE final scope and in line with NICE reference case.

	• health-related quality of life	• health-related quality of life	
Subgroups to be considered	If evidence allows, subgroup analyses by: • tumour site • previous therapy Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	Clinical data for the following subgroups are presented in this submission, though no economic analysis has been conducted on those groups: • Objective response rate by age, sex, ECOG performance status, etc. • Systemic efficacy by CNS status • Intracranial efficacy by CNS status	Due to the limited evidence base available for entrectinib, no meaningful subgroup analyses can be performed.

NICE: National Institute of Health and Care Excellence; NTRK: neurotrophic tyrosine receptor kinase

B.1.2 Description of the technology being appraised

Appendix C contains the summary of product characteristics for entrectinib.

Table 2: Technology being appraised

UK approved name and brand name	Entrectinib
Mechanism of action	Entrectinib is a CNS-active, potent inhibitor of tropomyosin receptor kinases A, B, and C (abbreviated as TrkA, TrkB, and TrkC), as well as anaplastic lymphoma kinase (ALK) and ROS proto-oncogene 1 receptor tyrosine kinase (ROS1) (5). Its ability to cross the blood-brain barrier has been demonstrated clinically (5). TrkA, TrkB, and TrkC are encoded by the genes NTRK1, NTRK2, and NTRK3, respectively and they are cell surface receptors expressed in neuronal tissues, where they play a critical role in the development of central and peripheral nervous systems (6, 7). Gene fusions, where the 3′ portion of the *NTRK* gene containing the catalytic tyrosine kinase domain, achieves an in-frame fusion to the 5′ portion of a partner gene that drives gene expression, may result in the constitutive activation or overexpression of Trk receptors, leading to downstream cell growth and proliferative pathways and oncogenesis (8, 9). To date, multiple fusion partners have been identified in NTRK1/2/3-rearranged tumours (6). Preclinical studies have shown that entrectinib selectively inhibits proliferative activity of cells expressing NTRK fusion proteins and can cause cell cycle arrest and apoptosis in these cells (10). This anti-proliferative activity correlates with inhibition of TrkA, TrkB, TrkC, ROS1, and ALK phosphorylation as well as the phosphorylation of key downstream mediators of the TRK signalling pathways and ALK signalling pathways (5, 11, 12).
Marketing authorisation/CE mark status	An application for marketing authorisation for entrectinib for *NTRK* fusion-positive locally advanced or metastatic solid tumours was made on [redacted]. Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated [redacted], with regulatory conditional approval expected in [redacted].
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated indication is as follows: • Entrectinib as [redacted] (11)

	• Entrectinib [redacted] (11) Entrectinib will only be contraindicated to people who demonstrate hypersensitivity to the medicinal product or any of its excipients.
Method of administration and dosage	The recommended dose of entrectinib for adults is 600 mg given orally, once daily. The recommended dose of entrectinib for paediatric patients who have the ability to swallow whole capsules is 300 mg/m^2 orally, once daily (11).
Additional tests or investigations	A validated assay is required for the selection of patients with *NTRK* fusion-positive locally advanced or metastatic solid tumours. *NTRK* fusion-positive status must be established prior to initiation of entrectinib therapy.
List price and average cost of a course of treatment	Proposed list price: £5,160.00 Average cost of a course of treatment (net): [redacted]
Patient access scheme (if applicable)	Commercial access agreement for duration of CDF: [redacted]

ALK, Anaplastic lymphoma kinase; CDF, Cancer Drugs Fund; CHMP, Committee for Medicinal Products for Human Use; CNS, central nervous system; NTRK, neurotrophic tyrosine receptor kinase; ROS1, ROS proto-oncogene 1 receptor tyrosine kinase; SmPC, summary of product characteristics; Trk, tropomyosin receptor kinase

B.1.3 Health condition and position of the technology in the treatment pathway

- **Precision medicine provides an opportunity to provide patients with better management of their health with timely diagnoses and personalised therapies**
- **Precision medicine is an integral part of the NHS Five Year Forward View and the 100,000 Genome Project has helped make progress on this vision. The Genomic Medicine Service was launched in October 2018 and a key part of this is the National Genomic Test Directory which provides the direction needed to embed genomic testing into clinical practice**
- ***NTRK* gene fusions are oncogenic drivers. They are found in a wide variety of cancers including NSCLC, CRC, breast cancer, pancreatic cancer, and rare tumour types such as sarcoma, papillary thyroid cancer, MASC**
- **A study found that patients *NTRK* gene fusions had poorer median OS when compared to patients without *NTRK*, *ALK*, or *ROS1* genomic alterations in their tumours**
- **CNS metastases are common in tumour types that are associated with *NTRK* gene fusions and is associated with high disease burden, reduced life expectancy and poor quality of life**
- **Cancer negatively impacts the HRQoL of patients, their family, and caregivers**
- **There is no standard of care for patients with *NTRK* fusion-positive cancer. The majority of patients are likely to receive a form of chemotherapy as the mainstay of management, although this is generally not effective and can be associated with notable toxicity**
- **Entrectinib is an oral, CNS-active, potent, anti-cancer agent for the treatment of patients with tumours of any type that harbour *NTRK1/2/3*, *ROS1*, and *ALK* rearrangements**
- **The efficacy and safety for entrectinib has been studied in four single-arm basket studies (ALKA, STARTRK-1, STARTRK-2, STARTRK-NG)**

B.1.3.1 The need for precision medicine

Medicine has been traditionally built around clinical teams specialising in a particular organ system working back from a patient's symptoms to arrive at a diagnosis. Precision medicine provides an opportunity to move away from this approach. The National Research Council (US) definition of precision medicine is as follows (13):

"The tailoring of medical treatment to the individual characteristics of each patient…to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventative or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not."

Although the concept of precision medicine is not new, recent technological and scientific advances have increased the possibilities to better understand how individual patients may respond to specific interventions. Knowledge of genetic variants responsible for an individual drug response can also be used to create a pharmacogenomic profile to help identify optimal treatment. This approach to targeted therapy will mean that treatments can be independent of tumour type; a concept known as 'tumour-agnostic'. Precision medicine therefore can avoid costly 'trial and error' prescription methods, as this could shorten time to diagnosis, thereby identifying the most appropriate management faster as well as avoiding lines of unnecessary and toxic therapies (1, 3). A systematic literature review of cost-effectiveness studies in metastatic CRC found that 76% of the studies identified confirmed the cost-effectiveness of biomarkers and 29% of the studies were able to confirm cost-saving upon biomarker use (14).

The National Health Service (NHS) has acknowledged the need to determine how it can best embed a precision medicine approach into mainstream healthcare to ensure the best care is provided to every patient, regardless of their illness (15). Precision medicine, with current scientific and technological advances in genomics at its core, is an integral component of delivering this vision (1).

B.1.3.2 Implementation of genomic testing

Currently, patients may undergo molecular diagnostic or genomic testing to determine specific genes that will govern treatment options; however, there is variation in the approach to the commissioning and funding of tests across England and by disease, creating inequity as not all eligible patients are currently able to access appropriate testing (16). NHS England

are introducing changes address this inequity and thereby provide improvements to the care of patients with rare cancers (16).

The importance of precision medicine to the future of healthcare at the NHS was acknowledged by the subsequent launch of the Genomic Medicine Service in October 2018, which evolves the role of existing clinical genetics services and the NHS Genomic Medicine Centres to provide comprehensive and equitable access to high quality genomic testing and management, regardless of condition and or geographical location (17). A key element of the service will be the National Genomic Test Directory that will identify the most appropriate test for each clinical indication; the technology by which it should be delivered; and when whole genome sequencing would be clinically appropriate, affordable and cost-effective to provide a better outcome for patients (18).

Part of the NHS Long Term Plan is that seriously ill children who are likely to have a rare genetic disorder, children with cancer, and adults suffering from certain rare conditions or specific cancers, will be offered whole genome sequencing from 2019 (19). Furthermore, patients with other cancers will be sequenced when clinically appropriate to do so, i.e. for tumour types already included in the Genomic Test Directory, such as sarcoma and mammary analogue secretory carcinoma (MASC) (19).

The economic investment in the Genomics Medicines strategy has been made with a clear focus on the future benefits to the NHS; while initially cost-incurring, it will become significantly more cost-effective as yet unknown targets are identified and efficiencies in screening and diagnosis are realised (3).

B.1.3.3 Current approaches to precision medicine and the potential for a tumour-agnostic medicine

There has been a dramatic increase in the use of precision medicines over recent years, particularly with the use of biomarkers to stratify patients that has led to targeted medicines being widely used in the NHS for many types of cancer (20). For example, NSCLC patients are currently tested for epidermal growth factor receptor (*EGFR*) mutations, *ALK* rearrangements, PD-L1 expression and *ROS1* rearrangements (21). Since many biomarkers occur in multiple tumour types, numerous new medicines are being marketed for multiple indications, each requiring separate regulatory submissions and health technology assessments. However, the emergence of precision medicine has led to a need to look at these medicines in a histology-agnostic manner and this approach will require collaboration between government, regulators, pharmaceutical and biotechnology industries, academia

and patient groups (22). Tumour-agnostic precision medicines that target genomic alterations rather than the tumour type/site will allow for quicker, equitable access to targeted therapies for some patients with rare cancers who would have previously been ineligible, thereby improving patient outcomes.

Evidence has been emerging in recent years on the benefit and potential value of precision medicine. A meta-analysis of 570 phase II, single-agent studies (including a total of 32,149 patients) in diverse cancer types observed that a personalised targeted treatment strategy was a key independent predictor of both improved outcomes and fewer deaths from treatment toxicity (23). A similar meta-analysis of phase I clinical trials also suggests personalised strategies employing a biomarker-based selection of patients to inform targeted treatment is associated with significantly better outcomes than a non-personalised strategy (24).

Two tumour-agnostic medicines have recently been approved by the FDA; in May 2017, the FDA granted accelerated approval for pembrolizumab for the treatment of adult and paediatric patients with unresectable or metastatic solid tumours that have been identified as having a biomarker referred to as microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR) (25), this was followed in November 2018 by the approval of larotrectinib for treating patients with solid tumours that have tested positive for *NTRK* genes (26). Other tumour-agnostic approaches currently in development, are summarised in Table 3. These new developments in oncology will address unmet needs, increase treatment options for patients, lower toxic side effects through avoidance of chemotherapy, and ultimately improve outcomes and quality of life for patients.

Table 3: Tumour-agnostic medicines which have been approved or are in development (27, 28)

Molecular target	FDA approved	Product in development
MSI-H (MMR-deficient solid tumours)	Pembrolizumab	Pembrolizumab LY3300054
NTRK fusions	Larotrectinib	Entrectinib* Merestinib TPX-0005 LOXO-195
RET fusions	-	RXDX-105 LOXO-292 BLU-667
FAP-high tumours	-	FAP-IL2v + atezolizumab
Mutant *BRAF/wtCRAF*	-	PLX8394
KIT mutations	-	PLX9486

NRG1-rearrangement	-	Anti-ERBB3

* Entrectinib granted priority review by the FDA (29)

B.1.3.4 *NTRK* gene fusions – a precision medicine target

The tropomyosin receptor kinase (TRK) family includes Trk A, B and C which are encoded by the neurotrophic tyrosine kinase (*NTRK*) receptor genes 1, 2 and 3, respectively (30). They are expressed in neuronal tissues, where they play a critical role in the development and function of neurons of the central and peripheral nervous systems, as well as a variety of non-neuronal tissues throughout development, including the cardiovascular, endocrine, reproductive, and immune systems (31). Gene fusions involving *NTRK1/2/3* (when the 3' region of the *NTRK* gene is joined with a 5' sequence of a fusion partner gene) result in a constitutive activation or overexpression of Trk receptors, potentially leading to oncogenesis (9); multiple fusion partners have been identified in *NTRK1/2/3*-rearranged tumours to date (30).

NTRK fusions require confirmation through genomic screening (e.g. Next Generation Sequencing [NGS]); they can be highly prevalent in rare tumour types (e.g. MASC, congenital fibrosarcoma) but less prevalent in more common solid tumour types (e.g. NSCLC, sarcomas), as summarised by prevalence data from various literature (Table 4) (32). Moreover, the presence of these genomic alterations tend to be mutually exclusive of other genomic aberrations, meaning that the *NTRK* fusion-positive population may not overlap with other known molecular targets (e.g. *ALK, ROS1, BRCA*) (33).

Table 4: Oncogenic TRK-fusions found across multiple tumour types (30)

Tumour type, frequency % (citation)	*NTRK1*	*NTRK2*	*NTRK3*
High prevalence of NTRK fusion			
Secretory carcinoma of salivary gland (MASC)			91–100% (34, 35)
Breast (secretory)			92% (36)
Mesoblastic nephroma			75–83% (37, 38)
Congenital fibrosarcoma			91–100% (38, 39)
Non-brainstem high-grade glioblastoma (paediatric)		40% (40)	
Melanoma (spitzoid)	21% (41)		
Papillary thyroid	<12% (42)		2–21% (43-45)
Low prevalence of NTRK fusion			
NSCLC	3% (46)	<1–3% (32, 47, 48)	
Head and neck cancer		<1% (47)	<1% (47)
Sarcoma	<1% (47)		
Colorectal cancer	<2% (49)	<1% (47, 49)	<1% (47)
Neuroendocrine tumour			<1% (50)

Glioblastoma (adult)	1% (51, 52)	1% (53)	<1% (53)
Low-grade gliomas		<1% (47)	
Pilocytic astrocytoma		3% (37)	
Infantile myofibroblastic tumour			3% (54)
Cholangiocarcinoma	4% (55)		
Acute myeloid leukaemia			<1% (56)
Thyroid carcinoma			<2% (47)
Skin cutaneous melanoma			<1% (47)
Gynaecological cancer	<2% (9)		<1% (9)
Pancreatic adenocarcinoma	1% (57)	<1% (57)	<2% (57)

MASC, mammary analogue secretory carcinoma; NSCLC; non-small cell lung cancer

Given the rarity of *NTRK* fusion genes and variation in testing methodologies, the exact frequency of *NTRK* fusion genes in solid tumours is not clear. To complement the table of figures presented in Table 4, some real-world evidence are also presented below.

Based on NGS profiling of 116,398 adult and paediatric tumour samples using the Foundation Medicine Inc. (FMI) NGS platform, an estimated prevalence of [redacted] has been observed (58). Overall, this is consistent with estimates of the prevalence of NTRK fusions by genomic profiling using high-throughput NGS on tumours from a large and broad cohort of cancer patients (0.25% [MSK IMPACT assay] (59)) and also specifically for paediatric/adolescent patients (0.44% (60); 0.49% (61)).

In summary, the overall frequency of *NTRK* fusions using different genomic NGS platforms and datasets is estimated to be in the range of [redacted] across all tumour types (Figure 1).

Figure 1: Estimated overall NTRK-positivity rate across solid tumour types (47, 58, 59, 62)

* This population may be enriched based on local pre-screening testing

ASCO, American Society of Clinical Oncology; CLIA, clinical laboratory improvement amendments; DNA, deoxyribonucleic acid; FM, Foundation Medicine; MSK, Memorial Sloan Kettering Cancer Center; NGS, next generation sequencing; PCR, polymerase chain reaction; RNA, ribonucleic acid; TCGA, The Cancer Genome Atlas

B.1.3.5 Burden of disease

There are limited data for the prognosis of patients with *NTRK*-fusion positive tumours, although Roche is currently carrying out a systematic review on this topic. Initial evidence suggests a worse outlook compared with patients without this genomic alteration. For instance, a study to compare 27 metastatic colorectal cancer (mCRC) patients bearing *NTRK/ALK/ROS1* rearranged tumours with 319 mCRC patients not bearing these rearrangements, found that those with *NTRK*, *ALK*, or *ROS1* rearranged tumours had a poorer median overall survival (OS) when compared with patients who had tumours without these genomic alterations, independent of tumour location (15.6 months vs 33.7 months respectively, $p<0.001$) (63). In an expression analysis of 119 patients with papillary thyroid carcinoma, cumulative survival analysis of *NTRK1* rearrangement-positive individuals demonstrated a worse outcome when compared with patients with expression of RET proto-oncogene hybrids (64).

Novel treatments for rarer tumour types with a high *NTRK* prevalence are difficult to develop due to the small population for clinical trials (65). As a result, treatment for rare, advanced cancers is often limited to standard chemotherapy which may be associated with significant toxicity.

Furthermore, CNS metastases are common in tumour types with a known prevalence of *NTRK* fusion genes; 10–20% of patients with advanced NSCLC (66), 2.5–23% of patients with mCRC (67) and 29% of patients with triple-negative breast cancer have all been shown to have brain metastases (68). The presence of CNS metastases is associated with a high disease burden, reduced life expectancy and poorer quality of life compared with other sites of metastases (69). The median survival of untreated patients with CNS metastasis is poor at less than 2 months, while active treatment may only extend this to 4–6 months (70).

Treatment for CNS metastasis is often limited by the ability of a drug to cross the blood-brain barrier (BBB) (66). Only a small number of targeted therapies are able to cross the BBB, for example, alectinib (Alecensa®) and gefitinib (Iressa®) (71, 72). Therefore, neurosurgery and radiotherapy are usually the primary treatment options for CNS metastasis, although both have significant treatment-associated risks (73). Despite widespread use, there has been rising concern about treatment-related toxicities with whole brain radiotherapy (WBRT), specifically neurocognitive toxicity (74).

Many of the current treatment regimens for patients with tumours harbouring *NTRK* fusions involve cytotoxic chemotherapy which is associated with significant toxicity, negatively

impacting HRQoL (75). Furthermore, psychological health can be profoundly affected by chemotherapy because of body image related issues caused by chemotherapy-induced hair loss and the lack of social activity due to the physical impact of the chemotherapy schedule (76, 77)

Several reviews have shown a greater prevalence of psychiatric disorders, in particular anxiety and depression, amongst caregivers of patients with cancer. Around 50% of caregivers of patients with advanced cancer show signs of emotional distress including depression, anxiety, insomnia, and decreased QoL (78).

A cancer diagnosis can also have significant financial impacts on both patients and their families; financial difficulties are a strong predictor of a poor QoL amongst cancer patients (79). A UK research survey by Macmillan Cancer Support found that 83% of those affected by cancer experienced negative financial repercussions, with the biggest impact coming from lost earnings due to stopping work permanently or temporarily (80).

B.1.3.6 Entrectinib: a tumour-agnostic precision medicine for patients with NTRK fusion-positive locally advanced or metastatic solid tumours

There is currently no biomarker-driven treatment available for patients with *NTRK* fusion-positive tumours in Europe. The discovery of oncogenic *NTRK* molecular alterations in various tumour types has led to the development of targeted therapies with the potential to provide patients with treatment that is both well-tolerated and effective (81). Larotrectinib was FDA-approved in the US for *NTRK* fusion-positive tumours in November 2018 (82, 83).

Entrectinib is an oral, CNS-active, potent, anti-cancer agent for the treatment of patients with tumours that harbour *NTRK1/2/3, ROS1*, and *ALK* rearrangements (10). Entrectinib has been shown to penetrate the BBB in multiple preclinical models as well as demonstrate potent anti-tumour activity in three TRK-driven intracranial tumour models, (11, 84).

The efficacy and safety for entrectinib has been studied in four single-arm basket studies (ALKA, STARTRK-1, STARTRK-2, STARTRK-NG) that have grouped patients depending on tumour genotypes; clinical data from these studies are presented in Section B.2.6.

B.1.3.7 Clinical evidence to support the NICE HTA submission for entrectinib in NTRK fusion-positive tumours

A conventional randomised controlled trial (RCT) of a tumour-agnostic treatment such as entrectinib in individual tumour types is not possible since the small number of patients with

the specific, rare genomic alteration in each tumour type would mean trial data are not statistically robust. Additionally, establishing a universal comparator standard of care across these tumour types was not feasible. Hence, basket trials were designed to allow investigators and manufacturers to conduct tumour-agnostic studies across multiple solid tumours by grouping cancer patients by their common genomic alterations (85). External regulatory bodies (Food and Drug Administration [FDA], European Medicines Agency [EMA], European Network For Health Technology Assessment [EUnetHTA]) have recognised basket studies as an acceptable method of technology for tumour-agnostic therapies (86, 87).

Compared to a medicine targeting a single indication, there are multiple challenges when assessing a tumour-agnostic medicine using standard HTA approaches, e.g. scoping, evidence synthesis and interpretation, health-economic modelling and interpretation of cost-effectiveness. In anticipation of this, Roche has been in collaboration with various stakeholders to gain advice on the optimal route for the evaluation and reimbursement of a tumour-agnostic medicine to avoid withholding patient access to drugs for licensed indications. Roche is also keen to work in partnership with the NHS and other relevant stakeholders to support the implementation of screening and the introduction of new biomarkers within the Genomic Test Directory.

B.1.3.8 Treatment pathway

There is currently no standard treatment pathway specifically for patients with *NTRK* fusion-positive cancer. For the tumour types in which *NTRK* fusion genes are prevalent, diagnosis and staging is performed following tumour type care pathway guidelines. In general, the majority of patients are likely to receive a form of chemotherapy as the mainstay of management at this stage of treatment. As standard chemotherapy is not necessarily targeted to specific cancer types, it can have 'off-target' effects and cause damage to both normal cells and tissues (88).

Based on the anticipated marketing authorisation indication, entrectinib monotherapy will be a treatment option for [redacted] (11). The position of entrectinib in clinical pathways is likely to vary by tumour type and current available therapies; examples are provided in Table 5 to illustrate this, while Table 6 provides an overview of where entrectinib might be positioned in clinical practice for all tumour types included in the integrated efficacy analysis.

Table 5: Examples of where entrectinib might be positioned by tumour types

MASC (example of a tumour type with a high NTRK-fusion frequency and limited systemic treatment options)	Soft tissue sarcoma (STS) (example of a tumour type with a low NTRK-fusion frequency and limited systemic treatment options)	NSCLC (example of a tumour type with a low NTRK-fusion frequency and numerous systemic treatment options)
MASC is a newly recognised variant of salivary gland malignancy that has been challenging to characterise; however, the presence of the *ETV6-NTRK3* fusion gene is pathognomonic to the disease (91–100% of the disease is positive for *NTRK3*) (34, 35). While MASC appears to follow an indolent course in most patients, a minority of cases appear predisposed to distant metastasis and increased mortality following attempts at curative surgery. The standard treatment algorithm for MASC is not well defined as most studies in the literature are retrospective in nature; however, a combination of radiation therapy and surgery are common (89, 90). Various treatment regimens have displayed modest response rates with unclear survival advantages in patients with metastatic salivary gland cancer (41, 43). With limited efficacy from available systemic agents for the treatment of MASC, there remains an unmet need for more effective and targeted treatment options, therefore entrectinib may be considered a first-line systemic therapy option in this setting.	STS are rare tumours, with Cancer Research UK estimating 11,700 people in the UK being diagnosed with connective and soft tissue sarcoma between 1991 and 2010 (91). While mapping the landscape of kinase fusions in cancer, Stransky and colleagues reported that 1% of sarcomas patients are positive of *NTRK1* gene fusions (47). Patients with metastatic sarcoma have a poor prognosis and their median OS doesn't exceed 18 months (92). In the UK, olaratumab in combination with doxorubicin is available via the Cancer Drugs Fund (CDF) as an option for advanced STS in adults; however, in January 2019, it was announced that the Phase III ANNOUNCE study failed to show a benefit in OS with olaratumab+doxorubicin compared with doxorubicin alone (93). Therefore, there still remains an unmet need for effective, first-line systemic therapies in this setting.	Lung cancer is the third most common cancer in the UK, accounting for 13% of all cancer cases, with 46,700 new lung cancer cases reported every year. It is responsible for 21% of all cancer deaths in the UK, making it the most common cause of cancer death; around 35,600 people die of lung cancer in the UK every year (94). Up-regulation or overexpression of oncogenes or driver mutations such as *EGFR* mutations, *ALK* and *ROS1* fusion genes, lead to uncontrolled cell division and increased cell survival. While screening for other genetic alterations, *NTRK1* and *NTRK2* have been implicated as a driving mutation in 1–3.3% of cases (46, 47). For the minority of patients expressing *ALK/EGFR* oncogenic driver mutations (a population which may not overlap with *NTRK* fusion patients) the development of targeted led to a paradigm shift that is now well established; however, disease progression is still inevitable as tumour resistance invariably develops (95). Following progression or tolerability issues with targeted tyrosine kinase inhibitor therapies, patients are likely to subsequently receive platinum-based chemotherapy, although a retrospective study reported only modest responses with a median PFS of around 4 months for erlotinib-resistant (*EGFR* inhibitor) patients who received second-line chemotherapy (96). Therefore, there also remains a need for more efficacious treatment options for these patients.

Table 6: Proposed positioning of entrectinib for the treatment of NTRK fusion-positive, locally advanced or metastatic solid tumours

Position of entrectinib in line of systemic therapy	
First-line*	**Second-line and beyond†**
MASC	NSCLC
Soft-tissue sarcoma	Breast
Pancreatic cancer	Thyroid cancer
Cholangiocarcinoma	Colorectal cancer
Gynaecological cancers	Neuroendocrine tumours

*Patients ineligible for curative surgery or radiotherapy with no immunotherapy or targeted therapy options

†Some patients may receive first-line entrectinib treatment if not eligible for targeted or immunotherapies

B.1.4 Equality considerations

No equality issues were identified.

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

A systematic literature review (SLR) of clinical study evidence on the efficacy, safety, and HRQoL of pharmacological interventions for the treatment the *NTRK* fusion-positive advanced cancer population was conducted.

Appendix D contains the full details of the process and methods used to identify and select the relevant clinical evidence.

B.2.2 List of relevant clinical effectiveness evidence

The efficacy and safety for entrectinib has been studied in four, single-arm basket studies that have grouped patients dependent on tumour genotypes (Table 7). Entrectinib was initially investigated as a single agent in the first-in-human study ALKA-372-001 (hereafter referred to as ALKA) conducted exclusively in Italy (97) and subsequently RXDX-101-01 (hereafter referred to as STARTRK-1) conducted in the US and Korea (98). Patients were enrolled into dose-escalation cohorts using a conventional “3+3” scheme until selection of the recommended Phase II dose (RP2D), followed by cohort expansion at the RP2D. Following determination of the RP2D, and early evidence of clinical activity observed with entrectinib in the Phase I studies (ALKA and STARTRK-1), the entrectinib clinical development program was expanded with the initiation of Phase II Study RXDX-101-02

(hereafter referred to as STARTRK-2) (99). Given the occurrence of *NTRK*-fusions in paediatric tumours, the entrectinib development program was expanded to include the ongoing paediatric study RXDX-101-03 (STARTRK-NG) conducted in the US (100).

Efficacy results from three of the studies in adult patients (ALKA, STARTRK-1, STARTRK-2) have been pooled and analysed collectively (58, 101) - this integrated efficacy analysis includes data for 54 adult patients who had at least 6 months' follow-up, and forms the basis for this submission and economic analysis. Data from STARTRK-NG, a Phase I/Ib study evaluating the effect of entrectinib in children, adolescent, and young adult patients is presented in Section B.2.6.6. Patient safety data from the ALKA, STARTRK-1, STARTRK-2, and STARTRK-NG studies have been pooled and analysed collectively as the integrated safety population (n=355) (102). Data up to clinical cut-off dates (CCOD) of 31st May 2018 and [redacted] are provided.

Table 7: Clinical effectiveness evidence

Study	**ALKA-372-001 (Phase I) ongoing (97)**	**RXDX-101-01 (STARTRK-1) (Phase I) ongoing (103)**	**RXDX-101-02 (STARTRK-2) (Phase II) ongoing (99)**	**RXDX-101-03 (STARTRK-NG) (Phase I/Ib) ongoing (100)**
Study design	First-in-human, multicentre, open-label, ascending-dose study with dose escalation according to a standard 3+3 scheme.	Multicentre, open-label, ascending-dose study with dose escalation according to a standard 3+3 scheme.	Registration enabling, global, multicentre, open-label, basket study.	Multicentre, 5-part, open-label, dose escalation and expansion study.
Population	Patients (≥18 years of age) with advanced/metastatic solid tumours, including patients with *NTRK1/2/3, ROS1*, or *ALK* molecular alterations. ***NTRK* efficacy evaluable analysis set n=1** **Patients evaluable for safety n=57**	Patients (≥18 years of age) with solid tumours with *NTRK1/2/3, ROS1*, or *ALK* molecular alterations. ***NTRK* efficacy evaluable analysis set n=2** **Patients evaluable for safety n=76**	Patients (≥18 years of age) with advanced or metastatic solid tumours that harbour an *NTRK1/2/3, ROS1*, or *ALK* gene fusion, excluding *ALK*-positive NSCLC. ***NTRK* efficacy evaluable analysis set n=51** **Non-measurable disease n=1**	Children and adolescents (2 to 22 years of age) with recurrent or refractory solid tumours and primary brain tumours, including tumours carrying *NTRK1/2/3, ROS1*, and *ALK* gene fusions. **Paediatric patient n=1[d]** **Patients evaluable for safety n=16**

Study	ALKA-372-001 (Phase I) ongoing (97)	RXDX-101-01 (STARTRK-1) (Phase I) ongoing (103)	RXDX-101-02 (STARTRK-2) (Phase II) ongoing (99)	RXDX-101-03 (STARTRK-NG) (Phase I/Ib) ongoing (100)
			Patients evaluable for safety n=206	
Intervention(s)	***Entrectinib:*** Schedule A: 100, 200, 400, 800, 1200, or 1600 mg/m^2 once daily (fasted) 4-days on, 3-days off schedule x 3 weeks followed by 7-day rest in a 4-week cycle[a] Schedule B: 200, 400 mg/m^2 or 600 mg continuous once daily (fed) in a 4-week cycle[b] Schedule C: 400 or 800 mg/m^2 once daily (fed) in a continuous 4-days on, 3-days off schedule in a 4-week cycle[c]	***Entrectinib:*** 100, 200, 400 mg/m^2 or 600, 800 mg continuous once daily (fed) on 28-day (i.e., 4-week) cycles.	***Entrectinib:*** 600 mg, orally, once daily on 28-day (i.e., 4-week) cycles.	***Entrectinib:*** Orally, once daily on 28-day (i.e., 4-week) cycles. Dosing nomogram based on BSA, ranging from 250 mg/m^2 to 750 mg/m^2.
Comparator(s)	None	None	None	None
Indicate if trial supports application for marketing authorisation	Yes – as integrated efficacy analysis (N=54) and pooled safety population (N=355)			Only as part of pooled safety population (N=355)
Indicate if trial used in the economic model	Yes – as integrated efficacy analysis (N=54) and pooled safety population (N=355)			Only as part of pooled safety population (N=355)
Rationale for use/non-use in the model	The specific objectives of each individual study were different with the primary endpoint of the STARTRK-2 study being an efficacy objective (BICR-ORR) and the primary objective of the Phase I dose-escalation studies ALKA, STARTRK-1, and STARTRK-NG being safety and dose determination (determination of MTD and/or RP2D). While the integrated analyses were not prespecified in the individual study protocols, considering the rarity of the patient population, an integrated statistical analysis plan was developed to maximise the number of gene fusion-positive patients available for safety and efficacy analyses, including patients from the Phase I studies. This proposal to pool safety and efficacy from the clinical studies was endorsed by the regulatory health authorities because of the rare disease setting (58).			Efficacy data from paediatric patients in STARTRK-NG were not included in the integrated analysis because these patients were assessed by investigator and only one patient in STARTRK-NG met the requirement for efficacy-evaluable analysis with at least 6 months follow-up at the time of initial submission. The results are presented in Section B.2.6.6.
Reported outcomes specified in the decision problem	• Primary endpoints: ORR (based on BICR assessment), DOR • Secondary endpoints: PFS, OS • Adverse effects of treatment • Patient-reported outcomes			Secondary endpoints (Parts A [expansion], C, and D): • safety, ORR, DOR, and PFS in all enrolled patients Secondary endpoints (Parts B and D): • DOR

Study	ALKA-372-001 (Phase I) ongoing (97)	RXDX-101-01 (STARTRK-1) (Phase I) ongoing (103)	RXDX-101-02 (STARTRK-2) (Phase II) ongoing (99)	RXDX-101-03 (STARTRK-NG) (Phase I/Ib) ongoing (100)
All other reported outcomes	• Primary endpoints: BOR • Secondary endpoints: CBR, time to CNS progression • Secondary endpoints in patients with CNS metastases at baseline: IC-ORR, IC-DOR, IC-PFS			Primary endpoint: • MTD or RP2D (Phase 1 Part A). Secondary endpoints: • PK of entrectinib in plasma • Parts A [expansion], C, and D: TTR, CBR • Parts B and D: intracranial tumour response, TTR, CNS-PFS

ALK, anaplastic lymphoma kinase; BICR, blinded independent central review; BOR, best overall response; BSA, body surface area; CNS, central nervous system; DLT, dose limiting toxicity; DOR, duration of response; IC, intracranial; MTD, maximum tolerated dose; NSCLC, non-small cell lung cancer; NTRK1/2/3, neurotrophic tyrosine receptor kinase 1/2/3; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; RP2D, recommended phase 2 dose; TTR; time in therapeutic range

[a] Terminated at 1600 mg/m^2/day because of a plateau in entrectinib exposure above 800 mg/m^2/day.

[b] Ongoing at 600 mg fixed dosing.

[c] Terminated at 800 mg/m^2/day, the highest dose evaluated.

[d] In addition, 4 paediatric patients with tumours harbouring an *NTRK* gene fusion have been enrolled after 30 November 2017 in the expansion portion of STARTRK-NG

B.2.3 Summary of methodology of clinical effectiveness studies

B.2.3.1 ALKA, STARTRK-1, STARTRK-2, and STARTRK-NG trials

A brief summary of the methodology of individual studies included in the integrated efficacy analysis set (ALKA, STARTRK-1 and STARTRK-2) is provided in Table 8.

Table 8: Summary of methodology of the ALKA, STARTRK-1, STARTRK-2, and STARTRK-NG trials

Trial number (acronym)	ALKA-372-001 (ALKA*) (97)	RXDX-101-01 (STARTRK-1*) (103)	RXDX-101-02 (STARTRK-2) (104)	RXDX-101-03 (STARTRK-NG) (100)
Primary objective	Determine the first cycle DLTs and the MTD of entrectinib	Evaluate the safety and preliminary antitumour activity of entrectinib in adult patients with any locally advanced or metastatic solid tumour confirmed to be positive for *NTRK1/2/3*, *ROS1*, or *ALK* molecular alterations	Determine the ORR of entrectinib, as assessed by BICR, in each patient population basket of solid tumours that harbour an *NTRK1/2/3*, *ROS1*, or *ALK* gene rearrangement.	Determine the MTD or recommended RP2D of entrectinib in paediatric patients (children and adolescents) with relapsed or refractory solid tumours (Phase 1 Part A).
Methodology	Entrectinib was administered orally in three dose schedules (97): • Schedule A: 100, 200, 400, 800, 1200, or 1600 mg/m^2 once daily (fasted) 4-days on, 3-days off schedule x 3 weeks followed by 7-day rest in a 4-week cycle; • Schedule B: 200, 400 mg/m^2 or 600 mg continuous once daily (fed) in a 4-week cycle; • Schedule C: 400 or 800 mg/m^2 once daily (fed) in a continuous 4-days on, 3-days off schedule in a 4-week cycle. The dose escalation for all schedules was planned to continue until the RP2D was determined or until the study was terminated at the discretion of the sponsor. For all schedules, a conventional "3+3" patient enrolment scheme was followed during the dose escalation. Patients were treated based on tumour molecular diagnosis: patients with *ALK* positive tumours or *ALK* negative tumours with *NTRK1* or *ROS1* genetic alterations, patients with *ALK* positive tumours or *ALK* negative tumours with *NTRK1/2/3* or *ROS1* genetic alterations, and	STARTRK-1 is comprised of 2 segments, a dose escalation segment and a dose expansion segment The primary objective of the dose escalation segment of this study was to determine the first cycle DLTs, MTD, and a biologically effective RP2D of orally administered entrectinib. The primary objective of the dose expansion segment was to assess ORR, defined as the proportion of patients with CR or PR. Each cycle in the dose escalation segment consisted of treatment for 28 consecutive days in repeated 4-week cycles. A standard "3+3" patient enrolment scheme was followed with an accelerated titration design. The starting dose was 100 mg/m^2 once daily in the fed condition; dose escalation began with an accelerated phase in which the dose was doubled in successive cohorts until one patient experienced a DLT in the first cycle; or two patients experienced adverse events at least possibly related to entrectinib that were Grade ≥2 severity, but not considered to be DLTs and occurred during the first cycle, whichever came first. Once this predetermined toxicity level was met,	*NTRK1/2/3* gene rearrangements were treated as a combined *NTRK1/2/3* gene rearrangement basket. Patients were enrolled in a "non-evaluable" basket if they were not assessable for the primary endpoints of the study (e.g., had non-measurable disease, co-occurring mutations, etc.) but could contribute to assessment of safety, PK, and other secondary endpoints. Based on the findings of the Phase I clinical studies, entrectinib was administered orally on a continuous daily dosing regimen at a dose of 600 mg once-daily in repeated 4-week cycle. Patients were followed for safety and efficacy as per the schedule of assessments and remained on study treatment until documented radiographic progression as assessed by BICR, development of unacceptable toxicity, or withdrawal of consent. At the discretion of the investigator and with the sponsor's approval,	The Phase 1 (Part A) dose escalation study was conducted to determine the MTD or RP2D, PK, and safety profile of entrectinib in children, adolescents, and young adult patients with relapsed or refractory extracranial solid tumours. Entrectinib was administered orally with food, QD, in repeated 4-week cycles. The starting dose in Part A was 250 mg/m^2 on a continuous daily dosing regimen. Up to four dose levels were evaluated. A "3+3" patient enrolment scheme was followed during the dose escalation. The RP2D was planned to be determined from DLT(s) derived from clinical and laboratory observations in the first treatment cycle (28 days). The MTD was defined as the dose level immediately below the dose level at which ≥2 patients from a cohort of 3 to 6 patients experienced a DLT. After MTD was established, based on the DLT assessment and an overall acceptable safety profile at the

	patients with tumours harbouring *NTRK1/2/3, ROS1*, or *ALK* genetic alterations. For patients with no prior diagnosis of *NTRK1/2/3, ROS1*, or *ALK* positive genetic alterations before study inclusion, pre-screening informed consent was requested to permit the molecular characterisation of the patient tumours for *NTRK1/2/3, ROS1*, and *ALK* genetic alterations.	escalation was planned to be followed by a modified Fibonacci scheme (50%, 40%, or 33% increments). Patients remained on study treatment until disease progression, unacceptable toxicity, or withdrawal of consent. In cases of progressive disease, after discussion with the sponsor, the patient could have continued treatment if the investigator believed that the patient might continue to derive clinical benefit.	patients could continue treatment with entrectinib after BICR-confirmed disease progression if the patient was perceived to be deriving clinical benefit. For these patients, tumour assessments were no longer submitted for BICR, but investigators were encouraged to continue to evaluate patients following a similar 8-week schedule.	MTD, this dose was selected as the RP2D for evaluation in the Phase 1b portion of the study. Dose Expansion Phase 1b were planned to be opened simultaneously after determination of the RP2D in Dose Escalation Phase 1. Phase 1b was designed to enrol additional patients with specific tumour types and molecular alterations. All patients in Phase 1b were planned to receive entrectinib at the paediatric RP2D, except for Part E, who were to initially receive entrectinib via alternative dosing methods at the -1 dose level de-escalation from the RP2D.
Main inclusion criteria for participants	• Age ≥18 • Patients with histologically or cytologically confirmed diagnosis of advanced/metastatic solid tumours with *ALK* positive alterations or *ALK* negative patients with *NTRK1/2/3* or *ROS1* genetic alterations, for whom no alternative effective standard therapy was available, standard therapy was considered unsuitable, or had been refused treatment • ECOG performance status ≤2 • Life expectancy of at least 3 months • Baseline laboratory data indicating acceptable haematologic status, liver and renal function	• Age ≥18 • Patients with histologically or cytologically confirmed diagnosis of relapsed or refractory locally advanced or metastatic solid tumours for whom no alternative effective standard therapy was available or for whom standard therapy was considered unsuitable or intolerable • A molecular alteration in *NTRK1/2/3*, *ROS1*, or *ALK* was preferred, but not a requirement for patient eligibility in the dose escalation segment • Eligible patients for the dose expansion segment are required to have locally advanced or metastatic solid tumours harbouring the following types of molecular alterations:	• Age ≥18 • Histologically- or cytologically-confirmed diagnosis of locally advanced or metastaticsolid tumour that harbours an *NTRK1/2/3, ROS1*, or *ALK* gene rearrangement that is predicted to translate into a fusion protein with a functional TRKA/B/C, ROS1, or ALK kinase domain, respectively, without a concomitant second oncodriver (e.g., epidermal growth factor receptor, KRAS) • Patients with CNS involvement, including leptomeningeal carcinomatosis, which is	• Patients ≥2 years and <22 years of age were eligible for Part A through Part D, and patients from birth to <22 years of age were eligible for Part E • Children, adolescents, and young adult patients with relapsed or refractory extracranial solid tumours (Phase 1; Part A), with additional expansion parts (Phase 1b) in children, adolescents, and young adult patients with primary brain tumours harbouring *NTRK1/2/3*, *ROS1*, or *ALK* molecular alterations (Part B), neuroblastoma (Part C), and other non-neuroblastoma, extracranial solid tumours

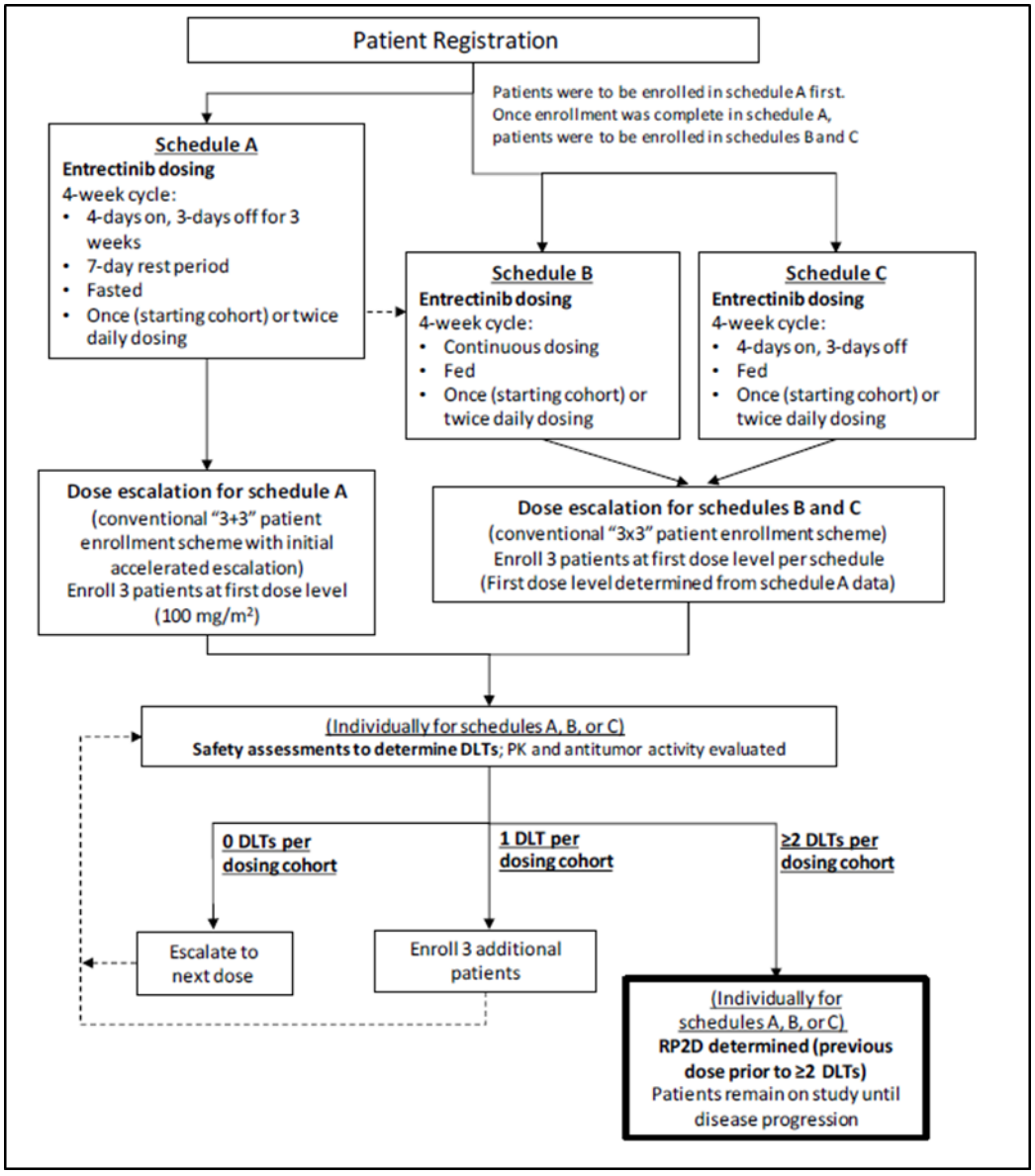
	• Resolution of any acute toxic effects (excluding alopecia) of any prior anticancer therapy • Patients with controlled asymptomatic CNS involvement, in absence of therapy with anticonvulsant or in presence of therapy with non-enzyme-inducing anti-epileptic drugs or requiring steroids at stable dose (≤4 mg/day dexamethasone or equivalent) for at least 2 weeks	○ *NTRK* fusions previously treated with other TRK inhibitors ○ *ALK* gene rearrangements with 1198 resistance single-nucleotide polymorphism ○ *ALK* alternative transcription initiation ○ *NTRK/ROS/ALK* overexpression ○ Activating splice variants ○ Other molecular alterations of interest, depending on biological rationale and after discussion with the sponsor	either asymptomatic or previously-treated and controlled, were allowed • Measurable disease as assessed locally using RECIST v1.1 • Prior anticancer therapy is allowed (excluding approved or investigational *TRK, ROS1*, or *ALK* [non-NSCLC patients only] inhibitors in patients who have tumours that harbour those respective gene rearrangements • ECOG performance status ≤2 and minimum life expectancy of at least 4 weeks	harbouring *NTRK1/2/3*, *ROS1*, or *ALK* gene fusions (Part D) • In addition, an exploratory cohort (Part E) enrols patients who were otherwise eligible but unable to swallow capsules
Settings and locations where the data were collected	2 centres in Italy	11 centres in the United States, Spain, and South Korea.	84 investigative sites in 15 countries globally, including the 3 centres in the United Kingdom.	Phase 1 (Part A): USA (8 centres) Phase Ib: Patients were enrolled at 4 of the 8 investigational sites involved in the Phase I, and 4 new sites in the US.
Tumour molecular characterisation	All patients must have tumour tissue available for central confirmation of a TrkA/B/C, ROS1, or ALK molecular alteration of interest by IHC, FISH, or NGS. These analyses will be performed at the Sponsor's CLIA laboratory in San Diego, California, USA.	Patients may be screened for the presence of molecular alterations by assays available to the given clinical site. These may include NGS, qPCR, FISH, and/or IHC. In addition, patients may also be screened at Roche's central CAP/CLIA laboratory.	For patients enrolled via local molecular testing, an archival or fresh tumour tissue (unless medically contraindicated) is required to be submitted for independent central molecular testing at Ignyta's CAP/CLIA laboratory post-enrolment. Testing for enrolment eligibility may be performed in one of two ways: 1. Tumour tissue may be submitted to Ignyta's CAP/CLIA laboratory in San Diego,	In order to determine enrolment eligibility for Parts B and D, molecular testing must be performed by a CLIA-certified or equivalently-accredited diagnostic laboratory; for detection of gene fusions, any nucleic acid-based diagnostic testing method that relies on direct assessment of gene fusions will be accepted. NGS, Sanger, RT-PCR, NanoString, and EdgeSeq are examples of acceptable methods; FISH is not an acceptable method. If potential study participants do

			California, USA, to be tested for the presence or absence of target gene rearrangements (fusions) via next generation sequencing 2. Alternatively, patient specimens may be tested locally using any nucleic acid-based diagnostic testing method that relies on direct assessment of gene rearrangements and is performed in a CLIA-certified or equivalently-accredited diagnostic laboratory. Eligible patients must have a reported gene rearrangement involving *NTRK1/2/3*, *ROS1*, or *ALK* that is predicted to translate into a fusion protein with a functional TrkA/B/C, ROS1, or ALK kinase domain, respectively. NGS, Sanger, RT-PCR, NanoString, and EdgeSeq are examples of acceptable methods; FISH is not an acceptable method.	not have access to an accepted molecular testing method to determine molecular eligibility for enrolment, sites may submit tissue for gene rearrangement screening to Foundation Medicine, Inc. in Cambridge, Massachusetts, U.S.A. For patients identified to have tumours harbouring relevant gene fusions based on local molecular testing, an archival tumour tissue from diagnosis, or preferably, from relapsed disease (preferably from the same tissue block and unless medically contraindicated) is also required to be submitted (preferably within 1 month of enrolment) for independent central molecular testing at Foundation Medicine.
Number of subjects (planned and analysed)	An overall sample size of approximately 70 treated patients was anticipated. This study report includes patients enrolled up to and including 30th Nov 2017 with a clinical data cutoff date of 31 May 2018. As of 30th Nov 2017, 58 patients were enrolled at 2 investigative sites; 57 received study drug treatment. Patient enrolment completed on 20th Mar 2018. The study is ongoing as of this report with 2 of 57 patients still receiving treatment.	At least 15 patients were anticipated to enrol into the dose escalation segment of the study. The actual number of patients enrolled was 76. As of the enrolment cut-off of November 30th 2017, no patients had been enrolled in the ongoing dose expansion segment.	A total of 207 patients were enrolled and 206 patients were treated with entrectinib (received at least one dose); 63 patients were enrolled in the *NTRK* population, 105 in the *ROS1* NSCLC population, and 38 patients in the other population basket.	**Phase 1 (Part A)** Planned: approximately 6–30 patients Enrolled: 16 patients **Phase 1b (Part B and D)** Planned: approximately 13 patients per basket (i.e., tumour type and molecular alteration combination) for the first stage. Up to an additional 49 patients into the second stage.

* STARTRK-1 and ALKA were concurrent studies and interdependent of each other in that dose escalation decisions in one study affected the conduct of the other.
BICR, blinded independent central review; CAP, College of American Pathologists; CLIA, Clinical Laboratory Improvement Amendments; CR, complete response; DLT, dose-limiting toxicities; ECOG, Eastern Cooperative Oncology Group; MTD, maximum tolerated dose; ORR, objective response rate; PR, partial response; QD, once a day; RP2D, recommended Phase 2 dose

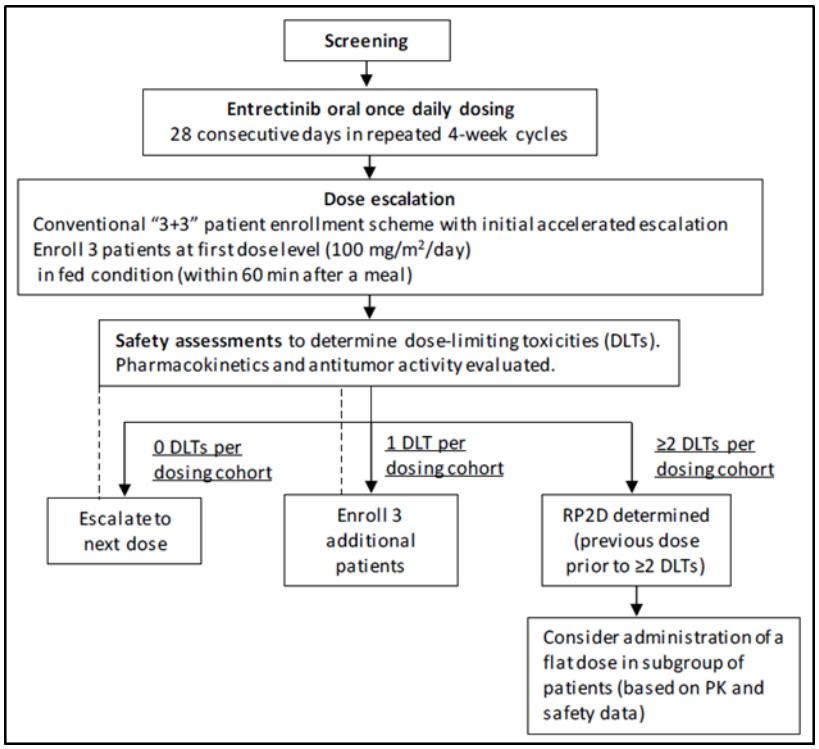
The study schema for the ALKA, STARTRK-1, STARTRK-2, and STARTRK-NG trials are presented in Figure 2, Figure 3, Figure 4, and Figure 5.

Figure 2: ALKA study design (97)



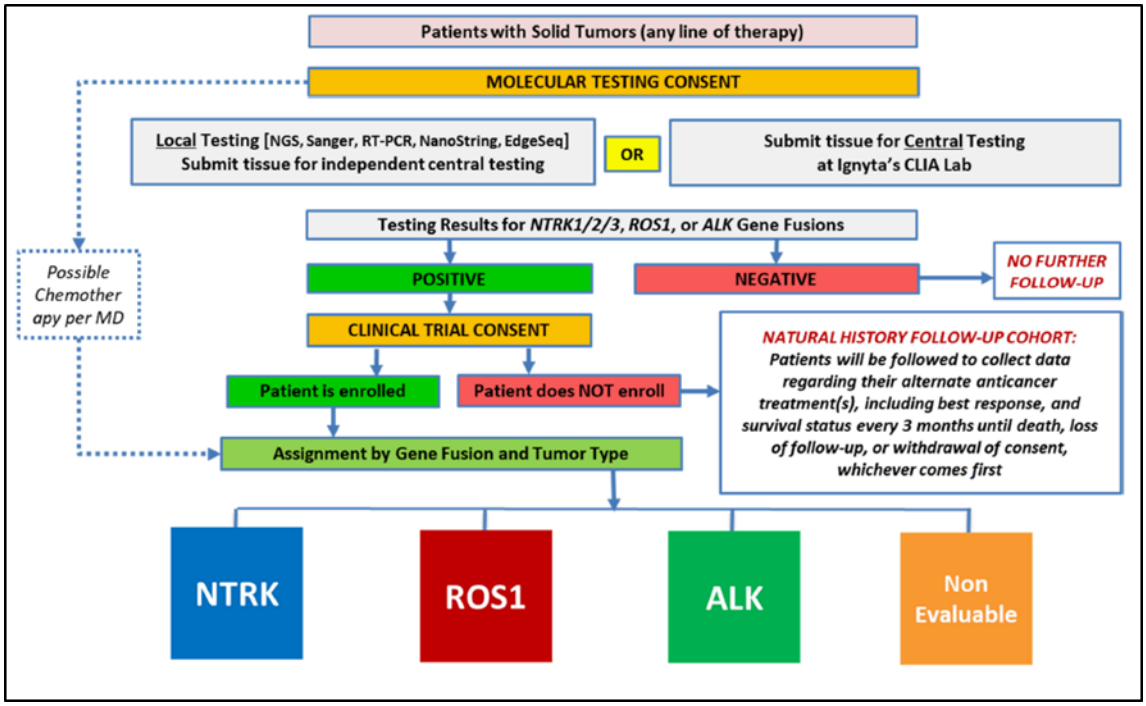
DLT, dose limiting toxitities

Figure 3: STARTRK-1 dose escalation study design (103)



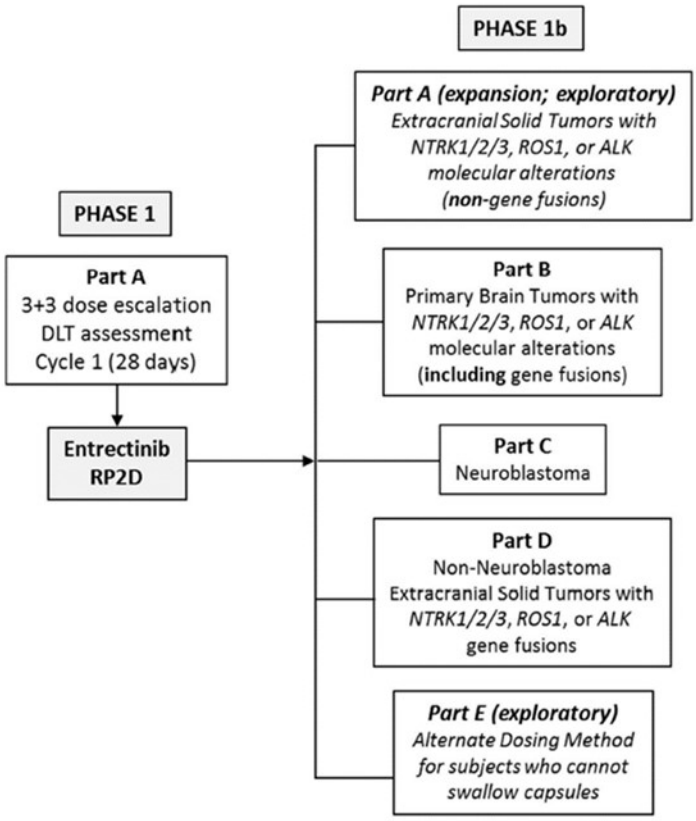
PK, pharmacokinetics; RP2D, recommended phase 2 dose

Figure 4: STARTRK-2 study design (99)



CLIA, Clinical Laboratory Improvement Amendments; NGS, next generation sequencing; RT-PCR, reverse transcription polymerase chain reaction

Figure 5: STARTRK-NG Study Design (104)



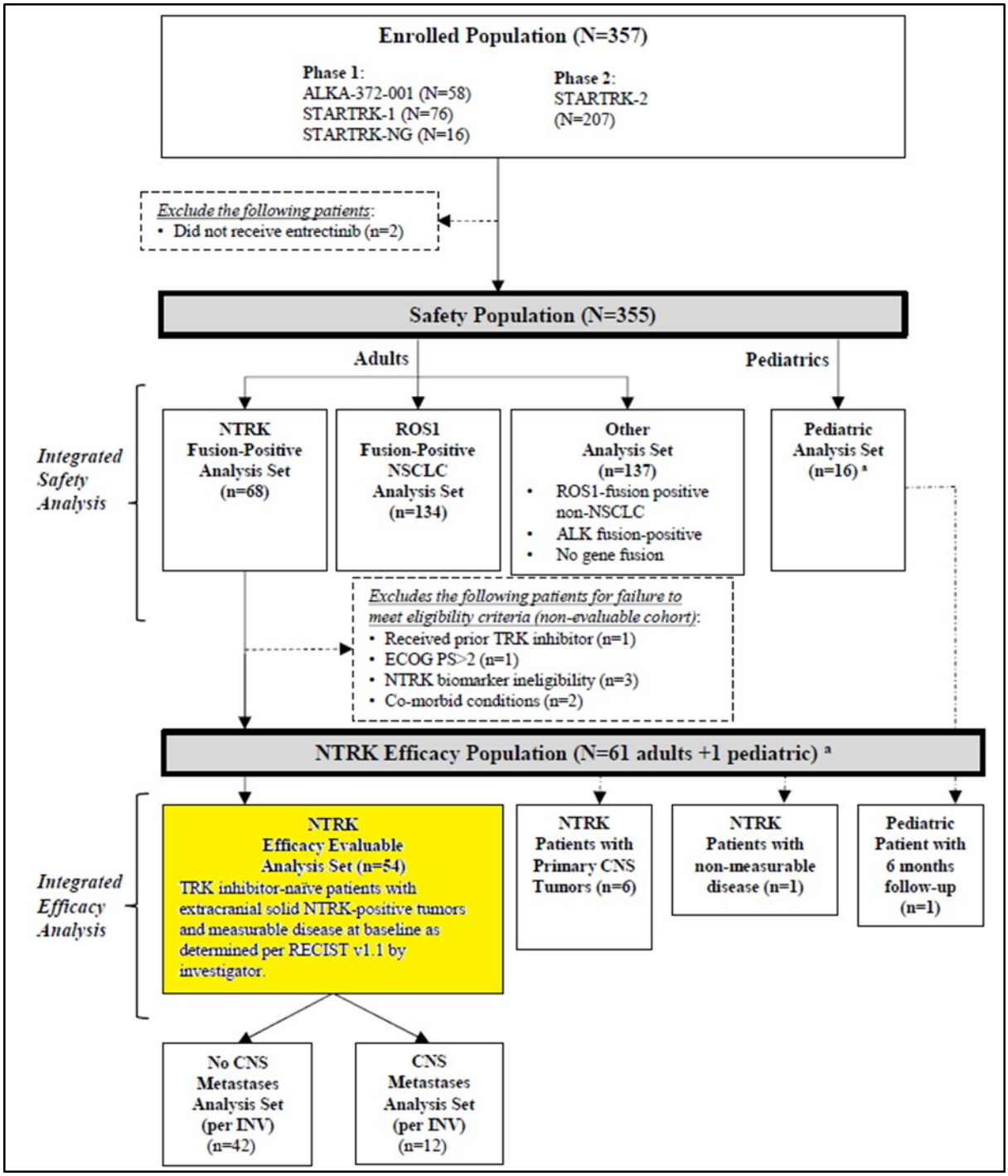
DLT, dose limiting toxicities; RP2D, recommended phase 2 dose

B.2.3.2 Integrated efficacy evaluable analysis

An integrated efficacy evaluable analysis was performed using data from ALKA, STARTRK-1, and STARTRK-2. The CCOD for all the efficacy evaluable analyses was May 31st 2018 which was based on a combined sample size of 54 adult patients (efficacy evaluable patients) with at least 6 months' follow-up enrolled into entrectinib studies up to November 30th 2017 (101).

A summary of the algorithm defining inclusion of enrolled patients in the *NTRK* efficacy evaluable population and subsets for the integrated analyses, and the overall disposition of enrolled patients (as of November 30th 2017) across the four clinical studies within these groups is shown in Figure 6.

Figure 6: Patient populations and analysis sets for patients with NTRK fusion-positive solid tumours (58)



[a] In addition, 4 paediatric patients with tumours harbouring an *NTRK* gene fusion have been enrolled after 30 November 2017 in the expansion portion of STARTRK-NG.

CNS, central nervous system; ECOG-PS, Eastern Cooperative Oncology Group performance status; INV, investigator; NSCLC, non-small cell lung cancer

The *NTRK* efficacy population (n=61 adults and n=1 paediatric) includes patients who met all of the following criteria and had at least 6 months' follow-up:

- Had tumours that harbour an *NTRK* gene fusion
- Received at least 1 dose of entrectinib
- Had not been previously treated with a TRK inhibitor

The *NTRK* efficacy population consisted of the following two mutually exclusive subgroups:

- ***NTRK* efficacy evaluable analysis set (n=54 adults):** TRK inhibitor-naïve patients with extracranial *NTRK* fusion-positive solid tumours and measurable disease at baseline as determined by the investigator using RECIST v1.1. The *NTRK* Efficacy Evaluable Analysis Set included two subgroup analysis sets based on the presence or absence of CNS metastases at baseline (i.e., no CNS metastases and CNS metastases analysis sets) as determined by investigator.
- ***NTRK* efficacy non-evaluable analysis set (n=7 adults + n=1 paediatric):** All other patients not included in the *NTRK* Efficacy Evaluable Analysis Set, including any patient enrolled with a primary CNS tumour and patients with non-measurable disease at baseline as assessed by the investigator. The single paediatric patient who had at least 6 months' follow-up was not included in the integrated efficacy analysis as this analysis only includes efficacy data from the three adult studies: ALKA, STARTRK-1, and STARTRK-2. A total of 6 adult patients with primary CNS tumours were excluded from the integrated efficacy analysis because these patients were assessed in the studies using RANO criteria, as is standard in clinical trial practice, rather than RECIST v1.1. Clinical status and corticosteroid use were not considered when determining RANO overall response by the BICR. One out of the 6 patients was a responder with DOR of 2.79 months and PFS of 6.34 months.

B.2.3.3 Integrated efficacy evaluable analysis: demographics and baseline characteristics

Among the 54 patients (59.3% female, 40.7% male) in the integrated efficacy evaluable analysis, the median age was 57.5 years (range: 21 to 83 years) at the time of enrolment (Table 9). By age group, most patients (63.0%) were <65 years old and 37.0% were elderly (≥65 years old). The majority of patients were white (79.6%) and 13.0% were Asian. Most

patients had an ECOG PS score of 0 (42.6%) or 1 (46.3%), and 11.1% had an ECOG PS score of 2 (58).

Table 9: Demographics and baseline characteristics (efficacy evaluable analysis) (58, 105)

	***NTRK* efficacy evaluable analysis set N=54**
Sex, n (%)	
Male	22 (40.7)
Female	32 (59.3)
Median age, years (range)	57.5 (21–83)
Age group, years, n (%)	
<65	34 (63.0)
≥65	20 (37.0)
Race, n (%)	
Asian	7 (13.0)
White	43 (79.6)
Not reported	4 (7.4)
Mean BSA, m^2 (SD)	1.85 (0.26)
Mean BMI, kg/m^2 (SD)	25.68 (5.30)
ECOG PS, n (%)	
0	23 (42.6)
1	25 (46.3)
2	6 (11.1)
History of smoking, n (%)	**(N=53)**
No	30 (56.6)
Yes	23 (43.4)
Gene fusion detected, n (%)	
NTRK1	22 (40.7)
NTRK2	1 (1.9)
NTRK 3	31 (57.4)
Median time since diagnosis, months (range)	21.4 (2.1–433.1)
Disease stage at initial diagnosis, n (%)	
0, I or II (A/B)	15 (28.3)[a]
III (A/B/C) or IV	33 (62.3)[a]
Unknown	5 (9.4)[a]
Metastatic disease, n (%)	
Any site	52 (96.3)
Brain metastases	12 (22.2)[b]
No. of lines of therapy since metastatic disease[c], n (%)	
0	20 (37.0)
1	11 (20.4)
2	14 (25.9)
3	4 (7.4)
≥4	5 (9.3)

Previous therapy, n (%)	
Any systemic therapy[d]	48 (88.9)
Surgery	43 (79.6)
Radiotherapy	36 (66.7)[e]

BMI, body mass index; BSA, body surface area; ECOG PS, Eastern Cooperative Oncology Group performance status; NE, not estimated

[a] Percentages calculated based on denominator of 53 patients as one patient in the ALKA study for whom the initial diagnosis field on the Case Report Form was blank was excluded.

[b] Includes two patients with measurable disease.

[c] Patients may have received other therapies in the adjuvant or neo-adjuvant setting that are not included as a line of therapy from the time of metastatic disease diagnosis.

[d] Includes chemotherapy, immunotherapy, targeted therapy or hormonal therapy.

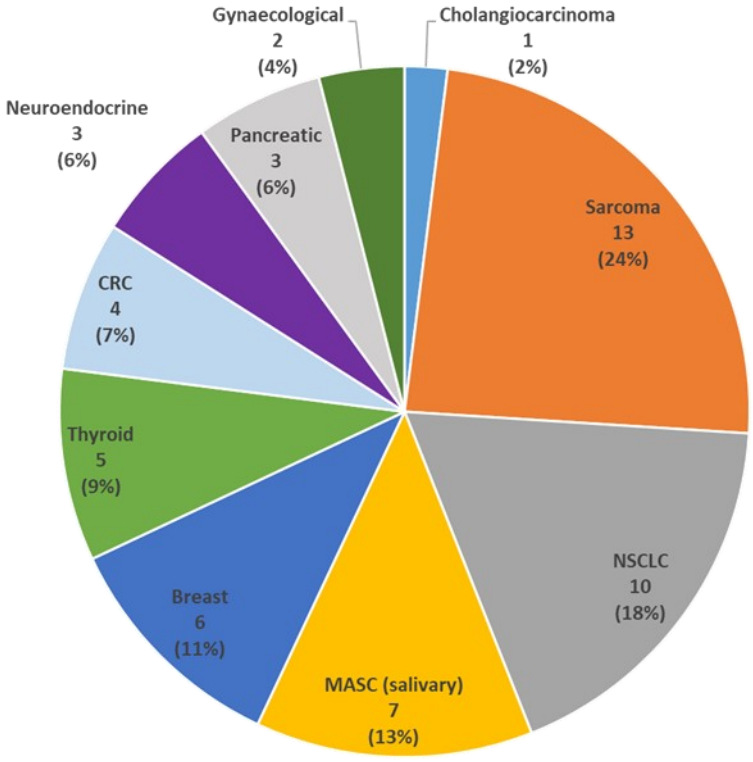
[e] Includes 7 patients who received prior radiotherapy of the brain.

Patient disease history was documented in the case report form and included both the cancer diagnosis and tumour histology. Prior to the integrated analyses, each patient was designated to a standardised tumour type per their unique diagnosis and histology data as reported in the case report form. Tumour types were classified according to high-level and low-level set of categories (105).

The most frequently represented solid tumour types (high-level category) in the efficacy evaluable analysis were sarcomas (24.1%), NSCLC (18.5%), salivary gland tumours (mammary analogue secretory carcinoma) (13.0%), and breast cancer (11.1%), which collectively accounted for approximately half of patients in the analysis (Figure 7). The low patient numbers for each tumour type has driven the basket trial approach and integrated analysis. The majority of patients (96.3%) presented with metastatic disease at baseline, of which the most common sites were lung (61.1%) and lymph nodes (55.6%) (105).

Tumours harbouring gene fusions of each of the *NTRK* genes were represented. Over half of the patients (57.4%) had *NTRK3* fusions (with 6 different fusion partners). *ETV6-NTRK3* was the most frequently represented fusion partner (46.3%) and detected in a range of tumour types. *NTRK1* gene fusions (with 13 different fusion partners) were detected in 40.7%) of patients, while an *NTRK2* gene fusion (*SQSTM1-NTRK2*) was detected in a single patient with a neuroendocrine tumour.

Figure 7: Tumour types in the efficacy evaluable analysis, N=54 (101)



CRC, colorectal cancer; MASC, mammary analogue secretory carcinoma; NSCLC, non-small cell lung cancer

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

B.2.4.1 Sample size

Table 10 summarises the planned per-protocol sample sizes for each of the three studies for which data are pooled for efficacy and provides the planned number of *NTRK* fusion-positive patients.

Assuming the true ORR by BICR (ORR-BICR) is 60%, a sample size of 56 patients will yield a 95% 2-sided confidence interval (CI) with precision ±14% that will exclude a lower limit of 30%. A response rate that excludes 30% or higher is considered clinically meaningful (58). This sample size is comparable to the number of patients contributing efficacy data for other agents that have been granted marketing authorisation for rare diseases, e.g. the 50 patients enrolled in the registration-enabling single arm study of crizotinib in *ROS1*-positive NSCLC (106).

Table 10: Planned sample sizes of *NTRK* fusion-positive patients by study for efficacy analyses (per individual study protocols) (107)

Study	Planned overall sample size	Planned sample size for *NTRK* fusion positive patients
ALKA	70	Not specified
STARTRK-1	15 (dose escalation) 50 (dose expansion)	Not specified
STARTRK-2	Up to 62 per gene fusion by tumour type bucket	Up to 62 per tumour type (e.g. *NTRK* sarcoma)

The pooled population of 355 safety-evaluable patients treated with entrectinib across all four clinical studies including adult patients (not only with *NTRK* fusion-positive tumours but also ROS1-positive tumours and other adult patients exposed to entrectinib) and paediatric patients with and without *NTRK*, *ALK*, and *ROS1* fusions is sufficient to adequately assess the safety of entrectinib. The size of this safety dataset meets the exposure of 300–600 patients, as recommended by the International Conference on Harmonisation guideline (108) to observe whether more frequently occurring events increase or decrease over time as well as to observe delayed events of reasonable frequency (e.g. in the general range of 0.5–5%) (58).

B.2.4.2 Analysis timing

An interim analysis of 19 NTRK fusion-positive patients was performed for the purpose of the FDA Breakthrough Therapy Designation submission on January 27th 2017 (107).

The final analysis of the integrated efficacy analysis set was planned to take place after approximately 56 NTRK fusion-positive patients had been enrolled across the three studies. All patients would have at least 6 months of efficacy follow-up from the time of response or would have discontinued study treatment at the time of final database snapshot for analysis. Safety parameters would be evaluated for all patients who received at least 1 dose of entrectinib and were enrolled on or before approximately 56 NTRK fusion-positive were enrolled across the three studies. Data that were integral to the analysis of safety and efficacy endpoints were reviewed for inconsistencies, queried, and finalised as a formal database lock prior to performing the final analysis (107).

Of note, STARTRK-2 enrolment continued even after reaching the integrated enrolment target of 56 NTRK fusion-positive patients across the three studies. Any patient enrolled after these approximately 56 patients have been enrolled will not be included in the primary integrated safety or efficacy analysis for the initial marketing application submission (107).

B.2.4.3 Integrated efficacy analysis endpoints

The endpoints in the integrated efficacy analysis were based on the Phase II STARTRK-2 study endpoints and are summarised in Table 11.

Table 11: Efficacy evaluable analysis endpoints

Primary endpoints (BICR assessment)
ORR: The proportion of patients with confirmed CR or PR that persisted on repeat-imaging ≥4 weeks after initial documentation of response. **DOR:** Measured from the date of first objective response to first documentation of radiographic disease progression or date of death due to any cause, whichever was earlier. For patients without disease progression or death, DOR was censored at the last tumour assessment prior to the CCOD. **BOR:** Best radiologic overall response (based on RECIST v1.1) recorded at any single timepoint from the start of treatment until disease progression.
Secondary endpoints
CBR: Proportion of patients with confirmed CR or PR and/or stable disease documented as lasting for at least 6 months following start of entrectinib. Patients without a post-baseline tumour assessment or patients who received at least one dose of entrectinib and who discontinued for any reason prior to undergoing one post-baseline response evaluation were counted as not achieving clinical benefit. **PFS:** Time (months) from first dose of entrectinib to first documentation of radiographic disease progression or death due to any cause. PFS data for patients without progression or death was censored on the date of the last tumour assessment (or, if no tumour assessment was performed after the baseline visit, at the date of first dose of entrectinib) prior to the CCOD. **Time to CNS progression:** Time (months) from first dose of entrectinib to first documentation of radiographic CNS disease progression (occurrence of a new CNS lesion or progression in any CNS lesion per RECIST v1.1 criteria) or death due to any cause. Patients without radiographic CNS progression or death were censored onthe date of the last tumour assessment. **OS:** Time (months) from the first dose of entrectinib to the date of death due to any cause. Patients who were alive at the time of the analysis were censored on the last known date that they were alive on or prior to the CCOD. In addition, the following censoring rules applied: • Patients with no post-baseline information were censored on the date of first dose of entrectinib • Patients who were lost to follow-up or withdrew consent for further follow-up were censored on the last known date that they were alive **Intracranial efficacy results according to CNS metastatic status at baseline, including the following endpoints:** **IC-ORR:** Selecting only CNS lesion(s) for each patient, the RECIST v1.1 algorithms for timepoint response and BOR assessment were used to determine IC response. A confirmed IC response was a CNS response that persisted on repeat-imaging ≥4 weeks after initial documentation of CNS response **IC-DOR:** Calculated only for IC responders and was measured from the date of first IC response to first documentation of radiographic CNS disease progression or date of death due to any cause, whichever was earlier. For patients without CNS disease progression and who had not died within 30 days of the last dose of study treatment, IC-DOR was censored at the last tumour assessment date prior to any date of subsequent anti-cancer therapy, including surgery or radiotherapy to the brain **IC-PFS:** Time (months) from first dose of entrectinib to first documentation of radiographic CNS disease progression or death due to any cause. Patients without radiographic IC progression or death were censored on the date of the last tumour assessment
Patient-reported outcomes

Quality of life and health status information were collected from self-administered instruments for patients enrolled in STARTRK-2 only. Therefore, the patient reported outcome endpoint analyses were based on STARTRK-2 and not based on data integrated across multiple studies. Analysis of patient reported outcome endpoints are based on the following instruments: QLQ-C30, QLQ-LC13, QLQ-CR29, and EQ-5D.
Subgroups analyses
Subgroups: • ORR by baseline demographics and clinical demographics • Efficacy by baseline CNS metastases

BICR, blinded independent central review; BOR, best overall response; CBR, clinical benefit rate; CCOD, clinical cut-off date; CNS, central nervous system; CR, complete response; DOR, duration of response; EQ-5D, EuroQol-5 Dimension; IC, intracranial; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; PR, partial response; QLQ-C30, Core Quality of Life Questionnaire; QLQ-LC13, Lung Cancer Module; QLQ-CR29, Colorectal Cancer Module

B.2.4.4 Statistical Analysis

Formal significance tests were not performed; therefore, P values were not reported. Instead, 95% 2-sided CIs for point estimates were utilised to estimate magnitude of effects.

Due to the rarity of this patient population and the expectation of significant clinical benefit, no statistical adjustment was made to address the sources of multiplicity associated with this integrated efficacy analysis. No other statistical adjustments were made to account for subgroup effects associated with pooling of data for this analysis.

Sensitivity analyses were performed to evaluate the robustness of therapeutic efficacy in patients with solid extracranial tumours and an *NTRK* gene fusion with measurable disease at baseline. The following efficacy endpoint and subgroups of patients were included in the analysis.

- ORR as determined by investigator (ORR-INV) estimated for the enrolled population
- ORR-BICR and ORR-INV estimated for the group of patients belonging to the efficacy evaluable analysis set in addition to any patients with extracranial solid tumours harbouring the *NTRK* gene fusion from the efficacy nonevaluable analysis set (e.g., nonmeasurable disease, baseline ECOG ≥3, etc.)

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

Quality assessment (conducted using the NICE Quality appraisal checklist [quantitative intervention studies: https://www.nice.org.uk/process/pmg4/chapter/appendix-f-quality-

appraisal-checklist-quantitative-intervention-studies]) was not performed for the seven included publications (Table 12) as they were not primary full publications (see Appendix D).

Table 12: Included studies from the systematic literature review

Author	Title	Journal	Year	Citation
Identified from electronic database searches (n=5)				
De Braud, F. G. (109)	*ALK*A-372-001: First-in-human, phase I study of entrectinib-an oral pan-trk, *ROS1*, and *ALK* inhibitor-in patients with advanced solid tumours with relevant molecular alterations	Journal of Clinical Oncology	2015	33 (15 Suppl. 1)
Desai, A. V. (110)	STARTRK-NG: A phase 1/1b study of entrectinib in children and adolescents with advanced solid tumours and primary CNS tumours, with or without TRK, *ROS1*, or *ALK* fusions	Cancer Research	2017	77 (13 Suppl. 1)
Drilon, A. (111)	STARTRK-2: A global phase 2, open-label, basket study of entrectinib in patients with locally advanced or metastatic solid tumours harboring TRK, *ROS1*, or *ALK* gene fusions	Cancer Research	2017	77 (13 Suppl. 1)
Drilon, A. (112)	Safety and antitumour activity of the multitargeted pan-TRK, *ROS1*, and *ALK* inhibitor entrectinib: Combined results from two phase I trials (*ALK*A-372-001 and STARTRK-1)	Cancer Discovery	2017	7(4):400-409.
Patel, M. R. (113)	STARTRK1: Phase 1/2a study of entrectinib, an oral Pa*NTRK*, *ROS1*, and *ALK* inhibitor, in patients with advanced solid tumours with relevant molecular alterations	Journal of Clinical Oncology	2015	33 (15 Suppl. 1)
Identified from supplementary hand searches (n=2)				
Desai, A. V. (114)	Phase 1 study of entrectinib (RXDX-101), a TRK, *ROS1*, and *ALK* inhibitor, in children, adolescents, and young adults with recurrent or refractory solid tumours.	ASCO 2018	2018	-
Demetri, G. D. (101)	Efficacy and safety of entrectinib in patients with *NTRK* fusion-positive tumours: Pooled analysis of STARTRK-2, STARTRK-1, and *ALK*A-372- 001	ESMO 2018	2018	-

B.2.6 Clinical effectiveness data

- **Entrectinib treatment resulted in a clinically meaningful ORR of 57.4% (95% CI: 43.2%, 70.8%) assessed by BICR. Four patients (7.4%) achieved CR and 27 patients (50.0%) had a PR. Responses were recorded in all solid tumour categories**
  - **At CCOD [redacted], entrectinib treatment resulted in a [redacted] assessed by BICR. [redacted] achieved CR and [redacted] had a PR**
- **Responses were durable, with a BICR-mDOR of 10.4 months in responders (95% CI: 7.1, NE) and a Kaplan-Meier estimated median PFS of 11.2 months (95% CI: 8.0, 14.9)**
  - **At CCOD [redacted], responses were durable, with a mDOR of [redacted] in responders ([redacted]) and a Kaplan-Meier estimated median PFS of [redacted]**
- **Results of the investigator-assessed responses and sensitivity analyses were consistent with the primary analyses**
- **The BICR-assessed clinical benefit rate was 64.8% (95% CI: 50.62, 77.32), indicating durable stable disease in some patients**
- **OS data are still immature with <30% of patients experiencing an event by the clinical cut-off for the integrated efficacy analyses**
  - **At CCOD [redacted], [redacted] had died and the Kaplan-Meier estimated median OS was [redacted]**
- **In Phase I Part A of STARTRK-NG (CCOD 31st May 2019), the overall ORR was [redacted] and the overall CBR was [redacted]**
- **At CCOD [redacted], antitumor activity was observed in [redacted] with solid tumours harbouring *NTRK* fusions**
  - **Objective responses were reported in [redacted] with tumours harbouring *NTRK* fusions**
- **Entrectinib has demonstrated a clinically meaningful benefit for adult patients whose tumours harbour *NTRK* gene fusions. Entrectinib has also demonstrated [redacted] in paediatric patients as [redacted]**

Unless otherwise stated, data presented in Section 2.6 are from the CSR 'Summary of Clinical Efficacy' (105). Data from the primary clinical cut-off date (CCOD) 31st May 2018 is presented for the integrated efficacy analysis, referred to as: 'efficacy evaluable analysis'. In addition, paediatric data from the primary CSR 'Primary Clinical Study Report - Report No. 1089445' and the supplementary results report 'Supplementary results report for study STARTRK-NG (CO40778) in paediatric patients' is presented. The economic analysis presented in Section B.3.2 has been carried out using only data from the integrated efficacy evaluable analysis with CCOD 31st May 2018. Data from CCOD [redacted] is also presented; this analysis was carried out in response to Day 75 questions from the FDA where additional data was requested.

At the time of the efficacy evaluable analysis, median duration of follow-up in adult responders from the time of first response was 13.1 months (range: 2.8–21.0) and median survival follow-up among all adult patients in the efficacy evaluable analysis set was 12.9 months (range: 0.6–24.7).

B.2.6.1 Primary efficacy endpoints

Objective response rate and best overall response (CCOD of 31st May 2018)

ORR was achieved in 57.4% of patients with the lower limit of the 95% CI excluding 30% (95% CI: 43.2%, 70.8%) demonstrating that entrectinib had a clinically meaningful effect (Table 13). [redacted] achieved CR and [redacted] had a PR. Disease progression was documented in 4 patients (7.4%). The majority of objective responses were achieved at the first tumour assessment after commencing entrectinib treatment (end of Cycle 1) (101).

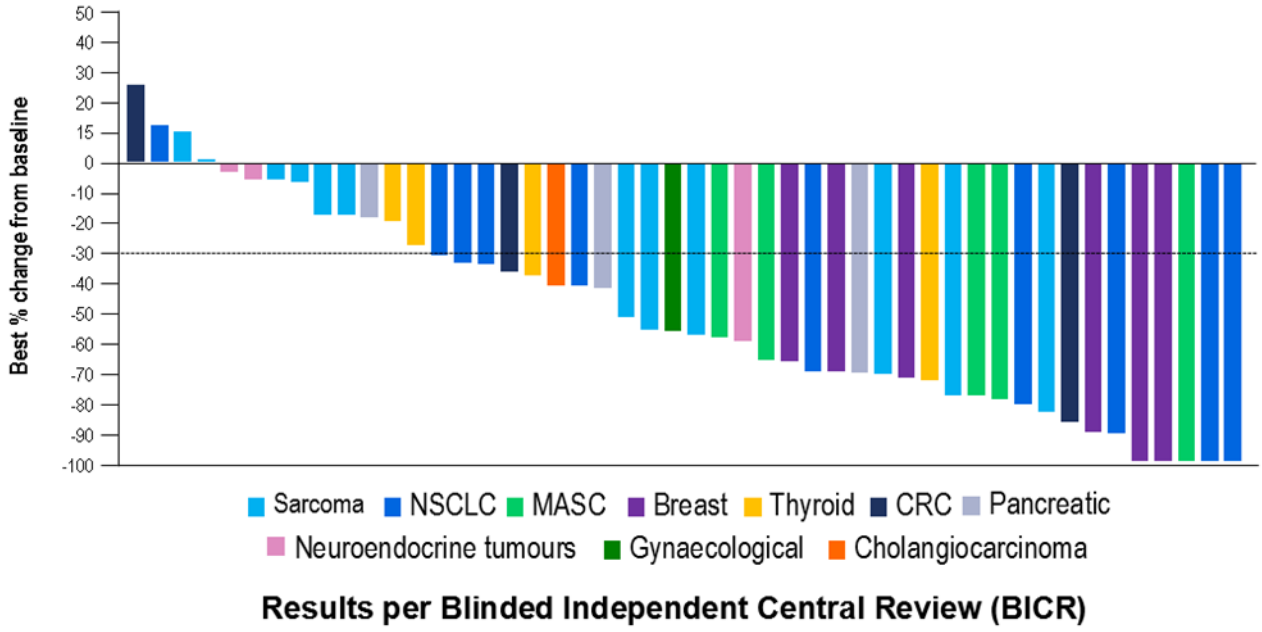
Table 13: Objective response and best overall response, BICR assessment (efficacy evaluable analysis) (101, 105)

n (%)	Total N=54
Responders Non-responders	31 (57.4) [redacted]
95% CI for response rates	(43.21, 70.77)
Complete response (CR)	[redacted]
Partial response (PR)	[redacted]
Stable disease (SD)	9 (16.7)
Progressive disease (PD)	4 (7.4)
Non CR/PD	3 (5.9)
Missing or unevaluable	7 (13.0)

Best Overall Response is derived per RECIST 1.1. Not Evaluable/Not Done category includes patients having on-study scans that could not be evaluated and patients who discontinued prior to obtaining adequate scans to evaluate or confirm response. SD and non CR/PD must be observed study day 35 or later, otherwise they count as NE. Objective response is defined as PR or CR confirmed by repeat imaging at least 28 days following first documentation of response. Otherwise, the patient is considered to be a non-responder. Patients were categorised as having non-CR/non-PD if they had non-target lesions (as assessed by BICR), but had measurable disease at baseline as assessed by Investigator.

Response to entrectinib treatment was observed across tumour types (Figure 8). In addition, responses were independent of the *NTRK* fusion gene (Figure 9). Except for 1 patient with an *NTRK2* fusion, patients had either an *NTRK1* or *NTRK3* fusion and ORRs in these patients were consistent with the overall efficacy evaluable analysis. Tumours with a *NTRK1* and *NTRK3* fusion gene displayed a 59.1% and 58.1% response rate to entrectinib, respectively.

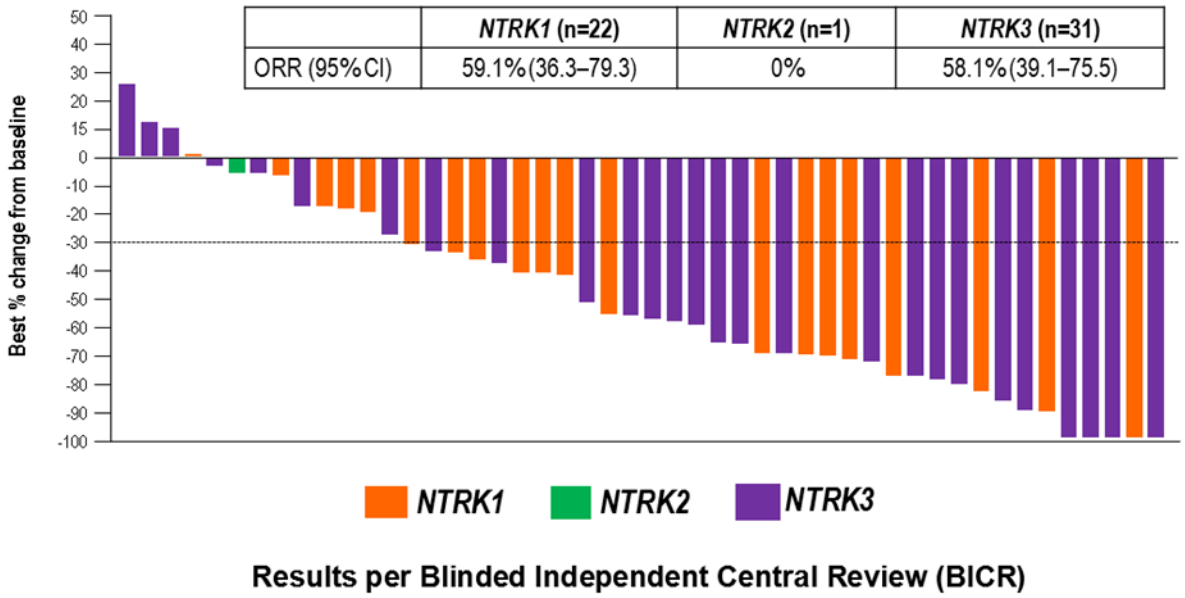
Figure 8: Entrectinib activity in *NTRK* fusion-positive solid tumours: individual patient responses by tumour type, BICR assessment (efficacy evaluable analysis) (101)



Patients with missing SLD percent change (N=6) were excluded from the plot.

BICR, blinded independent central review; CRC, colorectal cancer; MASC, mammary analogue of the salivary gland; NSCLC, non-small-cell lung cancer; SLD, sum of longest diameter

Figure 9: Entrectinib activity in *NTRK* fusion-positive solid tumours: individual patient responses by NTRK gene, BICR assessment (efficacy evaluable analysis) (101)



Patients with missing SLD percent change (n=6) were excluded from the plot.

BICR, blinded independent central review; NTRK, neurotrophic tyrosine receptor kinase; SLD, sum of longest diameter

Objective response rate and best overall response ([redacted])

ORR was achieved in [redacted] of patients (95% CI: [redacted]) demonstrating that entrectinib had a clinically meaningful effect. [redacted] achieved CR and [redacted] had a PR (115).

Table 14: Objective response rate and best overall response, BICR assessment (efficacy evaluable analysis) (115)

	BICR-assessed (n=54)
Responders (n)	[redacted]
ORR (95% CI)	[redacted]
Complete Response	[redacted]
Partial Response	[redacted]
Stable Disease	[redacted]
Progressive Disease	[redacted]
Non – CR/PD	[redacted]

Missing or Unevaluable	[redacted]
Clinical Benefit Rate* - (95% CI)	[redacted]

*Includes stable disease for a minimum of 6 months

CCOD of [redacted]

Duration of response (CCOD of 31st May 2018)

Responses were durable with a median DOR among responders, as assessed by the BICR, of 10.4 months (95% CI: 7.1, NE, Table 15). Approximately half (51.6%) of the 31 responders had an event (101). At the primary CCOD, [redacted] had responses lasting longer than 6 months, and the 6-month event-free rate was [redacted] (105).

Table 15: Kaplan-Meier event-free rates for duration of response, BICR assessment (efficacy evaluable analysis)

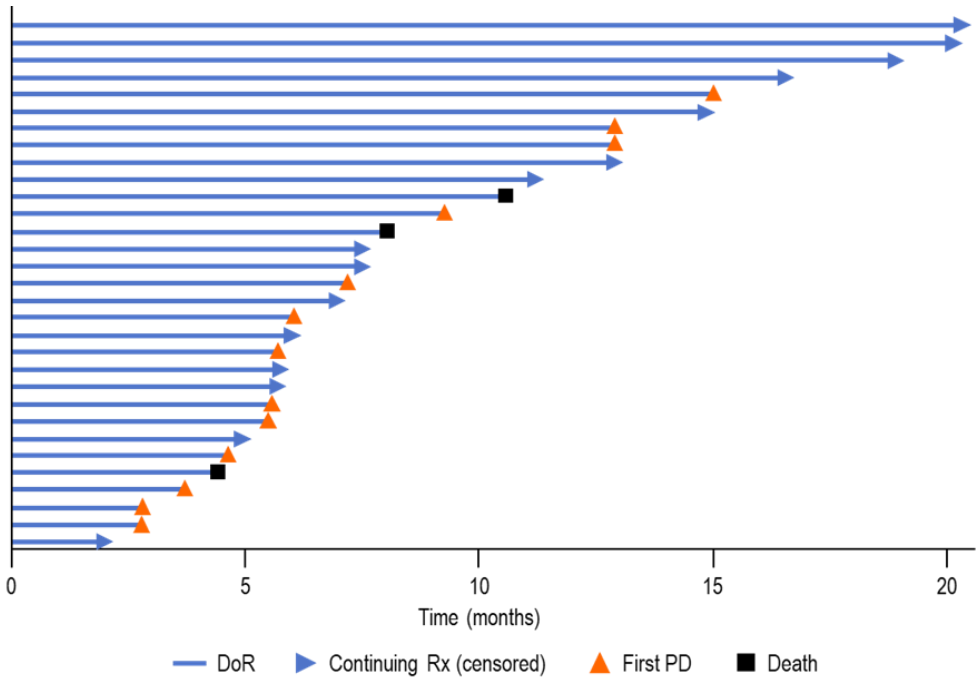
	Total N=31
Patients with event, n (%)	16 (51.6)
Earliest contributing event, n	
Disease progression	13
Death	3
Median time to event, months	10.4
95% CI	(7.1, NE)
6 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
9 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
12 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
18 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]

CI, confidence interval; NE, not estimated

Figure 10: Kaplan-Meier curve, BICR-assessed DOR (efficacy evaluable analysis) (115)

A swimmer plot for the 31 responses in the NTRK efficacy evaluable analysis set is shown in Figure 11.

Figure 11: Duration of response, PFS and OS, BICR assessment (efficacy evaluable analysis) (101)



Duration of response ([redacted])

Responses were durable with a median DOR among responders, as assessed by the BICR, of [redacted]. Approximately [redacted] of the 32 responders had an event.

Table 16: Duration of response, BICR assessment (efficacy evaluable analysis), updated analysis (115)

	BICR-assessed DOR
Pts included in analysis (Responders)	[redacted]
Pts with event (%)	[redacted]
Progressive Disease	[redacted]
Death	[redacted]
Time to Event (months)	
Median	[redacted]
95% CI for Median	[redacted]
25% and 75%-ile	[redacted]
Range	[redacted]

CCOD of [redacted]

* Subject to censoring

Figure 12: Kaplan-Meier curve for BICR-assessed DOR (efficacy evaluable analysis set), updated analysis (115)

[redacted]

CCOD of [redacted]

B.2.6.2 Secondary efficacy endpoints

Clinical benefit rate (CCOD 31st May 2018 and [redacted])

At CCOD of 31st May 2018, a total of 35 adult patients had confirmed [redacted], as assessed by the BICR, resulting in a CBR of 64.8%, indicating additional benefit of durable stable disease in some patients (Table 17).

At CCOD of [redacted], the CBR as assessed by BICR [redacted] (115).

Table 17: Clinical benefit rate, BICR assessment (efficacy evaluable analysis)

	Total N=54
Clinical benefit rate, n (%)	[redacted]
95% CI	[redacted]

Clinical benefit rate includes all patients with CR or PR plus patients with SD for at least 6 months after start of entrectinib. Otherwise, the patient is considered to not have clinical benefit.

Progression-free survival (CCOD of 31st May 2018)

The Kaplan-Meier estimated median PFS based on the BICR assessment was 11.2 months (95% CI: 8.0, 14.9), which excluded the lower limit of 6 months and indicated durability of entrectinib treatment effect (Table 18).

Table 18: Kaplan-Meier event-free rates for PFS, BICR assessment (efficacy evaluable analysis)

	Total N=54
Patients with event, n (%)	29 (53.7)
Earliest contributing event, n	
Disease progression	20
Death	9
Median time to event, months	11.2
95% CI	(8.0, 14.9)
6 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
9 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
12 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]

95% CI	[redacted]
18 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]

CI, confidence interval; NE, not estimated

Figure 13: Kaplan-Meier curve for BICR-assessed PFS (efficacy evaluable analysis) (115)

[redacted]

Progression-free survival (CCOD of [redacted])

The Kaplan-Meier estimated median PFS based on the BICR assessment was [redacted], Table 19) (115).

Table 19: Progression-free survival BICR assessment (efficacy evaluable analysis), updated analysis (115)

	BICR-assessed PFS (n = 54)
Patients with event (%)	[redacted]
Progressive Disease	[redacted]
Death	[redacted]

Time to Event (months)	
Median	[redacted]
95% CI for Median	[redacted]
25% and 75%-ile	[redacted]
Range	[redacted]

CCOD of [redacted]

* Subject to censoring

Figure 14: Kaplan-Meier curve for BICR-assessed PFS (efficacy evaluable analysis), updated analysis (115)

[redacted]

CCOD of [redacted]

Time to CNS progression (CCOD of 31st May 2018)

Durability of treatment effect was also observed for time to first documentation of radiographic CNS disease progression or death due to any cause with a time to CNS progression of [redacted], potentially indicating a durable protective effect against progression in the CNS (Table 20).

Table 20: Kaplan-Meier event-free rates for time to CNS progression, BICR assessment (efficacy evaluable analysis)

	Total N=54
Patients with event, n (%)	[redacted]
Earliest contributing event, n	

Disease progression	[redacted]
Death	[redacted]
Patients without event, n (%)	[redacted]
Median time to event, months	[redacted]
95% CI	[redacted]
6 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
9 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
12 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
18 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]

CI, confidence interval; NE, not estimated

Overall survival (CCOD of 31st May 2018)

At the time of the primary integrated efficacy analyses (CCOD of May 31st 2018), 16 patients (29.6%) had died and the Kaplan-Meier estimated median OS was 20.9 months (95% CI: 14.9, NE); however, these data are immature with <30% of patients experiencing events by the CCOD.

Table 21: Kaplan-Meier event-free rates for overall survival, BICR assessment (efficacy evaluable analysis)

	Total N=54
Patients with event, n (%)	[redacted]
Earliest contributing event, n	
Death	[redacted]
Patients without event, n (%)	[redacted]
Median time to event, months	20.9
95% CI	(14.9, NE)
6 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
9 months	

Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
12 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]
18 months	
Patients remaining at risk, n	[redacted]
Event free probability	[redacted]
95% CI	[redacted]

CI, confidence interval; NE, not estimated

Overall survival ([redacted])

At the CCOD of [redacted], [redacted] had died and the Kaplan-Meier estimated median OS was [redacted] (Table 22) (115).

Table 22: Overall survival, BICR assessment (efficacy evaluable analysis set), updated analysis (115)

	Overall survival (n = 54)
Pts with event (%)	[redacted]
Time to Event (months)	
Median	[redacted]
95% CI for Median	[redacted]
25% and 75%-ile	[redacted]
Range	[redacted]

CCOD of [redacted]

* Subject to censoring

B.2.6.3 Sensitivity analyses

The investigator-assessed ORR was [redacted], consistent with the BICR-assessed ORR (Table 23). [redacted] had CR and [redacted] had PR. Concordance between BICR- and investigator-assessed response was [redacted], with [redacted] identified by both the BICR and investigator. Discordance in the determination of PD (PD per investigator and no PD per BICR) was observed for [redacted]. Discordance in the time of PD (dates differed by >30 days) was observed for [redacted]; PD was determined by the investigator earlier than by the BICR for [redacted] and PD was determined by the investigator later than by the BICR for [redacted]. Median DOR for the [redacted] based on the investigator assessment was [redacted]. Results of the investigator-assessed responses and sensitivity analyses were consistent with the primary analyses.

Table 23: Overview of efficacy in adult patients with *NTRK* fusion-positive solid tumours as assessed by the investigator (efficacy evaluable analysis)

	N=54
Objective Response	
Patients with confirmed CR or PR, n	[redacted]
ORR, % (95% CI)	[redacted]
Patients with CR, n (%)	[redacted]
Patients with PR, n (%)	[redacted]
Patients with stable disease, n (%)	[redacted]
Duration of Response	
Median [a], months (95% CI)	[redacted]
Clinical Benefit Rate	
CBR (95% CI)	[redacted]

CBR, clinical benefit rate; CR, complete response; NTRK, neurotrophic tyrosine receptor kinase; ORR=objective response rate; PR, partial response.

[a] Median duration of response was estimated using Kaplan-Meier methods and measures of time from first response to death or progressive disease (censored at the last tumour assessment).

B.2.6.4 Patient-reported outcomes

Patient reported outcomes (PROs) were only evaluated in STARTRK-2, and were not included in the integrated efficacy analysis; the following data are a summary of the PRO assessments and results in the NTRK population from STARTRK-2.

Prior to the first dose of entrectinib on Cycle 1 Day 1, pre-dose on Day 1 of each subsequent treatment cycle, and at the End of Treatment, patients' health-related quality of life (HRQoL) was assessed through a self-administered validated questionnaire: the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (QLQ-C30). In addition, patients with NSCLC completed the lung cancer module, lung cancer module (QLQ-LC13) and patients with metastatic colorectal cancer (mCRC) completed the colorectal cancer module (QLQ-CR29). An EuroQol- 5 Dimension (EQ-5D) questionnaire was also administered; the results are presented in B.3.4.1 (Table 50).

All efficacy evaluable patients (N=51) completed the QLQ-C30 regardless of their tumour type. Nine patients with NSCLC completed the QLQ-LC13 and 3 patients with mCRC tumours completed the QLQ-CRC29. The completion rates for QLQ-C30, QLQ-LC13, and QLQ-CR29 were high at baseline (94.1%, 100%, and 100%, respectively) and the completion rate remained high (≥80%) at most study visits.

At baseline, patients reported [redacted] on a score ranging from 0 to 100, with 100 reflecting better functioning.

While receiving entrectinib, patients tended to [redacted] on the global health status). For functional scales (e.g., physical functioning, role functioning), patients continued to report [redacted] scores at most study visits, with a trend towards clinical improvement, with the exception of cognitive functioning, which while maintaining overall its [redacted], trended towards some [redacted] above the accepted within-arm clinical meaningful threshold of 10-points (worst mean change score of [redacted] at Cycle 20 Day 1).

According to the QLQ-C30, patients in the safety analysis population generally reported they [redacted] in the past week. In addition, [redacted] number of patients reported experiencing treatment-related symptoms such [redacted] at some time points while receiving treatment (99).

B.2.6.5 Patient disposition for entrectinib treatment

As of the CCOD of 31st May 2018, A total of 31 patients had discontinued entrectinib treatment. The main reason for discontinuation of entrectinib treatment was [redacted] (Table 24). The median duration of treatment with entrectinib in the efficacy evaluable analysis was [redacted].

Table 24: Patient disposition for entrectinib treatment (efficacy evaluable analysis) (CCOD of 31st May 2018)

	Total (N=54)
Discontinued Treatment	[redacted]
Adverse Event	[redacted]
Informed Consent Withdrawn	[redacted]
Progressive Disease	[redacted]

B.2.6.6 STARTRK-NG paediatric efficacy results

Primary analysis (CCOD 31st May 2018)

The STARTRK-NG study is a Phase I/Ib, 5-part, multicentre, open-label study evaluating the effect of entrectinib in children, adolescent, and young adult patients. The study consisted of a dose escalation phase (Phase I) in patients with relapsed or refractory extracranial solid tumours, with or without molecular alterations (Part A), plus expansion parts (Phase Ib) in patients with primary brain tumours harbouring *NTRK1/2/3, ROS1*, or *ALK* molecular alterations (Part B), neuroblastoma (Part C), other non-neuroblastoma, extracranial solid tumours harbouring *NTRK1/2/3, ROS1*, or *ALK* gene fusions (Part D) and an exploratory cohort of patients who were otherwise eligible but unable to swallow capsules (Part E) (104).

A total of 16 patients were enrolled into the Phase 1 (dose escalation) portion and received at least one dose of entrectinib. One of these patients with infantile fibrosarcoma had an *EML4-NTRK3* gene fusion. At the time of clinical cut-off (31st May 2018), a total of [redacted]) were discontinued from the study. [redacted] remained on study treatment and [redacted] were under survival follow-up. The median age of patients enrolled was 9.5 years (range: 4 - 20 years). Other than [redacted] enrolled in the 400 mg/m^2 dose group, [redacted], with the majority of patients [redacted] (104).

Efficacy analysis for the Phase 1 (dose escalation) portion of the study was conducted in the safety evaluable population (n=16). Among the 16 patients, objective responses were reported in each of the [redacted] with tumours harbouring gene fusions. All objective responses were achieved within [redacted] of first treatment administration and were [redacted] at the last tumour assessment visit prior to the clinical data cutoff date. The overall ORR was [redacted]. The overall CBR was [redacted] (104).

Table 25: STARTRK-NG Phase I dose escalation (Part A) – Summary of overall response (Safety Population) (104)

	250 mg/m^2/day (n=3)	400 mg/m^2/day (n=3)	550 mg/m^2/day (n=7)	750 mg/m^2/day (n=3)	Subtotal (N=16)
Responders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Non-Responders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI for Response Rates	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Complete Response (CR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
95% CI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Partial Response (PR)					
95% CI					
Stable Disease (SD)					
95% CI					
Progressive Disease (PD)					
95% CI					
Non CR/PD					
95% CI					
Missing or unevaluable					

Note: Percentages of subjects are calculated based on the number of subjects in each assigned dose level in Phase 1 dose escalation and each tumour type cohort/part in Phase 1b. Confidence Interval is calculated using Clopper-Pearson exact confidence interval.

Table 26: STARTRK-NG Phase I dose escalation (Part A) - Clinical benefit rate (Safety Population) (104)

	250 mg/m^2/day (n=3)	**400 mg/m^2/day (n=3)**	**550 mg/m^2/day (n=7)**	**750 mg/m^2/day (n=3)**	**Subtotal (N=16)**
Clinical Benefit Rate					
95% CI					

CI: confidence interval

STARTK-NG analysis (date of data cut-off)

The results in this updated analysis are an aggregate of safety, efficacy and PK data from all patients enrolled into the study as of 31st May 2018 (n=26), including an update of the 16 patients in Part A as well as results from 10 patients enrolled into the Phase Ib expansion (which includes the 4 patients with <6 months follow up provided in the initial application to support efficacy). The results in this report are presented for the 26 pooled paediatric patients and are based on new analyses. The results presented are based on data collected up to a clinical cut-off date of (100).

Table 27: Summary of tumour response in paediatric patients treated with entrectinib in phase I and Ib of study STARTK-NG by tumour type and gene alteration (Safety Population) (100)

	Gene alteration			
	NTRK fusion	**ROS1 fusion**	**ALK fusion**	**other/unknown**
All Patients	(n=5)	(n=2)	(n=2)	(n=17)[a]
Objective Response Rate[b]				
No. of patients with confirmed CR or PR, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Complete Response, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Partial Response, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Stable Disease, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Missing or unevaluable	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Benefit Rate[b]				
No. of patients with confirmed CR or PR, or SD ≥6 months, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Best Overall Response by Tumour Type				
Patients with non-neuroblastoma extracranial solid tumours	(n=3)	(n=1)	(n=2)	(n=2)[a]
	[redacted]	[redacted]	[redacted]	[redacted]
Patients with primary CNS tumours	(n=2)	(n=1)	-	-
	[redacted]	[redacted]	[redacted]	[redacted]
Patients with neuroblastoma	-	-	-	(n=15)
	[redacted]	[redacted]	[redacted]	[redacted]

CCOD [redacted]

[a] One patient with [redacted]

[b] Tumour response and progression was assessed using criteria applicable to the appropriate imaging modality for the primary malignancy, i.e., RECIST v 1.1 (measureable extracranial solid tumours), with or without Curie Scale (neuroblastoma with MIBG-avid lesions), or RANO (measurable primary CNS disease).

[c] Patient ##19038/05003 achieved an [redacted]

[d] Patient ##19026/05005 [redacted]

The time to first response, DOR and duration of treatment for each patient with solid tumours harbouring NTRK fusions at the time of the CCOD is shown in Table 28.

Table 28: Efficacy Listing of Paediatric Patients with NTRK Fusion-Positive Solid Tumours in STARTRK-NG (100)

Study Phase/ Part	Assigned dose level (mg/m^2)[a]	Patient ID	Age/sex	NTRK Gene Fusion	Tumor type	BOR (INV) (at visit Cycle Day)	Clinical benefit[b]	DOR (INV) (months)	Time to response (months)	PFS (INV) (months)	OS (months)
Phase I											
Part A	750	19038/01015	4/F	*EML4-NTRK3*	infantile fibrosarcoma	PR (C3D1)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Phase Ib											
Part B	550	19029/02001	3/F	*ETV6-NTRK3*	epitheloid glioblastoma	CR[c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Part E	400	19038/05002	4/F	*TPR-NTRK1*	high grade glioma	PR[c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	400	19038/05003	4.5 mo/M	*ETV6-NTRK3*	infantile fibrosarcoma	SD[e]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	400	19043/05001	4/F	*ETV6-NTRK3*	metastatic melanoma	PR (C3D1)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CCOD [redacted]

[a] All patients with NTRK-fusion-positive tumours in STARTRK-NG received the F01 formulation.

[b] Patient with either CR, PR or stable disease (SD) at 6 months after the first dose of entrectinib, as assessed by RECIST v1.1.

[c] response assessment was by RANO criteria.

[d] response ongoing at time of clinical cutoff date (31 October 2018).

[e] A partial response was recorded at the last tumour assessment (C10D1), but had not been confirmed by the time of the clinical cutoff date.

[f] Patient continued to receive entrectinib at the time of the clinical cutoff date.

B.2.7 Subgroup analysis

- **In general, ORRs in most subgroups were consistent in the efficacy evaluable analysis. While numerical differences in ORR were observed in some subgroups, the sample sizes were too small to make any meaningful inference**
- **Patients receiving entrectinib had clinically equivalent responses regardless of the presence or absence of CNS disease. The ORR for patients with CNS disease was 50.0% (95% CI: 21.1, 78.9) compared to 59.5% (95% CI: 43.3, 74.4) in patients without baseline CNS metastases**
  - **At CCOD [redacted], ORRs were [redacted] and [redacted] in patients with and without baseline CNS metastases, respectively**
- **Median DOR was not estimable in patients with baseline CNS metastases; durable response was observed in patients without baseline CNS metastases with a median DOR of 12.9 months (95% CI: 7.1, NE)**
  - **At CCOD [redacted], the median DOR [redacted] with baseline CNS metastases; durable response was [redacted] without baseline CNS metastases with a median DOR of [redacted]**
- **The Kaplan-Meier estimated median PFS was [redacted] in patients with baseline CNS metastases and [redacted] in patients without baseline CNS metastases**
  - **At CCOD [redacted], the Kaplan-Meier estimated median PFS [redacted] in patients with baseline CNS metastases and [redacted] in patients without baseline CNS metastases**
- **In patients with baseline CNS metastases confirmed by the BICR, clinically meaningful intracranial efficacy (ORR of 54.5% [95% CI: 23.4, 83.3]) was of a similar magnitude to the systemic response**
  - **At CCOD [redacted], the intracranial ORR [redacted]**

Please refer to Appendix E for full details of the subgroup analyses from the integrated efficacy analysis population. Unless otherwise stated, the data presented in Section B.2.7 is from the 'Integrated Analysis CSR - Summary of Clinical Efficacy' (102).

B.2.7.1 Objective response rate by subgroups, BICR assessment

In general, ORR in most subgroups was consistent with the efficacy evaluable analysis. Some subgroups, such as ECOG PS 2 and patients with four or more prior anticancer radiation therapies, appeared to have numerically worse response to treatment with entrectinib than others. These would seem to align with comorbidities and prognostic factors; however, the small sample sizes prohibited meaningful interpretation. Numerical differences in ORR with overlapping CIs were observed in the subgroups (data presented in Appendix E).

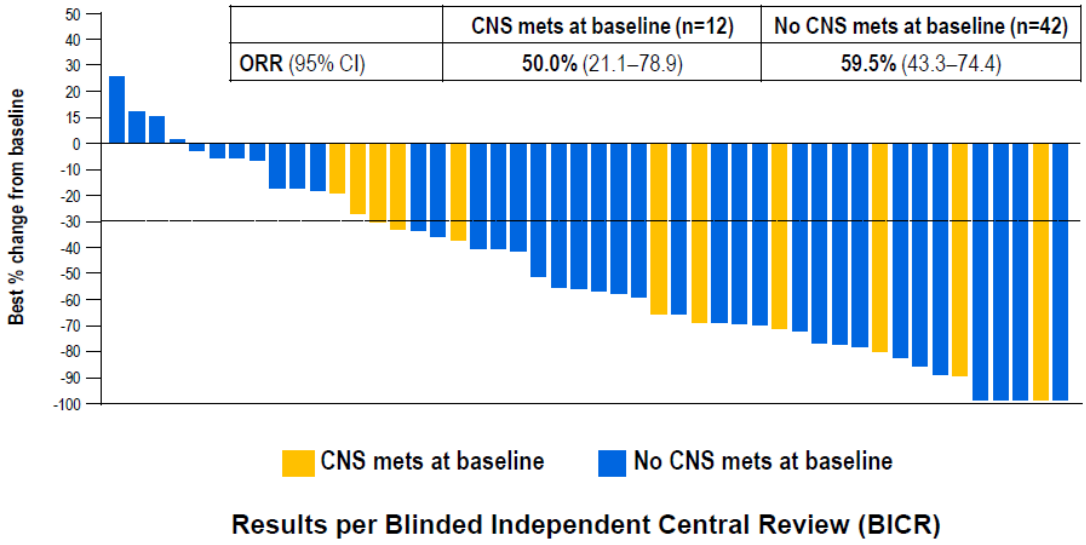
As described in Section 2.6.1, the majority of patients had either an *NTRK1* or *NTRK3* fusion and ORRs were similar in these patients (59.1% [95% CI: 36.4, 79.3] and 58.1% [95% CI: 39.1, 75.5], respectively). The 1 patient with an *NTRK2* fusion was a non-responder. *ETV6-NTRK3* (25 patients [46.3%]) was the only gene fusion partner reported in more than 2 patients at the time of enrolment and ORR for patients with *ETV6-NTRK3* was 68.0% (95% CI: 46.5, 85.1).

B.2.7.2 Systemic efficacy by BICR for baseline CNS metastases subgroups (CCOD of 31st May 2018 and [redacted])

At CCOD 31st May 2018, within the NTRK efficacy evaluable analysis set, 42 patients were included in the no CNS metastases analysis set and 12 patients were included in the CNS metastases analysis set, based on the presence or absence of CNS metastases as determined by the investigator at baseline. Entrectinib demonstrated similar response rates regardless of the presence of CNS metastatic disease at baseline (105):

- ORRs were 50.0% (95% CI: 21.1, 78.9) and 59.5% (95% CI: 43.3, 74.4) in patients with and without baseline CNS metastases, respectively (Figure 15)
- Median DOR was not estimable in patients with baseline CNS metastases; durable response was observed in patients without baseline CNS metastases with a median DOR of 12.9 months (95% CI: 7.1, NE)
- The KM estimated median PFS was [redacted] in patients with baseline CNS metastases and [redacted] in patients without baseline CNS metastases
- The OS was [redacted] in patients with baseline CNS metastases and [redacted] in patients without baseline CNS metastases

Figure 15: Entrectinib activity in NTRK fusion-positive solid tumours: individual patient responses by CNS mets status, BICR assessment (101)



BICR, blinded independent central review; NTRK, neurotrophic tyrosine receptor kinase; SLD, sum of longest diameter

At CCOD [redacted] (115):

- ORRs were [redacted] and [redacted] in patients with and without baseline CNS metastases, respectively (Table 29)
- Median DOR was [redacted] in patients with baseline CNS metastases; durable response was observed in patients without baseline CNS metastases with a median DOR of [redacted], Table 30)
- The KM estimated median PFS was [redacted] in patients with baseline CNS metastases and [redacted] in patients without baseline CNS metastases

Table 29: Systemic ORR by Baseline CNS Metastatic Disease Status[a], BICR Assessment, updated analysis (efficacy evaluable analysis)

	CCOD 31st May 2018		CCOD [redacted]	
	CNS Disease at baseline (n=12)	No CNS Disease at Baseline (n=42)	CNS Disease at baseline (n=12)	No CNS Disease at Baseline (n=42)
Responders	6 (50.0%)	25 (59.5%)	[redacted]	[redacted]
95% CI (%)	(21.09, 78.91)	(43.28, 74.37)	[redacted]	[redacted]
Complete Response	[redacted]	[redacted]	[redacted]	[redacted]

Partial Response	[redacted]	[redacted]	[redacted]	[redacted]
Stable Disease	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease	[redacted]	[redacted]	[redacted]	[redacted]
Non – CR/PD	[redacted]	[redacted]	[redacted]	[redacted]
Missing or Unevaluable	[redacted]	[redacted]	[redacted]	[redacted]

CCOD of 31st May 2018 and [redacted]

[a]CNS disease status determined by Investigator

Table 30: Systemic DOR by Baseline CNS Metastatic Disease Status[a], BICR Assessment, updated analysis (efficacy evaluable analysis)

	CCOD 31st May 2018		CCOD [redacted]	
	CNS Disease at baseline (n=12)	**No CNS Disease at Baseline (n=42)**	**CNS Disease at baseline (n=12)**	**No CNS Disease at Baseline (n=42)**
Patients included in analysis (Responders)	[redacted]	[redacted]	[redacted]	[redacted]
Patients with event (%)	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease	[redacted]	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]	[redacted]
Time to Event (months)	[redacted]	[redacted]	[redacted]	[redacted]
Median	[redacted]	[redacted]	[redacted]	[redacted]
95% CI for Median	[redacted]	[redacted]	[redacted]	[redacted]
25% and 75%-ile	[redacted]	[redacted]	[redacted]	[redacted]
Range	[redacted]	[redacted]	[redacted]	[redacted]

CCOD of 31st May 2018 and [redacted]

[a]CNS disease status determined by Investigator

B.2.7.3 Intracranial efficacy in patients with baseline CNS metastatic disease, BICR assessment (CCOD of 31st May 2018 and [redacted])

CNS metastatic disease, as assessed by the investigator, was documented at baseline in 12 patients. Of these patients, CNS metastatic disease was confirmed by the BICR in 11 patients of which CNS metastases was measurable in 7 patients (101, 105). At the CCOD of 31st May 2018, in patients who had CNS metastatic disease at baseline (n=11) treated with entrectinib, BICR assessment observed clinically meaningful efficacy of a similar magnitude to the systemic response (Table 31). The IC-ORR was 54.5% (95% CI: 23.4, 83.3), with 3 patients (27.3%) achieving a CR and 3 patients (27.3%) achieving PR. The response in

patients with measurable CNS metastases was similar with an intracranial ORR of 57.1% (95% CI: 18.4, 90.1).

A summary of the intracranial assessment results of 11 patients with confirmed baseline CNS metastatic disease is provided in Table 31.

Table 31: Overview of intracranial efficacy in patients with baseline CNS metastatic disease, BICR assessment (efficacy evaluable analysis) (101, 105)

	All patients N=11	Patients with measurable disease n=7
Objective response		
Responders, n	6	[redacted]
ORR, % (95% CI)	54.5 (23.38, 83.25)	[redacted]
Best overall response		
Patients with, n (%)		
CR	3 (27.3)	[redacted]
PR	3 (27.3)	[redacted]
SD	1 (9.1)	[redacted]
PD	1 (9.1)	[redacted]
Non CR/PD	2 (18.2)	[redacted]
Missing of unevaluable response	1 (9.1)	[redacted]
Duration of intracranial response		
Patients with event, n (%)	[redacted]	[redacted]
Median, months (95% CI)	NE (5.0, NE)	[redacted]
Progression-free survival		
Patients with event, n (%)	[redacted]	[redacted]
Median, months (95% CI)	14.3 (5.1, NE)	[redacted]

CI, confidence interval; CR, complete response; NE, not estimated; ORR, objective response rate; PD, progressive disease; PR, partial response; SD, stable disease

At the CCOD of [redacted], the number of responders remain unchanged (Table 32) (115).

Table 32: Intracranial ORR in patients with Baseline CNS Metastatic Disease Status[a], BICR Assessment, updated analysis (efficacy evaluable analysis) (115)

	[redacted] n =11
Responders	[redacted]
95% CI (%)	[redacted]
Complete Response	[redacted]
Partial Response	[redacted]
Stable Disease	[redacted]

Progressive Disease	[redacted]
Non – CR/PD	[redacted]
Missing or Unevaluable	[redacted]

[a]CNS disease status determined by BICR

The median intracranial DOR was not estimable given that majority of the responders among patients with CNS metastatic disease at baseline were still ongoing without an event at the time of CCOD (Table 33).

Table 33: Intracranial DOR in patients with baseline CNS metastatic disease status[a], BICR assessment (efficacy evaluable analysis) (115)

	CCOD 31st May 2018 n=6	**CCOD [redacted] n=6**
Patients with event (%)	[redacted]	[redacted]
Progression	[redacted]	[redacted]
Death	[redacted]	[redacted]
Time to Event		
Median	[redacted]	[redacted]
95% CI for Median	[redacted]	[redacted]
25% and 75%-ile	[redacted]	[redacted]
Range	[redacted]	[redacted]

[a]CNS disease status determined by BICR

At CCOD 31st May 2018, the KM estimated median intracranial PFS based on the BICR assessment was 14.3 months (95% CI: 5.1, NE), reflecting the durability of entrectinib treatment effect in CNS metastatic lesions (Table 34). At the CCOD [redacted], the KM estimated median intracranial PFS based on the BICR assessment remained at [redacted] (Table 34).

Table 34: Intracranial PFS in patients with Baseline CNS Metastatic Disease Status[a], BICR Assessment (efficacy evaluable analysis) (115)

	CCOD 31st May 2018 (n = 11)	**CCOD [redacted] (n = 11)**
Patients with event (%)	[redacted]	[redacted]
Progression	[redacted]	[redacted]
Death	[redacted]	[redacted]
Time to Event		
Median	14.3	[redacted]
95% CI for Median	(5.1, NE)	[redacted]
25% and 75%-ile	[redacted]	[redacted]
Range	[redacted]	[redacted]

[a]CNS disease status determined by BICR

B.2.8 Meta-analysis

Due to the significant heterogeneity between patient and disease characteristics, tumour types and potential comparator therapies meeting the definition of 'established management without entrectinib', a network meta-analysis was not feasible.

B.2.9 Indirect and mixed treatment comparisons

A formal mixed treatment comparison was not deemed feasible due to the significant heterogeneity between patient and disease characteristics, tumour types and potential comparator therapies meeting the definition of 'established care'. For the purposes of economic evaluation, a naïve weighted comparison was therefore developed using published data for a population of patients where *NTRK* fusion-positive status was not reported (Further discussed in Section 3.3.1 and data extraction table available in Appendix L). Given the broad indication covered by the decision problem, a full systematic review was not deemed feasible; an initial screening search identified >1,000,000 publications. Instead, the approach described below was developed to provide an overview of potential comparator outcomes, while pragmatically reducing the number of evidence sources identified.

B.2.9.1 Defining the decision criteria

Firstly, a set of decision criteria were applied to identify relevant comparators. Based on the anticipated marketing authorisation, the classes of treatments described in Table 35 were considered to be relevant to the decision problem. Clinical advice suggests that these criteria are likely to be generally applicable for the majority of clinical scenarios (116).

Table 35: Decision criteria for selection of NICE-recommended comparator interventions

Include	Exclude
Chemotherapy Hormone therapy Best supportive care	Surgery with curative intent Radiotherapy (unless palliative)* Immunotherapy Targeted agents Biological therapy

*Included in 'best supportive care'

B.2.9.2 NICE Pathways search

A search was conducted as described in Appendix L, utilising NICE Pathways (https://pathways.nice.org.uk/) to identify relevant comparators meeting the criteria in Table 35.

B.2.9.3 Data aggregation

The search outputs (described in Appendix L) enabled synthesis of an average estimate of the median PFS and OS outcomes likely to be expected with relevant NICE-recommended therapies across the tumour types recruited to the integrated efficacy analysis. Methods for subsequently deriving mean PFS and OS estimates are described in Sections B.3.3.2 and B.3.3.3.

These median and mean survival estimates were applied at an individual patient level, allowing a weighted average to be calculated and providing a naïve indication of outcomes that could potentially be achieved within current clinical practice. However, these do not account for important prognostic factors such as *NTRK*-fusion positive status, or the high frequency of central nervous system involvement within the entrectinib-treated cohort (20.4%). Furthermore, although conducted using a naïve, unadjusted comparison, these outcomes may reflect an overestimation of the PFS and OS likely to be achieved for the comparator, which could lead to a more conservative estimate of entrectinib benefit.

The comparator data derived using this methodology was then compared to integrated efficacy analysis data in a naïve fashion (see section B.3.3 for further details).

B.2.10 Adverse reactions

- **AEs infrequently led to treatment discontinuation ([redacted]) and of these, 3.9% were assessed as related to treatment by the investigator**
- **The safety profile (nature and severity of events) was consistent between the *NTRK* fusion and overall population and between adult and paediatric populations**
- **The vast majority of AEs were Grade 1–2 ([redacted]) and non-serious ([redacted]). Grade 3–4 AEs were experienced by [redacted] of patients, and SAEs by [redacted] of patients, reflective of the advanced disease status under study**
- **Most deaths in the adult population were due to disease progression. AEs that resulted in deaths occurred in [redacted] patients and were reported in the context of worsening or complications of the underlying malignancy, none being considered related to entrectinib by the investigator**
- **[redacted] AE was reported in paediatric patients**
- **Overall, entrectinib is generally safe and well tolerated in children, adolescent, and young adult patients with relapsed or refractory solid tumours during the Phase 1 portion of the study**

Unless otherwise stated, all data is from the Integrated Analysis CSR - Summary of Clinical Safety report (102).

The clinical safety data supporting this submission are derived primarily from three ongoing adult studies ALKA (n=57), STARTRK-1 (n=76), and STARTRK-2 (n=206) and one paediatric study STARTRK-NG in children >4 months of age, adolescents, and young adults (n=16), which in total provide safety data on 355 patients. All patients from Study STARTRK-NG are herein referred to as paediatric patients and all patients from Studies ALKA, STARTRK-1, and STARTRK-2 are referred to as adult patients. Patient safety data from the above mentioned studies have been pooled and analysed collectively as the 'integrated safety population' with a CCOD of 31st May 2018 (patients enrolled up to 30th November 2017). The enrolment cut-off date was set to ensure that patients had at least 6 months of follow-up at the CCOD.

Additional safety data are available in Appendix F.

B.2.10.1 Demographic and baseline characteristics

In the overall integrated safety population (n=355, 45.4% male, 54.6% female), the median age was 55.0 years (range: 4–86 years), the majority of patients were <65 years of age (74.6%) and white (66.4%, 235/354), and 23.2% (82/354) of patients were Asian (Table 35.

The median weight at baseline was 67.55 kg (range: 13.5–130.2 kg) and the median BMI was 23.67 kg/m^2. The majority of adult patients had no history of smoking (57.2%, 183/320); the remaining 42.8% (137/320) were current or previous smokers. The vast majority of adult patients had ECOG performance status of 0 or 1 (91.4%, 310/339), and 8.6% (29/339) of adult patients had ECOG performance status of ≥2. The paediatric patients were graded by Karnofsky (patients >16 years old) or Lansky (patients ≤16 years) performance scores, and the range in baseline score was 70–100.

Table 36: Demographic and baseline disease characteristics of the integrated safety population

	Integrated safety population N=355
Sex, n (%)	
Male	161 (45.4)
Female	194 (54.6)
Median age, years (range)	55.0 (4–86)
Age group, years, n (%)	
<65	265 (74.6)
≥65	90 (25.4)
Race, n (%)	
Asian	82 (23.2)
White	235 (66.4)
Black of African American	16 (4.5)
Other	5 (1.4)
Not reported	16 (4.5)
Mean BSA, m^2 (SD)	1.76 (0.30)
Mean BMI, kg/m^2 (SD)	24.45 (5.36)
ECOG PS, n (%)	n=339
0	140 (41.3)
1	170 (50.1)
2	25 (7.4)
3	3 (0.9)
4	1 (0.3)
Metastatic disease at baseline, n (%)	
Any site	311 (87.6)
CNS lesions*	138 (38.8)

*measure or present CNS lesions, as determined by the Investigator

BMI, body mass index; BSA, body surface area; CNS, central nervous system; ECOG PS, Eastern Cooperative Oncology Group performance status; SD, standard deviation

B.2.10.2 Exposure to entrectinib

As of the CCOD, 355 patients had received at least one dose of entrectinib. Most patients received all their planned doses of entrectinib, with few missed doses; the median number of missed doses was [redacted], Table 37). In the overall integrated safety population, the median duration of exposure to entrectinib was [redacted] months) corresponding to a median of [redacted].

Table 37: Summary of extent of exposure to entrectinib in the integrated safety population

	NTRK adult patients n=68	ROS1 NSCLC adult patients n=134	Other adult patients n=137	All Adult patients n=339	All paediatric patients n=16	All patients N=355
Median treatment duration, months (range)[a]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median no. of cycles (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median no. of missed doses (range)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean cumulative dose, mg (SD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Median dose intensity, % (range)[b]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] Treatment duration is the date of the last dose of study medication minus the date of the first dose plus one day.

[b] Defined as total cumulative dose actually received/total planned dose x 100%. Factors contributing to dose intensity >100% included patients enrolled during the dose finding portion of the Phase I studies who underwent intra-patient dose escalation after determination of the recommended Phase II dose.

B.2.10.3 Integrated safety population – patient status

As of the CCOD, a total of [redacted] in the overall integrated safety population had discontinued treatment. The primary reason for discontinuation was disease progression ([redacted] of those who discontinued]), followed by AEs ([redacted]) and withdrawal by subject ([redacted]).

Up to the CCOD [redacted] of the enrolled population had withdrawn from the study, [redacted] had completed the study, and [redacted] were on study. The most common reason for study withdrawal was death ([redacted] of enrolled patients).

B.2.10.4 Integrated safety population – safety profile

Overall, the safety data indicate that entrectinib has a favourable safety profile and is well tolerated. The overall safety profile observed was generally similar between paediatrics and adults, except where noted. An overview of the safety profile in patients treated with entrectinib in the overall integrated safety population is provided in Table 38

Table 38: Overview of AEs in the integrated safety population

n, (%)	NTRK adult patients n=68	ROS1 NSCLC adult patients n=134	Other adult patients n=137	All adult patients n=339	All paediatric patients n=16	All patients N=355
Patients with AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with related AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with SAE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with related SAE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with Grade ≥3 AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with related Grade ≥3 AE	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with AE leading to discontinuation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with related AE leading to discontinuation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with AE leading to dose reduction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with related AE leading to dose reduction	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with AE leading to drug interruption	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with related AE leading to drug interruption	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients with AE leading to death	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AE, adverse event; SAE, serious AE

B.2.10.5 Common adverse events

Almost [redacted] patients in the overall integrated safety population ([redacted]) experienced at least one AE of any grade. The vast majority were Grade 1–2 ([redacted]) and non-serious ([redacted]). The most frequently reported events were from the system organ class (SOC) of [redacted]

[redacted]

The most frequently reported (≥25% of patients) AEs by preferred term (PT) were [redacted]. Please refer to Appendix F for the full data table.

B.2.10.6 Treatment-related adverse events

[redacted] patients in the overall integrated safety population had at least one AE that was considered by the investigator to be related to entrectinib treatment. The most frequently reported (≥10% of patients) treatment-related AEs were [redacted].

B.2.10.7 Grade 3–4 adverse events

Grade 3 or 4 AEs were experienced by [redacted] of patients in the overall integrated safety population, of which about [redacted] had an event that was assessed by the investigator as related to entrectinib.

The most frequently reported (≥2% of patients) Grade 3 or 4 AEs by PT were [redacted]. Please refer to Appendix F for the full data table.

B.2.10.8 Deaths

Deaths from adverse events occurred in [redacted] patients; none of which were assessed by the investigator as related to entrectinib. All deaths occurred in the adult population. Please refer to Appendix F for the full data table.

There were a total of [redacted] deaths in the overall integrated safety population. The rate of deaths that occurred within 30 days of the last dose of entrectinib ([redacted]) was similar to the rate of deaths that occurred more than 30 days after the last dose of entrectinib

([redacted]) (Table 39). The most common reason for death was [redacted], which accounted for [redacted] of all deaths.

Table 39: Deaths by cause

n, (%)	NTRK adult patients n=68	ROS1 NSCLC adult patients n=134	Other adult patients n=137	All adult patients n=339	All paediatric patients n=16	All patients N=355
Total number of deaths	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Death within 30 days of last dose of entrectinib						
Total number of deaths	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Unknown	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Death more than 30 days after last dose of entrectinib						
Total number of deaths	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progressive disease	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Other	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Unknown	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Nb. Cause of death is defined differently for each study: ALKA - 'Progressive Disease' if selected by investigator, 'Unknown' if selected by investigator or no cause given, 'Other' for any other reason; STARTRK-1 - Cause of death was not collected (all 'Unknown'); STARTRK-2 and STARTRK-NG - 'Progressive Disease' if death is related to cancer, 'Other if death is not related to cancer, 'Unknown' if death has unknown relation to cancer.

B.2.10.9 Serious adverse events

In the overall integrated safety population, [redacted] patients experienced at least one SAE. Treatment-related SAEs were reported in [redacted] patients. The most frequently reported SAEs regardless of causality by SOC (≥5% of patients, any grade) in the overall integrated safety population were as follows:

Respiratory thoracic and mediastinal [redacted]

Infections and infestations [redacted]

Nervous system disorders [redacted] the patients with these events were noted to have brain metastases at baseline.

No particular pattern was observed in the type and frequency of SAEs reported. A [redacted]) of paediatric patients experienced SAEs compared to adults. There was no SAE with an incidence that was [redacted]. Please refer to Appendix F for the full data table.

B.2.1-.10 Adverse events leading to treatment discontinuation, dose interruptions and dose reductions

Entrectinib was well-tolerated and there was a small number of patients [redacted] with AEs leading to study drug discontinuation in the overall integrated safety population. AEs leading to withdrawal were reported across a variety of SOCs with the most frequently reported (≥1% of patients) being [redacted] (2.0% each), and [redacted], (1.1% each). There was no a predominant AE that led to withdrawal of entrectinib (58).

A total of [redacted] patients in the overall integrated safety population experienced at least one AE that led to a dose interruption. By PT, the most frequently reported AEs leading to entrectinib dose interruption (≥1% of patients) were [redacted] (58).

A total of [redacted] patients in the overall integrated safety population experienced at least one AE that led to a dose reduction. By PT, the most frequently reported AEs leading to entrectinib dose reduction (≥1% of patients) were [redacted] (58).

B.2.11 Ongoing studies

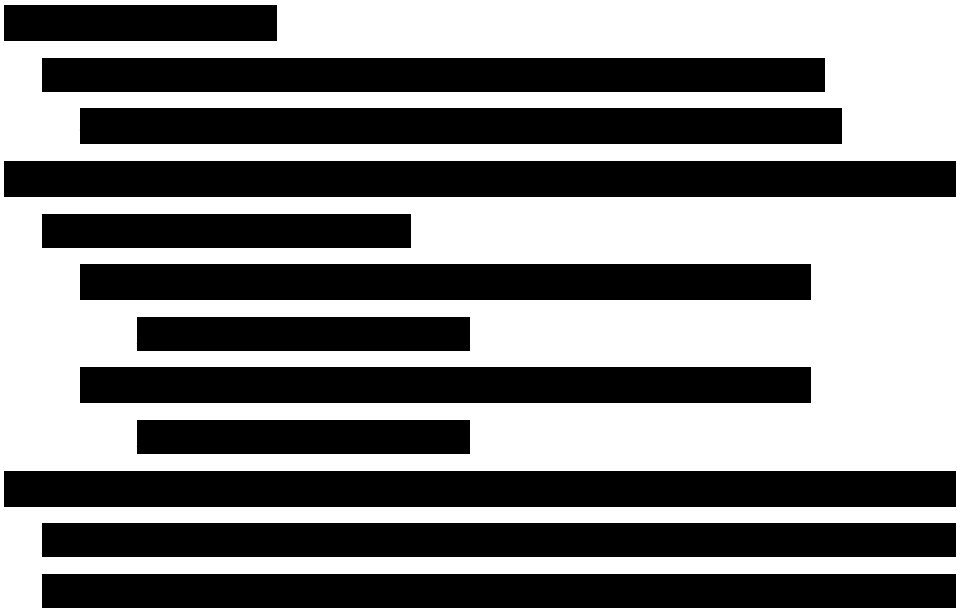
Survival follow-up of the patients with NTRK fusion-positive solid tumours treated with entrectinib in the studies listed in Table 7 are ongoing. Plans for the next analysis of the integrated efficacy analysis (from ALKA, STARTRK-1 and STARTRK-2) are currently being discussed with regulators.

B.2.11.1 Additional data collection proposals

The process of developing this submission has identified a number of challenges for health technology assessment including generalisability of the trial population, understanding baseline clinical outcomes, and the assessment of a highly heterogeneous population, which are likely to be common issues for other multiple tumour-agnostic indications nearing marketing authorisation within the next 3–5 years. For this reason, a proposal has been put forward for entrectinib to enter the Cancer Drugs Fund under a commercial access agreement to collect further data and reduce key uncertainties arising during the course of the appraisal.

A data collection agreement taking into account for the key clinical and cost-effectiveness uncertainties identified through the appraisal process will therefore be formulated with input from NICE, NHS England, Public Health England and Roche.

While the individual data collection activities and key uncertainties are expected to evolve as the appraisal proceeds, Roche proposes the following provisional concepts:



These data collection proposals should be considered provisional until the key clinical and cost-effectiveness uncertainties have been identified through the appraisal process. It should also be noted that some of the activities proposed are not yet confirmed to be taking place.

B.2.12 Innovation

Entrectinib is a novel targeted therapy addressing the underlying cause of the disease: inhibition of dysregulated tyrosine kinase activity arising from NTRK gene fusions (101). It has the ability to cross the BBB and remain within the CNS (12) and is the only NTRK fusion inhibitor in development with a once-per-day dosing regimen (11). As one of the first tumour-agnostic indications to be appraised by NICE, entrectinib represents a step-change in the treatment of cancer, changing the focus from the origin of the primary cancer to the underlying oncogenic driver, regardless of histology, and providing important benefits to a group of patients with tumour types where treatment options have been historically limited, such as in MASC and pancreatic cancer. It is a CNS-active NTRK inhibitor where the clinically equivalent responses were observed regardless of the presence or absence of CNS disease responses (22% of all enrolled patients in the integrated efficacy analysis were positive for CNS metastases). These results mean that patients with CNS tumours who previously had limited treatment options and poor prognosis could benefit from the availability of molecularly targeted CNS-active treatment options. Utilising novel genomic

technologies such as NGS to identify *NTRK* fusion-positive patients may also provide benefits to patient health and cost efficiencies for health care systems (2).

Regulatory bodies have formally recognised the innovative nature of entrectinib by granting entrectinib FDA Breakthrough Therapy Designation in May 2017 and EMA Priority Medicine Designation in October 2017 (117, 118). Furthermore, in December 2018, entrectinib was granted a Promising Innovative Medicine (PIM) designation by the Medicines and Healthcare Products Regulatory Agency (MHRA), demonstrating that entrectinib is a novel treatment for patients with NTRK-positive tumours regardless of age and the site of tumour.

There is a clear unmet medical need for molecularly-targeted treatment options, including those that are CNS active, that reduces the risk of over treatment with the offer of greater efficacy and lesser toxicity relative to conventional cytotoxic chemotherapy, for the rare population of adult and paediatric cancer patients with relapsed or refractory *NTRK* fusion-positive solid tumours. Taken together, the available data indicate that when an *NTRK* fusion is present, the anti-tumour activity of entrectinib is agnostic to tumour histology, patient age, and extracranial or intracranial disease.

B.2.13 Interpretation of clinical effectiveness and safety evidence

B.2.13.1 Integrated efficacy analysis data

The integrated efficacy analysis data pooled from ALKA, STARTRK-1, and STARTRK-2, has demonstrated that entrectinib provides a clinically meaningful benefit for adult patients whose tumours harbour *NTRK* gene fusions. The integrated safety data pooled from ALKA, STARTRK-1, STARTRK-2, and STARTRK-NG demonstrated that entrectinib is CNS active, well-tolerated and has a manageable safety profile in patients whose tumours harbour *NTRK* gene fusions. The STARTRK-NG data showed that antitumour activity was observed in the majority of paediatric patients with solid tumours harbouring *NTRK* fusions and demonstrated a generally safe and well-tolerated safety profile in children, adolescent and young adult patients enrolled in this study. Overall, these studies show that entrectinib is a transformative medicine that can address a great unmet need in this rare population with no established standard of care.

The BICR-ORR was 57.4% (95% CI: 43.2%, 70.8%) including four patients who achieved a CR (7.4%). The majority of objective responses were achieved at the first tumour assessment after commencing entrectinib treatment (end of Cycle 1). Responses were

durable with a median BICR-DOR in responders of 10.4 months (95% CI: 7.1; NE) with over half of responders having responses lasting longer than 6 months.

The lower limit of the 95% CI for the BICR-ORR (43.2%, 70.8%) excluded (was greater than) 30% and thus exceeded the historical response rates (typically <30%) for available treatments in later lines or salvage therapy in patients with advanced or metastatic solid tumour types that have been reported to harbour *NTRK* fusions.

The median PFS of 11.2 months was similar to the mDOR of 10.4 months and compare favourably with currently approved treatment options for patients with NSCLC or CRC who have failed previous lines of therapy (119-122), or tumour types for which there are no standards of care (e.g. MASC and soft tissue sarcomas) (123, 124). For patients with CNS metastases at baseline, the KM estimated median intracranial PFS of 14.3 months (95% CI: 5.1, NE) reflected the durability of the treatment effect of entrectinib on CNS metastatic lesions. The median OS of 20.9 months (95% CI: 14.9, NE) was maintained at CCOD [redacted] with a median OS of [redacted]. These data are encouraging when considering that median OS reported in clinical trials for patients treated with front-line chemotherapy regimens for NSCLC or breast cancer with brain metastases at diagnosis is typically one year or less (125-127). However, PFS and OS data should be interpreted with caution considering the limitations already described in this submission. Nonetheless, the consistency of these secondary endpoints provides supportive information about the effectiveness of entrectinib demonstrated by ORR and DOR. Survival follow-up of the patients with *NTRK* fusion-positive solid tumours treated with entrectinib is ongoing.

In the STARTRK-NG trial, the overall ORR was 18.8% (95% CI: 4.05, 45.65) and the overall CBR was 18.8% (95% CI: 4.05, 45.65) for the analysis population in Phase 1 portion of the study) (104). In the updated analysis [redacted] with diverse tumour types including primary brain tumours in which *NTRK* gene fusions, as well as *ROS1* or *ALK* gene fusions were detected.

The efficacy benefit described above occurred in the context of a well-tolerated and manageable safety profile considering the advanced nature of the disease in the patient population under study. Of the overall AEs reported in the integrated safety population, the vast majority were Grade 1–2 (90.6%) and non-serious (96.3%). Grade ≥3 AEs were experienced by 61.1% of patients, and SAEs by 38.6% of patients, reflective of the advanced disease status under study. The tolerability of entrectinib was evident by the low

discontinuation rates. Discontinuation of entrectinib due to safety reasons (AEs) was relatively infrequent (8.5%) and accounted for a small proportion of all patients overall who discontinued entrectinib. Most AEs requiring intervention could be adequately managed with dose interruptions (withholding of dose), dose reduction, and/or supportive care.

There are some remaining scientific questions of interest related to entrectinib's activity against *NTRK* fusion-positive tumours, namely the ability to extrapolate activity to tumour types where there is no clinical or nonclinical experience to date, mechanisms of resistance to entrectinib, and the impact of co-occurring oncodrivers on the treatment effect of entrectinib.

B.2.13.2 Study design of the entrectinib trials

Regulatory bodies (FDA and EMA) and the oncology community (European Society of Molecular Oncology (ESMO)) have acknowledged the challenges in assessing novel treatments that target rare genomic alterations and have accepted the use of single-arm basket trials across tumour types as sufficient evidence for approval (128, 129). Basket trials were used for the investigation of entrectinib due to the low prevalence of NTRK fusions and associated challenges in identifying patients when routine *NTRK* mutation testing is not in place, the heterogeneity of the *NTRK* fusion-positive population and a lack of a consistent comparator. The pooling of data in the STARTRK-2 trial for patients harbouring any NTRK1, NTRK2 or NTRK3 rearrangement was justified because the fusion proteins derived from these genes share the same tyrosine kinase biology and were expected to behave similarly (130). While the integrated analyses were not prespecified in the individual study protocols, considering the rarity of the patient population, an integrated statistical analysis plan was developed to maximise the number of gene fusion-positive patients available for safety and efficacy analyses, including patients from the Phase I studies. This proposal to pool safety and efficacy from the clinical studies was endorsed by the regulatory health authorities because of the rare disease setting.

B.2.13.3 End-of-life criteria

Entrectinib meets end-of-life criteria compared to the current standard of care for tumour histologies represented in the entrectinib integrated efficacy analysis, with due consideration given to its position in the care pathway as dictated by its proposed licence, that is in [redacted] (Table 40).

The approach taken to establish end-of-life criteria is derived from the overall survival of the entrectinib integrated efficacy analysis population and the matched comparator population established in the economic model. This was necessary due to the absence of a comparator arm in the entrectinib trials.

The benchmark for end-of-life is therefore established by a matched hypothetical comparator population consisting of the same tumour histology proportions as in the entrectinib integrated efficacy analysis population treated with medicines that are currently available NICE baseline funded medicines. Please refer to section B.2.9 and B.3.3.1 for further details on how the comparator arm was derived.

Table 40: End-of-life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Entrectinib is expected to be licensed for any tumour histology harbouring an *NTRK* gene fusion in patients [redacted] (11). According to the approach taken in the model, the comparator-matched population have a collective median OS of 15.7 months.	B.3.3.3, page 101
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	After a median survival follow-up of 12.9 months, only 29.6% of patients in the entrectinib integrated efficacy analysis population had died, meaning median OS has not yet been reached. However, extrapolation of existing OS data suggests an anticipated median OS of 26.5 months. This would provide a median OS benefit of 10.8 months.	B.3.3.3, page 101

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

A SLR was conducted to identify published cost-effectiveness studies for treatment of patients in *NTRK* fusion-positive solid tumours. To ensure that all relevant publications were captured - the population of interest was kept broad and included adult patients with advanced or metastatic solid tumours, regardless of line of therapy.

Detailed descriptions of the search strategy and extraction methods are provided in Appendix D and G. No previous cost-effectiveness studies have been identified.

B.3.2 Economic analysis

A systematic literature review was conducted to identify relevant previous cost-effectiveness studies. As no economic evaluations have previously been reported which align with the decision problem, a de novo economic model was developed.

B.3.2.1 Patient population

The patient population included in the base-case cost-effectiveness covers adult patients with advanced or metastatic *NTRK* fusion-positive solid tumours, who have progressed following prior therapies or for whom no acceptable standard therapy exists. This reflects the population included in the integrated efficacy analysis.

The baseline demographics of the patients within the integrated efficacy analysis, and therefore informing the base case cost-effectiveness analysis, are summarised in Table 41.

Table 41: Cost-effectiveness analysis - patient population

Characteristic	Description	NTRK efficacy cohort (n=54)
Age	Median	57.5 years old
Gender	Female	59.3%
Race	White	79.6%
	Asian	13.0%
Performance status	ECOG 0	42.6%
	ECOG 1	46.3%
	ECOG ≥2	11.1%
Smoking status	Never-smoker	56.6%
Baseline CNS metastases	Present	20.4%

CNS, central nervous system; ECOG, Eastern Cooperative Oncology Group

B.3.2.2 Model structure

In the absence of previous cost-effectiveness studies, a de novo model was developed. The economic evaluation was developed in Microsoft Excel, as a partitioned survival model which evaluates the cost-effectiveness of entrectinib for the treatment of *NTRK*-fusion solid tumours when compared with established practice (standard of care).

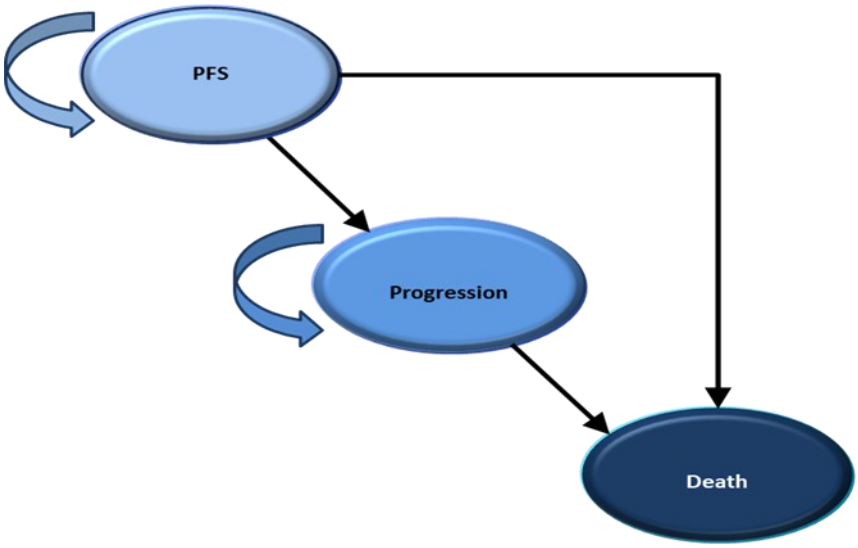
Partitioned survival modelling involves partitioning OS into states of interest. The three states of interests are PFS, PD, and the absorbing health state of death (Figure 16). The partitioning of OS is achieved using PFS. The two trial outcomes PFS and OS in the integrated efficacy analysis are each modelled directly using parametric regression to allow for extrapolation.

However, this approach does not consider post-progression survival directly: instead, the mean time in PD is derived from the difference in the area under the two survival outcomes.

Similarly, the partitioned survival principle is applied between time-to-off-treatment (TTOT) and OS in order to assess the states on and off treatment. This reduces the number of assumptions required when assessing and extrapolating immature survival data from the entrectinib clinical trial results.

This model structure was selected, as per NICE decision support unit (DSU) guidance (131), in order to allow for full use of the mature PFS and OS study data from IMpower150 and to be able to incorporate external evidence for additional comparators in the economic model.

Figure 16: Partitioned survival model structure



Patients enter the model in a progression-free health state. Transition from the PFS state to the progression state is defined according to the RECISTv1.1 criteria, as is commonly used within clinical trials (132). Patients may transition to Death, the absorbing health state, from the progression state or directly from the PFS state.

The model inputs (efficacy, safety/tolerability) for the intervention arm were based on the results of the integrated analysis presented in Section B.2.6. In terms of the modelling of established practice, the health states are still based on the partition survival principle but the progression free survival and the overall survival estimates are not derived from extrapolated KM curves. Synthesis of different tumour types and treatment lines KM curves would not have been feasible. Therefore, reported median PFS and OS from the literature, for each tumour type within the entrectinib trials, has been converted to mean values. This conversion is based on an exponential extrapolation assumption, in order to simulate an exponential area under the curve.

Using this method, an effort is made to simulate the PFS and OS benefit of the comparator as applied on the entrectinib cohort, to provide a benchmark for comparison.

Duration of treatment for entrectinib was assumed to be aligned with the anticipated marketing authorisation, i.e. until disease-progression or unacceptable toxicity. For the base case, treatment with standard of care was also assumed to continue according to relevant SmPC guidance.

Costs and health-related utilities are allocated to each health state and multiplied by state occupancy to calculate the weighted costs and quality adjusted life years (QALYs) per cycle.

The economic model base case uses a time horizon of 30 years, which was considered to be sufficiently long enough to reflect all important differences in costs or outcomes between the technologies being compared. This takes into consideration: 1) the median age of the patient population in the cost-effectiveness analysis of 57.5 years and 2) the maximum plausible impact of improved outcomes following treatment with entrectinib. Scenario analyses are provided that consider shorter time horizons.

The model has been designed to use a weekly cycle, with the proportion of patients in each health state calculated each week. Transition between health states can occur at any time within the cycle. To account for the over or under estimation of transitions occurring at the beginning or end of the cycle, half-cycle corrections were applied to each time interval in the Markov trace sheets of the model. This is also consistent with previous NICE single

technology appraisals (STAs) in the oncology area. details the main features of this economic analysis.

Table 42: Features of the economic analysis

	Current appraisal	
Factor	**Chosen values**	**Justification**
Time horizon	30 years	This period is expected to allow for consideration of all costs and outcomes for the relevant population. Sensitivity analyses have been conducted to test this assumption.
Source of clinical effectiveness data	Entrectinib: integrated analysis of clinical trials Established management: Search of NICE Pathways website to identify technology appraisals for relevant NICE-recommended comparators	Entrectinib: aligned with methods guide Established care: formal evidence synthesis methods (e.g. network meta-analysis) are not feasible due the high number of relevant chemotherapies. Weighted outcomes for the comparators have been derived from data extracted from NICE appraisals.
Treatment waning effect?	None	No treatment waning is a plausible but conservative assumption based on the method of administration and mechanism of action of entrectinib.
Source of utilities	Entrectinib: integrated analysis Established management: systematic search (see Appendix H)	Entrectinib: aligned with NICE methods guide (2013) Established care: a deviation was required from the NICE methods guide due to the high number of search results identified (>1m). Therefore, where possible utilities were derived from previous NICE appraisals in relevant tumour types.
Source of costs	Systematic search of relevant NICE guidance Only NHS and social care costs included	In line with the NICE 'Guide to the Methods of Health Technology Appraisals 2013'
Were health effects measured in QALYs; if not, what was used?	Yes, measured in QALYs	In line with the NICE 'Guide to the Methods of Health Technology Appraisals 2013'
Discount rate for costs and effects	3.5%	In line with the NICE 'Guide to the Methods of Health Technology Appraisals 2013'

Cycle length	Weekly	In line with the NICE 'Guide to the Methods of Health Technology Appraisals 2013'
Half-cycle correction	Applied	In line with the NICE 'Guide to the Methods of Health Technology Appraisals 2013'
Perspective (NHS/PSS)	UK NHS	In line with the NICE 'Guide to the Methods of Health Technology Appraisals 2013'

PSS, personal social services; QALYs, quality-adjusted life years

Intervention technology and comparators

The final scope for this appraisal includes entrectinib in comparison with 'established management'.

Entrectinib is administered as an oral treatment, with a full pack of 90 capsules administered once per month (3 per day). These capsules are assumed to be provided at a scheduled monthly outpatient appointment at a specialist cancer centre. Treatment is assumed to continue until a patient experiences disease progression, unacceptable toxicity or death.

For the purposes of modelling established management, a simulated chemotherapy comparator was created by producing an average of clinical outcomes derived from NICE appraisals weighted by the proportions of tumour types represented in the integrated analysis population (see B.2.9 and Appendix L). This aimed to provide an estimate of the outcomes anticipated for the tumour types represented within the integrated efficacy analysis. Treatment-specific details such as route of administration (e.g. oral, intravenous) and frequency of scheduled clinic visits was applied at a tumour-specific level, before costs were weighted and aggregated to reflect an 'average' for the trial population. No adjustments were made to the weighted comparator to reflect important prognostic factors such as the influence of *NTRK*-fusions, or the presence of central nervous system (CNS) metastases. However, the influence of changes in prognosis for this synthetic comparator were tested extensively using sensitivity analyses (see Section B.3.8).

B.3.3 Clinical parameters and variables

B.3.3.1 Incorporation of clinical data into the economic model

Intervention

The primary data source for entrectinib within the economic model is the integrated efficacy analysis of three clinical studies, the phase II STARTRK-2 study, and the phase I STARTRK-1 and ALKA-372-001 studies (101). These studies are the data source for the clinical

outcomes (progression-free and overall survival), adverse events and quality of life for entrectinib (the intervention).

Due to the rarity of *NTRK* gene fusions and the myriad possible comparator products, it was not possible to conduct randomised study with a control arm. To account for this, a simulated comparator arm has been synthesised as described in B.2.9 and Appendix L.

Parametric extrapolation of entrectinib OS, PFS and TTOT from the integrated efficacy analysis population was required, for the proportion of patients that had not progressed or died within the follow-up period of the trials (after 12.9 months' follow-up for both PFS and OS).

NICE DSU guidance (Technical Support Document 14) (133) was therefore followed to identify base case parametric survival models for OS, PFS and TTOT.

All parametric models were assessed against the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for statistical fit to the observed data. To aid the assessment of the appropriateness of the tail a summary of the extrapolation in the trend of the hazard was considered. Curves were visually inspected and validated against clinical expert opinion to help identify the most plausible survival model.

The parametric distribution which displayed the best statistical fit, with a plausible hazard trend, was selected for use in the base case analysis.

Comparator

The methods used to define comparators and derive relevant comparator data are described in section B.2.9. Once comparators were chosen, the following steps were taken to synthesise a comparator arm and incorporate it into the model:

1. Median PFS and OS outcomes for each comparator within a specific tumour type were averaged to ascertain overall outcomes for that tumour.

2. These averaged outcomes were applied at an individual patient level in the integrated analysis population.

3. These median outcomes were then converted to mean values through exponential extrapolation, in alignment with the recommendations of NICE Technical Support Document 14, which states that in cases where only published summary statistics are available, the conversion of median to exponential mean represents a reasonable

approach where individual patient data are lacking (134). The formula used to convert the median survival to exponential mean is 1/(-LN(0.5)/median).

4. The estimated mean PFS and OS for each tumour type was then averaged, and weighted by the proportion of tumour types present in the pooled entrectinib trials NTRK+ cohort.

5. This weighted comparator arm was then compared in a naïve fashion with the entrectinib integrated analysis population data, which allows a Partitioned Survival Analysis to be conducted as described in section B.3.2.

Choice of comparator therapy and associated outcomes for most tumour types in the trial were validated with a clinical expert specialising in each given tumour type (116).

The resulting survival estimates for the comparator cohort are shown in Table 43.

Table 43: Estimated exponential mean for comparator PFS and OS

Tumour groups	PFS median	PFS SE	PFS exponential mean	OS median	OS SE	OS exponential mean
	(Months)					
CRC	2.63	0.150	3.80	9.07	0.150	13.08
MASC	4.35	0.150	6.27	13.80	0.150	19.91
Papillary thyroid	4.55	0.150	6.56	30.95	0.150	44.65
Other thyroid	4.55	0.150	6.56	30.95	0.150	44.65
Squamous NSCLC	3.75	0.150	5.41	10.65	0.150	15.36
Non-squamous NSCLC	3.75	0.150	5.41	10.65	0.150	15.36
Pancreatic	5.20	0.150	7.50	8.80	0.150	12.70
Sarcoma	3.90	0.150	5.63	14.30	0.150	20.63
Neuroendocrine	8.03	0.150	11.58	39.61	0.150	57.14
Secretory Breast	3.03	0.150	4.36	12.18	0.150	17.56
Triple negative breast	3.03	0.150	4.36	12.18	0.150	17.56
Other	4.35	0.150	6.27	17.23	0.150	24.86

B.3.3.2 Entrectinib progression-free survival extrapolation

Extrapolation beyond the pooled entrectinib trials clinical follow-up period was performed by fitting a parametric distribution to the observed time to event data from the pooled entrectinib trials trial for the relevant basket population (*NTRK*-fusion positive).

Several parametric distributions were fitted to the time to event data in order to extrapolate PFS beyond the observation period: Exponential, Weibull, Log-Logistic, Log-normal, Generalised gamma and Gompertz.

The statistical fit and hazard trend, are described in Table 44. Landmark PFS rates for each parametric distribution are shown in Table 45. Individual graphical distributions are shown in Figure 17 and Figure 18. An exponential distribution was selected for the base case analysis, as it represents a conservative but statistically and clinically plausible estimate of progression-free survival for entrectinib-treated patients.

Table 44: Statistical goodness-of-fit and hazard trend for PFS

Parametric distribution	AIC	BIC	Hazard trend*
Exponential	220.7	222.6	Stable
Weibull	221.4	225.3	Increasing
Log-normal	223.0	227.0	Decreasing
Generalised gamma	223.4	229.3	Increasing
Log-logistic	222.6	226.6	Decreasing
Gompertz	221.6	225.5	Increasing

Table 45: Landmark PFS rates for each parametric distribution

Parametric distribution	2 years	5 years	10 years	15 years	20 years	Clinically plausible?
Exponential	22%	2.3%	0.1%	0.0%	0.0%	Yes
Weibull	17%	0.5%	0.0%	0.0%	0.0%	Yes
Log-normal	25.7%	8.3%	2.6%	1.2%	0.6%	No
Generalised gamma	16.2%	0.3%	0.0%	0.0%	0.0%	Yes
Log-logistic	24.1%	7.8%	3.0%	1.7%	1.1%	No
Gompertz	14.4%	0.0%	0.0%	0.0%	0.0%	Yes

Figure 17: Parametric extrapolation of entrectinib PFS: exponential distribution

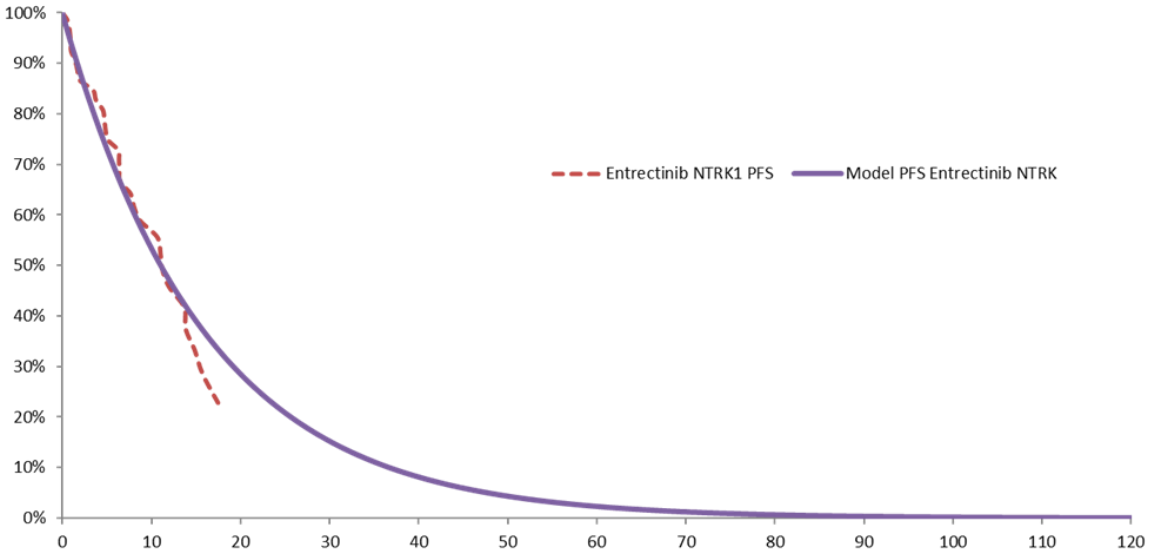
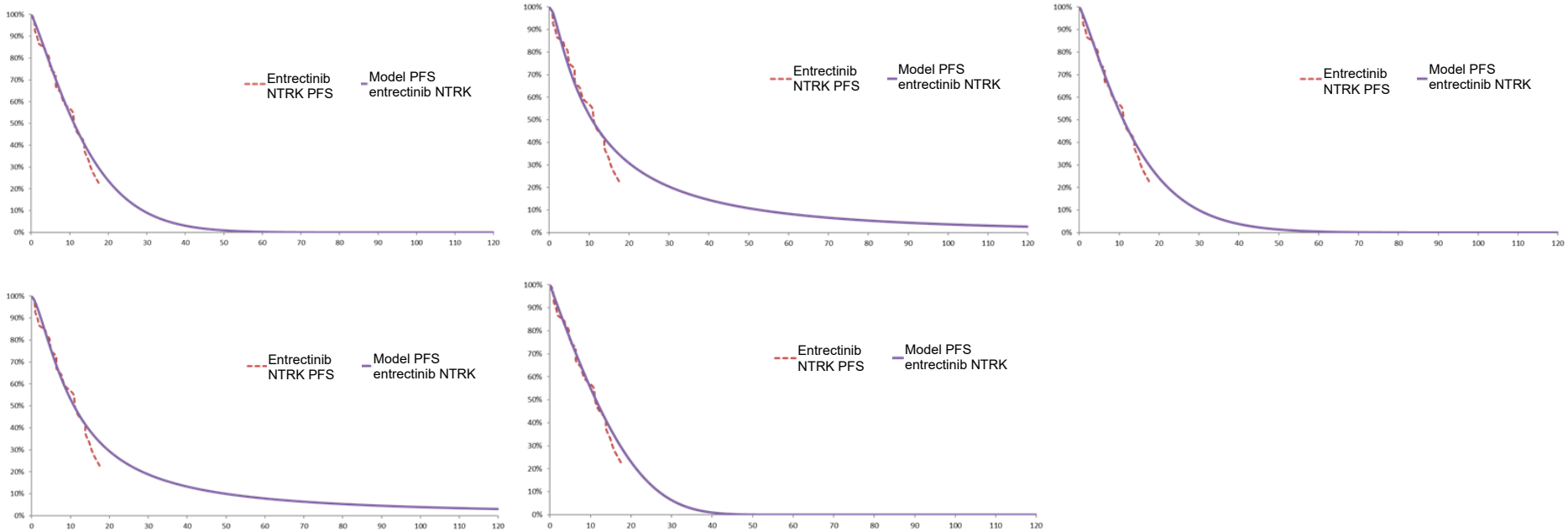


Figure 18: Alternative PFS distributions (1st row, left to right: Weibull, Log-Normal, Gamma; 2nd row, left to right: Log-logistic, Gompertz)



B.3.3.3 Entrectinib overall survival extrapolation

As was conducted for PFS, several parametric distributions were fitted to the time to event data in order to extrapolate OS beyond the observation period: Exponential, Weibull, Log-Logistic, Log-normal, Generalised Gamma and Gompertz.

The statistical fit and hazard trend are described in Table 46. Landmark OS rates for each parametric distribution are shown in Table 47. Individual graphical distributions are shown in Figure 19 and Figure 6. An exponential distribution was selected for the base case analysis, as it demonstrated the best statistical fit to the observed data.

Table 46: Statistical goodness-of-fit and hazard trend for OS

Parametric distribution	AIC	BIC	Hazard trend*
Exponential	150.5	152.5	Stable
Weibull	151.2	155.2	Increasing
Log-normal	153.3	157.3	Decreasing
Generalised gamma	152.7	158.6	Increasing
Log-logistic	152	156	Stable/ Decreasing
Gompertz	150.5	154.5	Increasing

Table 47: Landmark OS rates for each parametric distribution

Parametric distribution	2 years	5 years	10 years	15 years	20 years	30 years	Clinically plausible?
Exponential	53%	20.7%	4.3%	0.9%	0.2%	0.0%	Yes
Weibull	47.5%	9.0%	0.3%	0.0%	0.0%	0.0%	Yes
Log-normal	55.1%	29.6%	15.3%	9.5%	6.5%	3.6%	No
Generalised gamma	45.7%	0.0%	0.0%	0.0%	0.0%	0.0%	No
Log-logistic	52.1%	22.3%	9.7%	5.7%	3.9%	2.2%	No
Gompertz	44.3%	0.0%	0.0%	0.0%	0.0%	0.0%	No

Figure 19: Parametric extrapolation of OS: exponential distribution

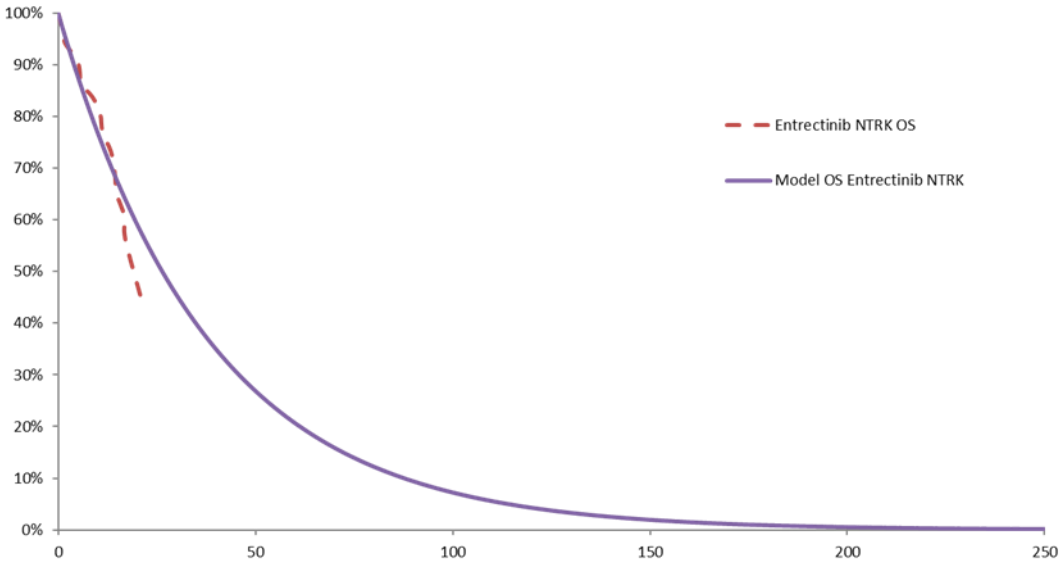
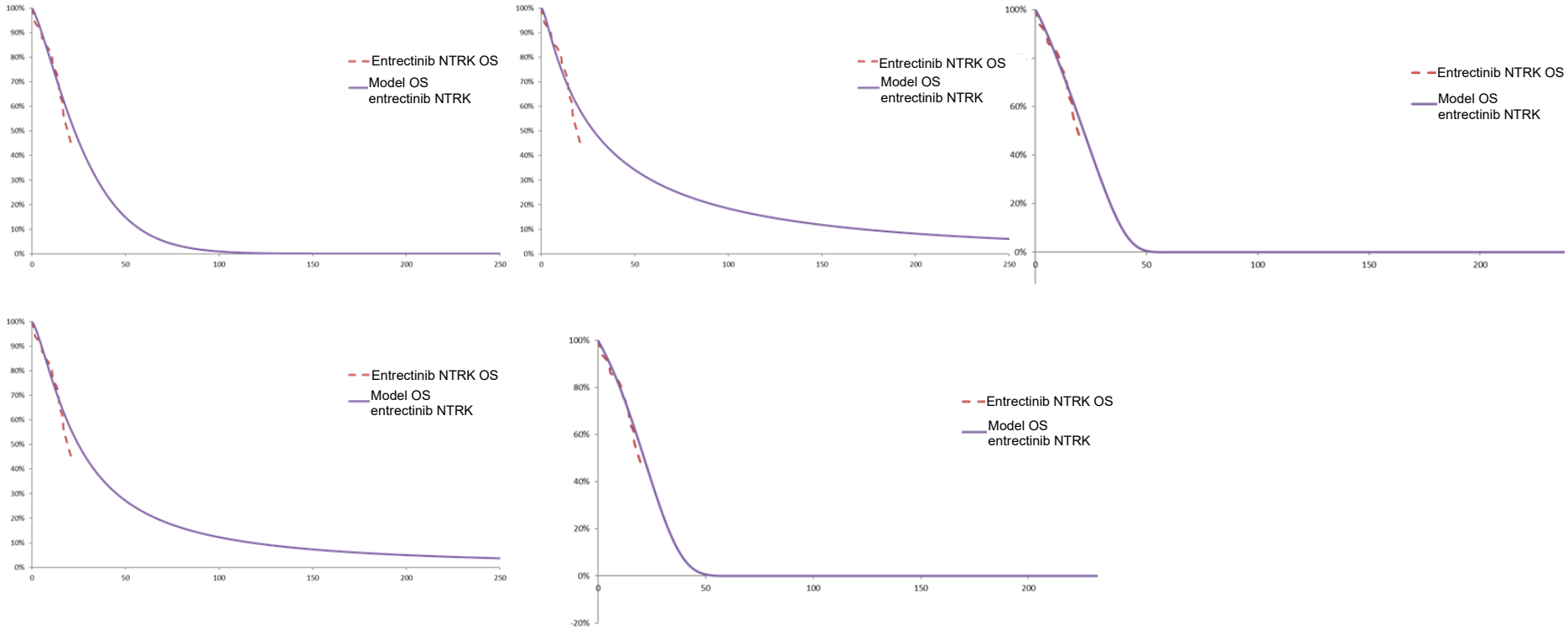


Figure 20: Alternative OS distributions (1st row, left to right: Weibull, Log-normal, Gamma; 2nd row, Log-logistic, Gompertz)



B.3.3.4 Entrectinib time to off treatment (TTOT) extrapolation

As was conducted for PFS and OS, several parametric distributions were fitted to the time to event data in order to extrapolate TTOT beyond the observation period: Exponential, Weibull, Log-Logistic, Log-normal, Generalised Gamma and Gompertz.

The statistical fit and hazard trend are described in Table 48. Landmark TTOT rates for each parametric distribution are shown in Table 49. Individual graphical distributions are shown in Figure 21 and Figure 22. An exponential distribution was selected for the base case analysis, as it demonstrated the best statistical fit to the observed data.

Table 48: Statistical goodness-of-fit and hazard trend for TTOT

Parametric distribution	AIC NTRK	BIC NTRK	Hazard trend*
Exponential	236.4	238.4	N/A
Weibull	238.3	242.3	N/A
Log-normal	241.3	245.3	N/A
Generalised gamma	240	246	N/A
Log-logistic	240.1	244.1	N/A
Gompertz	238.1	242.1	N/A

*Hazard trend indicates the trend observed in the parametric extrapolation after Month 10 with Decreasing indicating a decreasing risk of event and increasing indicating an increasing risk of the event. Since TTOT for the comparator arm cannot be modelled, hazard trend does not apply.

Table 49: Landmark TTOT rates for each parametric distribution

Parametric distribution	2 years	5 years	10 years	15 years	20 years
Exponential	22.4%	2.4%	0.1%	0.0%	0.0%
Weibull	21.2%	1.8%	0.0%	0.0%	0.0%
Log-normal	29.4%	12.2%	5.0%	2.8%	1.7%
Generalised gamma	17.4%	0.0%	0.0%	0.0%	0.0%
Log-logistic	27.5%	10.8%	4.9%	3.0%	2.1%
Gompertz	18.5%	0.2%	0.0%	0.0%	0.0%

Figure 21: Parametric extrapolation of TTOT: exponential distribution

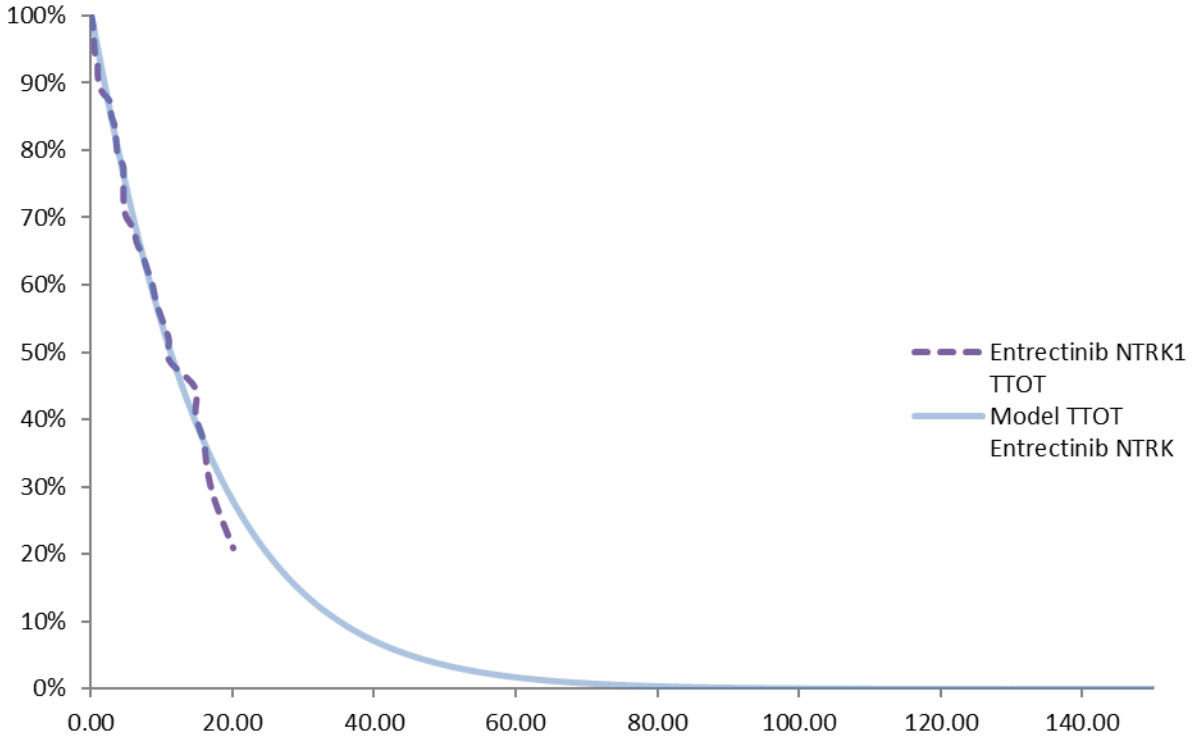
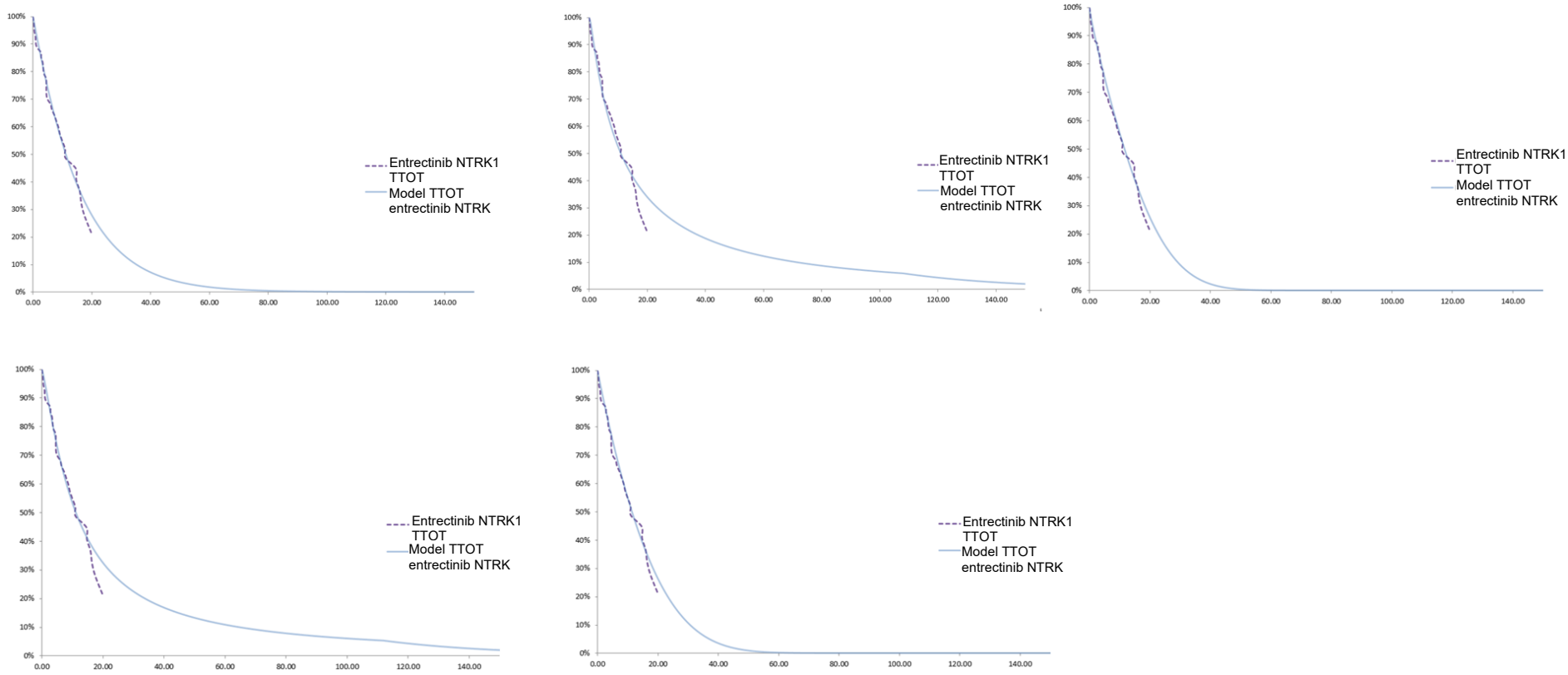


Figure 22: Alternative TTOT distributions (1st row, left to right: Weibull, Log-normal, Gamma; 2nd row, Log-logistic, Gompertz)



B.3.3.5 Comparator PFS, OS, and TTOT

Comparator PFS and OS data were extrapolated (to exponential means) as described in section B.3.3.1. Due to the nature of the construction of the comparator arm and the lack of patient-level data, TTOT data is not available. Consequently, for the purposes of the model, comparator TTOT is assumed to be equivalent to undiscounted mean PFS.

B.3.3.6 Validation of clinical parameters

It is acknowledged that given the novelty of the *NTRK*-fusion indication, and the absence of comparative natural history data, uncertainty exists with the extrapolation of both progression-free and overall survival for entrectinib, as well as accurately estimating outcomes for the comparator population. As such, a range of sensitivity analyses have been reported in Section B.3.8. In addition, [redacted] regarding both entrectinib and an NTRK-fusion positive population untreated with entrectinib (see Section B.2.11.1).

Validation of extrapolations

Validation of entrectinib extrapolations with clinical experts is challenging, as any given clinician is typically likely to be an expert in only a subset of tumour histologies represented in the integrated analysis population, and there is limited knowledge of entrectinib in the clinical community. As such, clinical validation of PFS and OS extrapolations was sought from investigators at two UK sites (The Christie NHS Foundation Trust and Cambridge University Hospitals NHS Foundation Trust) participating in the STARTRK-2 study, who specialise in early phase research of solid malignancies, with clinical interests in lung and genito-urinary cancers. This was conducted through their visual inspection of all six extrapolations of the PFS and OS curves for entrectinib; emphasis was placed on OS extrapolation due to its importance in the model. Summaries of their assessments of clinical plausibility are shown in Table 45 and Table 47. Both investigators concluded that the exponential extrapolation for PFS resulted in a clinically plausible prediction of outcomes across the time horizon, alongside Weibull, Gamma and Gompertz. For overall survival, both investigators were of the view that the exponential and Weibull extrapolations were clinically plausible, while others were either too optimistic (log normal and log logistic) or too pessimistic (Gamma and Gompertz). All distributions are explored in scenario analyses.

Validation of comparator choice and data

As previously discussed, comparators have been chosen based on anticipated clinical positioning of entrectinib and the relevant NICE recommended therapies (or best supportive care) available in the clinical care pathways for the tumour types included in the integrated analysis population.

The specific treatment choices for each tumour type were discussed with a clinical expert in each of the following tumour types, which cover the majority of those represented in the integrated analysis population:

- Non-small cell lung cancer
- Breast cancer
- Sarcoma
- Thyroid Cancer
- Neuroendocrine tumours
- Colorectal cancer
- Pancreatic cancer

Comparators were validated in each case and where necessary were amended according to recommendations from clinical experts, though were kept broadly in line with the therapies listed in NICE Pathways. Average PFS and OS outcomes identified in relevant NICE technology appraisals were also discussed with a view to establishing whether they reflected what is seen in clinical practice; clinical experts endorsed the data, with the caveat that that they were trial data, and some comparator OS outcomes were confounded by crossover, and therefore in some cases comparators exhibited better-than-expected outcomes where adjusted data could not be found. These are noted in Appendix L. Some comparator outcomes therefore represent a conservative estimate with regard to the relative efficacy of entrectinib.

Validation of tumour type proportions represented in the integrated analysis

The proportions of tumour types represented in the integrated analysis population shown in section B.2.3.2 represent a cross-section of patients recruited globally, including in the UK. This was the result of an extensive genetic screening programme of approximately [redacted] patients performed by the central laboratory used in the entrectinib clinical trial programme. Given the volume of screened patients, it is reasonable to expect that the resulting tumour proportions in the integrated analysis population may reflect that seen in clinical practice.

Feedback from two UK trial investigators suggests that the frequencies of the tumour types seen may reflect clinical practice, with the possible exception that MASC is over represented, potentially due to the fact that 91-100% of MASC tumours exhibit an *NTRK* gene fusion (34, 35). Therefore, an *NTRK* gene fusion population may be expected to be enriched for this tumour type.

However, there is still uncertainty as to whether the integrated analysis population represents that which may be seen in clinical practice in terms of frequency of tumour types. Therefore, this has been explored in scenario analyses whereby maximum (100%) weighting has been given to the most and least cost-effective tumour types (MASC and pancreatic, respectively) in order to demonstrate a range of ICERs. This scenario analysis was conducted by using the integrated analysis population data as a proxy of entrectinib's performance in any given tumour type, which was then compared with a 100% weighting of comparators relevant to the two tumour types. Due to the very small number of patients in the integrated efficacy analysis population with any given tumour type, direct subgroup comparisons were not deemed to be informative.

B.3.4 Measurement and valuation of health effects

Health-related quality of life data were obtained from the STARTRK-2 trial for entrectinib, and separately *via* a review of published literature for the comparator.

B.3.4.1 Health-related quality-of-life data from clinical trials

Health-related quality of life data were collected within the STARTRK-2 study using the condition-specific European Organisation for the Research and Treatment of Cancer (EORTC) Quality of Life-Core 30 (QLQ-C30) questionnaire, and generic preference-based measure EuroQoL EQ-5D-3L. HRQoL utilities incorporated in the cost-effectiveness model were derived from this trial. Evaluation of HRQoL using the EQ-5D-3L directly from patients is consistent with the NICE reference case, the methods of which and results for this instrument are described below. EORTC QLQ-C30 results are provided in the overview of clinical effectiveness (Section B.2.6.4).

Patients completed the EQ-5D-3L questionnaires at baseline (Cycle 1 Day 1); and then on Day 1 of each subsequent treatment cycle thereafter; and at the end of treatment visit. Questionnaires were also completed in the period after end of treatment. The UK tariff was used to estimate utilities (135). For the purposes of analysis, the measurements were categorised as follows:

- Baseline – pre treatment
- PFS – measurements after treatment start but prior to disease progression
- PPS – measurements after IRC assessed progression

Results for the integrated population are shown in Table 50. Adverse event utilities were not applied in the base case, as these are represented within the trial-derived estimates.

Table 50: Utility estimates from entrectinib integrated efficacy analysis

State	Number of Observations	Mean	Minimum	Maximum	Median
Baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PPS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

B.3.4.2 Mapping

As EQ-5D-3L was collected within the STARTRK-2 clinical trial, no mapping methods were utilised for the estimation of HSUVs from trial-derived HRQoL data.

However, given the small sample size, some adjustments were made to the PFS utility value derived from STARTRK-2. A linear mixed model was fit to the data; given limited observations, the model was fit only to PFS observations adjusting for Sex, Tumour type, Age and Time. The best fitting model included intercept and slope as random effects but no fixed effects were retained.

The final model results in population mean estimate with 95% CI for PFS utility of 0.8119 (0.76, 0.86) with a 95% CI estimated using bootstrapping. The use of a nested random effects model changed the utility only slightly with PFS utility of [redacted]. The nested random effects estimate has been used as base-case in the model. This makes the assumption that tumours were randomly sampled from a population of possible tumours and that patients were then sampled randomly from within this tumour pool.

As is evident from Table 50, the PPS utility value for entrectinib derived from STARTRK-2 is based on a very limited number of observations and is somewhat implausible as it is higher than the PFS value. For this reason, in the base case, post-progression utility for entrectinib was assumed to be equal to established management.

In the absence of individual-level data, mapping of estimates from the comparator population was not deemed feasible. As with survival outcomes, where possible PFS and PPS utilities for the comparator arm were derived from weighted averages of values used in NICE

technology appraisals for relevant tumour types. Further details on how these were calculated are shown in B.3.4.3.

B.3.4.3 Health-related quality-of-life studies

Systematic search

Although a search was conducted to identify utility values specific to *NTRK*-fusion positive solid tumours, no publications with relevant data were identified (see appendix H).

In the absence of utility data specific to the *NTRK*-fusion positive cohort, utility estimates were identified using the same search approach as utilised for the identification of clinical outcomes estimates (B.3.3.1). Further details are provided in Appendix H.

Selection of base case utility estimates

In order to identify preferred utility sources for each tumour type, the methods used for each of the utility sources in Appendix H were evaluated for consistency with a number of criteria related to the decision problem and NICE reference case:

- Data collected from patients with relevant tumour type
- Metastatic/advanced stage disease
- Alignment with model structure (PFS/post-progression)
- Level of consistency with NICE reference case

The utility sources which most closely aligned with the criteria described above was selected as the relevant estimates for each tumour type (Table 51).

A weighted average for the indicative comparator cohort was then calculated according to the proportion of patients with each tumour type within the trial

Table 51: Selected utility sources for comparator tumour types

Tumour type	N	Utility estimate – PFS	Measure of uncertainty (SE)	Utility estimate – PPS	Measure of uncertainty (SE)	Source
Colorectal cancer	4	0.73	0.14	0.64	0.14	TA405
MASC	7	0.725	0.14	0.60	0.14	Assumption: average of known
Thyroid cancer (papillary and anaplastic)	5	0.72	0.14	0.64	0.14	TA535
Non-small-cell lung cancer	10	0.74	0.18	0.59	0.06	TA428

(squamous and non-squamous)						
Pancreatic cancer	3	0.70	0.14	0.65	0.14	TA476
Sarcoma	13	0.72	0.14	0.56	0.14	TA465
Neuroendocrine tumours	3	0.767	0.14	0.725	0.14	TA539
Breast cancer (including secretory)	6	0.705	0.14	0.496	0.14	TA515
Other (average of known)	3	0.725	0.14	0.65	0.14	Assumption: average of known
Weighted average		0.73		0.59		Calculation

B.3.4.4 Adverse event disutilities

There are two approaches that could be taken regarding the inclusion of AE impacts on HRQoL:

1. The assumption that any disutility has already been incorporated in to the base case health state utilities through trial derived EQ-5D utilities, and incorporating an additional disutility could be considered double counting;

2. The assumption that averaged trial-derived utilities underestimate disutilities associated with adverse events, and therefore an additional disutility must be applied.

The base case analysis takes the former assumption (disutility has already been incorporated). Scenario analyses conducted to assess the impact of disutilities associated with the adverse events listed in Table 52 showed that these had minimal impact on the ICER, and are therefore not included.

Based on clinical advice, the only adverse event which is expected to be higher for entrectinib than chemotherapy is weight gain ("weight increased"). Disutility was applied to the level described below (Table 52).

Table 52: Adverse event disutilities

Adverse event	Disutility	Confidence intervals	Frequency	Reference
Entrectinib-related adverse events				

Weight increased	-0.0472	(-0.0375, -0.0569)	7%	Lane et al 2014 (136)
Chemotherapy-related adverse events				
Diarrhoea	-0.006	(-0.026,0.014)	10%	TA515 (137)
Febrile neutropenia	-0.012	(-0.041,0.017)	10%	TA515 (137)
Palmar-plantar erythro-dysaesthesia syndrome	0.000	(-0.013,0.012)	10%	TA515 (137)
Peripheral sensory neuropathy	-0.014	(-0.030,0.002)	10%	TA515 (137)

B.3.4.5 Health-related quality-of-life data used in the cost-effectiveness analysis

An overview of the base-case health state utility values are described in Table 53. Utility values are assigned according to the relevant health states, with progression-free estimates separated from progressed disease.

Utility values for entrectinib in the progression-free state were derived directly from the integrated efficacy analysis, whilst utility estimates for the relevant comparators were obtained from the preferred sources previously described in B.3.3.1. It is recognised that the PFS utility used in the base case for entrectinib ([redacted]) is higher than that derived for the comparator (0.73). A plausible explanation for this is that entrectinib is an oral TKI therapy with a more convenient administration and relatively tolerable safety profile when compared with traditional cytotoxic chemotherapies, which form the majority of comparator products. Although for the purposes of the model a conservative assumption has been made that the

side effect profiles of entrectinib and the comparator are the same (with the exception of weight increase on entrectinib), these chemotherapies are commonly associated with adverse events such as nausea, vomiting, diarrhoea, alopecia and neutropenic sepsis. However, in recognition of this difference in PFS utility values, a scenario analysis has been conducted whereby the PFS utility of entrectinib has been reduced to the level of the comparator.

Utility estimates for tumour types in progressed disease were also obtained from the sources most closely meeting the criteria described above in B.3.4. For the base case analysis, post-progression utility for entrectinib was assumed to be equal to established management. Due to the small sample size and associated uncertainty, the post-progression utility from the integrated efficacy analysis was not used.

Table 53: Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (standard error)	95% confidence interval	Justification
Progression-free survival			
Entrectinib	[redacted]	[redacted]	Utility derived from clinical trial and valued according to UK societal preferences
Established management weighted average	0.73	Applied at individual tumour level	Weighted average of tumour-specific utilities
Progressed disease			
Entrectinib	0.59	Applied at individual tumour level	Assumption of equivalent PPS utility to comparator
Established management	0.59	Applied at individual tumour level	Weighted average of tumour-specific utilities

HS, health state; AR, adverse reaction

B.3.5 Cost and healthcare resource use identification, measurement and valuation

An SLR was conducted to identify costs and healthcare resource use evidence for *NTRK* fusion-positive patients (see Appendix D). However, no relevant sources were identified.

B.3.5.1 Intervention and comparators' costs and resource use

Drug acquisition costs – intervention and comparator

Drug acquisition costs used in the model for entrectinib and an indicative estimate for the comparator are presented in Table 54. The [redacted] as used in the base case analysis, is shown as [redacted]. Details of entrectinib dosing are as follows:

- Entrectinib: For adults, the recommended dose of entrectinib is 600 mg administered orally once daily. The proposed list price of entrectinib is £5160.00 per month.

Table 54: Drug acquisition costs for entrectinib

Drug	Pack concentration	Pack volume	Dose per pack	Cost per pack	Source
Entrectinib	100 mg	30	3,000 mg	[redacted] (£860.00)	[redacted] (and list) price
Entrectinib	200 mg	90	18,000 mg	[redacted] (£5,160.00)	[redacted] (and list) price

Although a number of the chemotherapy comparators that inform the comparator price are subject to patient access schemes, the extent of these is unknown. Drug acquisition costs for the comparator are therefore shown at list price, as described in the British National Formulary (Table 55).

Table 55: Individual comparator acquisition costs (138)

Drug	Formulation	Composition (ml or tablets)	Cost (£)/pack	Cost (£)/mg	Cycle length	Dose per cycle	Source
Capecitabine	Tablet	150 mg/tablet	30.00	0.0033	2 weeks	1250 mg/m^2	BNF
Eribulin	Vial	0.88 mg/2ml	361.00	410.23	3 weeks	2.26 mg/m^2	BNF
Vinerolbine	Vial	10 mg/1ml	29.00	2.90	Weekly	25-30 mg/m^2	BNF
Gemcitabine	Vial	1 g/10ml	13.09	0.01	3 weeks	2500 mg/m^2	BNF
Paclitaxel	Vial	100 mg/16.7ml	200.35	2.00	3 weeks	175 mg/m^2	BNF
Docetaxel	Vial	20 mg/ml	91.51	4.58	3 weeks	75 mg/m^2	BNF
Irinotecan	Vial	40 mg/2ml	39.38	0.98	2 weeks	180 mg/m^2	BNF

Folinic acid	Vial	50 mg/5ml	20.00	2.00	2 weeks	500 mg/m^2	BNF
Fluorouracil	Vial	500 mg/10ml	6.08	0.01	2 weeks	12 mg/kg	BNF
5FU	Vial	2.5 g/50ml	32.00	0.01	2 weeks	600 mg/m^2	BNF
Oxaliplatin	Vial	50 mg/10ml	155.00	3.10	2 weeks	85 mg/m^2	BNF
Trifluridine-tipiracil	Tablet	15 mg	25.00	1.67	4 weeks	700 mg/m^2	BNF
Everolimus	Tablet	10 mg	2673.00	8.91	Daily	10 mg	BNF
Nab-paclitaxel	Vial	100 mg	246.00	2.46	4 weeks	375 mg/m^2	BNF
Leucovorin	Vial	100 mg/10ml	37.50	0.38	2 weeks	200 mg/m^2	BNF
Doxorubicin	Vial	200 mg/100ml	391.40	1.96	3 weeks	60 – 75 mg/m^2	BNF
Trabectedin	Vial	0.25 mg	363.00	1,452.00	3 weeks	1.5 mg/m^2	BNF
Nintedanib	Capsule	100 mg/capsule	2,151.10	17.93	3 weeks	8000 mg	BNF

BNF, British National Formulary

Drug acquisition costs – comparator by tumour type

An aggregated monthly drug acquisition cost for each tumour type is shown in Table 56. These are an average of the monthly acquisition costs for each identified comparator for the given tumour type.

Table 56: Tumour-specific monthly drug acquisition – average by tumour type

Tumour type	Cost per month
Colorectal cancer	£1,878.09
MASC	£0
Thyroid cancer (papillary and anaplastic)	£0
Non-small-cell lung cancer (squamous and non-squamous)	£1952.05
Pancreatic cancer	£1,507.37
Sarcoma	£3,096.16
Neuroendocrine tumours	£1,354.32
Breast cancer (including secretory)	£1,178.76
Other (average of known)	£1,281.60

MASC, mammary analogue secretory carcinoma

Drug acquisition costs – subsequent therapies

Drug acquisition costs for subsequent therapies are assumed to be the same as for the comparator, and therefore comprise a weighted average of these costs. The rationale for this is that subsequent therapy costs are only applied to entrectinib patients, as they may be eligible to receive treatment after progression on entrectinib. Meanwhile, patients receiving

comparator therapies are generally considered to be at the end of the treatment pathway. These assumptions are tested within scenario analyses.

Drug administration costs – intervention and comparator

Although a number of the chemotherapy combinations utilised in standard treatment involve significant administration time, some over multiple days, a simplifying assumption was made that each treatment should be sorted into three categories. Administration cost per month was therefore calculated based on the proportion of chemotherapy types fitting each of three categories of administration (oral; simple IV; complex IV). Average monthly administration costs were then calculated according to tumour type (Table 57).

Table 57: Administration costs – intervention and comparator

Drug	Type of administration	NHS reference code	Administration unit cost (2017/18)	Administration cost per month	Average monthly admin cost
Intervention					
Entrectinib	Oral	Pharmacist preparation (oral) 12 minutes/ month: £14.59	£14.59	£14.59	£14.59
Comparators					
NSCLC					
Docetaxel	Simple IV chemotherapy	SB12Z	£229.00	£331.69	£338.98
Docetaxel + nintedanib	Simple IV chemotherapy + oral	SB12Z + pharm prep	£243.59	£346.28	
Colorectal cancer					
FOLFIRI	Complex IV chemotherapy	SB14Z	£337.00	£488.12	£278.13
Oxaliplatin	Simple IV chemotherapy	SB12Z	£229.00	£331.69	
Irinotecan	Simple IV chemotherapy	SB12Z	£229.00	£331.69	
Trifluridine-tipiracil	Oral	Pharm prep	£14.59	£14.59	
Breast cancer					
Capecitabine	Oral	Pharm prep	£14.59	£14.59	£330.63
Eribulin	Simple IV chemotherapy	SB12Z	£229.0	£331.69	
Vinorelbine	Complex IV chemotherapy	SB14Z	£337.00	£488.12	
Gemcitabine + paclitaxel	Complex IV chemotherapy	SB14Z	£337.00	£488.12	

MASC					
BSC	N/A	N/A	N/A	N/A	0
Soft tissue sarcoma					
Doxorubicin	Simple IV chemotherapy	SB12Z	£229.0	£331.69	£331.69
Trabectedin	Simple IV chemotherapy	SB12Z	£229.0	£331.69	
Pancreatic cancer					
FOLFIRINOX	Complex IV chemotherapy	SB14Z	£337.00	£488.12	£435.97
Gemcitabine + nab-paclitaxel	Complex IV chemotherapy	SB14Z	£337.00	£488.12	
Gemcitabine	Simple IV chemotherapy	SB12Z	£229.0	£331.69	
Thyroid cancer					
BSC	N/A	N/A	N/A	N/A	0
Neuroendocrine cancer					
Everolimus	Oral	Pharm prep	£14.59	£14.59	£7.30
BSC	N/A	N/A	N/A	N/A	

BSC, best supportive care; IV, intravenous; MASC, mammary analogue secretory carcinoma; N/A, not available; NSCLC, non-small cell lung cancer

Subsequent therapies

For the purpose of this appraisal, it is assumed that a proportion of entrectinib patients will still be eligible for established management following disease progression. At the date of primary analysis of the integrated population, [redacted] of patients received a subsequent anti-cancer therapy after progression on entrectinib (105). Therefore, the base case assumes that [redacted] of entrectinib patients who experience disease progression will incur equivalent monthly drug acquisition and administration costs to those attributed to the comparator in PFS state. In the case of the comparator, given the advanced stage of disease, no drug acquisition or administration costs are assumed following progression (Table 58). These assumptions are tested within scenario analyses.

Table 58: Subsequent therapy following progression

	% receiving active therapy	Monthly cost
Entrectinib	[redacted]	£1,581.00
Comparator	0	£0.00

Due to the variation between chemotherapy types informing the average comparator price, wastage has not been applied to intervention or comparator.

B.3.5.2 Health-state unit costs and resource use

In order to estimate health state costs across a range of chemotherapy types, it was necessary to apply a simplifying assumption that treatments with similar routes of administration are likely to be associated with similar routine healthcare costs across the different tumour types.

The following classifications were used to categorise anti-cancer therapies according to their route of administration:

- Oral: entrectinib, any oral chemotherapy (e.g. capecitabine)
- Simple IV: single-agent chemotherapy (e.g. gemcitabine)
- Complex IV: combination chemotherapy involving at least one IV formulation (e.g. gemcitabine + capecitabine; FOLFIRI)

Healthcare resource use estimates for the relevant method of administration were obtained from the most recent NICE technology appraisal (TA) identified within the search in B.3.3.1; these were TA515 (Eribulin for treating locally advanced or metastatic breast cancer after 1 chemotherapy regimen (137)) for oral chemotherapy, TA520 (Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy (139)) for simple IV chemotherapy and TA476 (Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer (140)) for complex IV chemotherapy. As well as being recent examples of accepted costs for each administration type, these TAs also represent relevant tumour types and lines of therapy to this appraisal. Due to the complexity of validating costs for multiple different tumour types, the clinical expert validation of costs in TA476, TA515 and TA520 was accepted as being generalisable to the costs associated with tumour types covered in this appraisal. The costs broadly reflect the reference TAs with the exceptions that: costs were updated to the most recent NHS reference and Personal Social Services Research Unit costs (141, 142), costs of certain tumour-specific tests are not included (e.g., tumour marker CA19-9 test in TA476), and adjustments have been made for logical consistency across the TAs. The summary of cost components is shown in Table 59, Table 60, and Table 61.

Table 59: PFS health state: oral treatment HCRU (from TA515 (137))

Item	Number used	% of patients	Unit cost	Monthly cost	Reference*
Medical oncology, outpatient visit	1	100	£162.05	£162.05	Total Outpatient Attendances - row 370 (Outpatient,

					consultant-led)
GP surgery visit	1	100	£37.40	£37.40	PSSRU 2018 (143): 10.3b GP unit costs (9.22 minutes patient time)
CT scan	1	33	£132.75	£44.21	RD22Z (CT scan of one area, pre- and post-contrast)
Full blood count	1.55	100	£2.51	£3.89	DAPS05 (haematology)
Liver function tests	1.66	100	£1.11	£1.84	DAPS04 (clinical biochemistry)

*NHS Reference costs 2017-2018 unless otherwise stated (141); CT: computerised tomography

Table 60: PFS health state: simple IV chemotherapy HCRU (from TA520 (139))

Item	Number used	% of patients	Unit cost	Monthly cost	Reference*
Medical oncology, outpatient visit	1.33	100	£162.05	£215.53	Total Outpatient Attendances - row 370 (Outpatient, consultant-led)
GP surgery visit	1	100	£37.40	£37.40	PSSRU 2018 (143): 10.3b GP unit costs (9.22 minutes patient time)
CT scan	1	33	£132.75	£44.21	RD22Z (CT scan of one area, pre- and post-contrast)
Full blood count	1.55	100	£2.51	£3.89	DAPS05 (haematology)
Liver function tests	1.66	100	£1.11	£1.84	DAPS04 (clinical biochemistry)

*NHS Reference costs 2017-2018 unless otherwise stated (141); CT: computerised tomography

Table 61: PFS health state: complex IV chemotherapy HCRU (from TA476 (140))

Item	Number used	% of patients	Unit cost	Monthly cost	Reference*
Medical oncology, outpatient visit	1.33	100	£162.05	£215.53	Total Outpatient Attendances - row 370

					(Outpatient, consultant-led)
Medical oncology, outpatient, nurse-led	1	50	£104.00	£52.00	Total Outpatient Attendances - row 370 (Outpatient, non-consultant-led)
GP surgery visit	1	100	£37.40	£37.40	PSSRU 2018 (143): 10.3b GP unit costs (9.22 minutes patient time)
Nurse community visit	1	50	£42.00	£21.00	PSSRU 2018 (143)- 10.2 GP practice nurse unit costs (assumes 1 hour patient contact)
CT scan	1	33	£132.75	£44.21	RD22Z (CT scan of one area, pre- and post-contrast)
Full blood count	1.55	100	£2.51	£3.89	DAPS05 (haematology)
Liver function tests	1.66	100	£1.11	£1.84	DAPS04 (clinical biochemistry)

*NHS Reference costs 2017-2018 unless otherwise stated (141); CT: computerised tomography

Monthly HCRU costs for entrectinib were assumed to be entirely associated with those of an oral therapy. For the indicative chemotherapy comparator, the proportion of each chemotherapy category was applied to create a weighted comparator cost (Table 62). Scenario analysis was applied to test the influence of changes in this estimate, as described in Section B.3.8.

Table 62: Monthly HCRU costs by category of chemotherapy

Category	Tests	Monitoring	Healthcare staff costs	Proportion of entrectinib HCRU costs	Proportion of comparator HCRU costs
Oral	£5.73	£44.21	£199.45	100%	21%
Simple IV	£5.73	£44.21	£252.93	0%	43%
Complex IV	£5.73	£44.21	£325.93	0%	36%

Due to the positioning of entrectinib towards the end of the treatment pathway, it was assumed that patients assigned comparator treatment would receive palliative care following disease progression with no active therapy, and hence no monitoring or testing costs were applied following disease progression. TA515 provides the most recent estimate for these palliative care costs and was used in the base case (Table 63).

Table 63: Progressed disease health state costs (from TA515 (137))

Item	Number used	% of patients	Unit cost	Monthly cost	Reference*
Medical oncology, outpatient visit	1	100	£162.05	£162.05	Total Outpatient Attendances - row 370 (Outpatient, consultant-led)
Medical oncology, outpatient, nurse-led	1	100	£104.00	£104.00	Total Outpatient Attendances - row 370 (Outpatient, non consultant-led)
GP home visit	1	100	£37.40	£37.40	PSSRU 2018 (143): 10.3b GP unit costs (9.22 minutes patient time)
Nurse community visit	1	67	£42.00	£30.15	PSSRU 2018 (143): 10.2 GP practice nurse unit costs (assumes 1 hour patient contact)

*NHS Reference costs 2017-2018 unless otherwise stated (141)

To reflect the increasing costs of unplanned healthcare requirements towards the end of life, end-of-life care costs were applied at the transition from PD to death. Costs from Georghiou and Bardsley (144), adjusted for inflation to 2017-2018 (143), are shown in Table 64. To avoid double-counting, the equivalent of three months of PD health state costs was subtracted from the end-of-life care cost.

Table 64: Terminal care costs

Component	Mean cost, last 3 months (2017 – 2018)	Mean cost/month, last 3 months (2017 – 2018)
Emergency inpatient admission	£4049.29	£1349.76
Non-emergency inpatient admission	£1352.75	£450.92
Outpatient attendance	£375.98	£125.33
A&E visits	£79.57	£26.52
Social care	£441.63	£147.21
District nursing care	£584.86	£194.95
GP visits	£363.05	£121.02

B.3.5.3 Adverse reaction unit costs and resource use

Unit costs and resource use for the management of adverse events occurring at a rate of ≥5% in the entrectinib integrated analysis were considered in the base case. Clinical input suggests that increased weight is the only adverse event which is expected to be higher for entrectinib than chemotherapy, therefore all adverse events except increased weight were considered to occur at the same rate for the comparator.

In practice, chemotherapy-related adverse events are anticipated to have a significant cost and resource implication compared to a targeted therapy such as entrectinib; as such the base-case represents a conservative assumption in terms of the cost-effectiveness of entrectinib.

As adverse events will typically emerge towards the start of treatment with an anti-cancer therapy, costs of managing each adverse event were applied at the start of the first cycle for both entrectinib and the comparator, rather than applying a monthly probability throughout.

HRG codes were sourced for the relevant management activities and are described in Table 65.

Table 65: HRG codes for adverse events

Toxicities grade 3/4	Percentage reported in integrated analysis	Costs 2017 - 2018	HRG code/reference	Description
Base case analysis				
Anaemia	13%	£505.00	SA04K	Iron deficiency anaemia with cc score 2-5, non-elective short stay

Fatigue	6%	£42.00	Not applicable	Assumption: community nurse visit (1 hour patient contact)
Neutropenia	7%	£273.00	SA12J	Assumption: same as thrombocytopenia with cc score 2-4, day case
Weight increased	7%	£0.00	Not applicable	Assumption: no additional management required

B.3.5.4 Miscellaneous unit costs and resource use

A projected cost of screening for eligible entrectinib patients is included in the base case analysis. Although Roche's preference is to work with the NHS to support the implementation of genomic screening, guidance regarding the process of introducing and reimbursing a new target within the NHS Genomic Testing Directory has not yet been established. For the purposes of including screening costs within the model, a hierarchical approach is therefore proposed where immunohistochemistry testing (IHC) is conducted to identify patients with tumours expressing NTRK protein, followed by confirmatory screening with a next-generation sequencing panel to establish whether these patients have specific NTRK gene fusions.

The proposed approach for NTRK screening assumes that an IHC assay (Ventana pan-TRK [EPR17341] assay) is reimbursed at the standard tariff within the NHS, and that NGS costs associated with acquiring a commercially-available test are only attributed to those patients with a potential fusion.

Two steps were involved in the calculation of expected screening costs:

- Estimation of number-needed-to-screen
- Attribution of screening costs

Estimation of number-needed-to-screen

The frequency of NTRK fusions has been evaluated by Roche and informed the fusion rates used in the model (Table 66). The number needed-to-screen (NNS) to identify an eligible entrectinib patient was then estimated using the following equation:

$$NNS = \frac{1}{NTRK\ fusion\ rate\ (\%)}$$

Table 66: Frequency of NTRK fusions in enrolled tumour types (30, 145)

Tumour type	NTRK Fusion Rate	Number needed to screen
CRC	[redacted]	[redacted]
NSCLC (squamous and non-squamous)	[redacted]	[redacted]
Pancreatic	[redacted]	[redacted]
Non-secretory breast cancer	[redacted]	[redacted]
Secretory Breast Carcinoma (0.02% HER2-)	[redacted]	[redacted]
Thyroid (papillary/anaplastic)	[redacted]	[redacted]
Neuroendocrine tumours	[redacted]	[redacted]
Sarcoma (non-paediatric)	[redacted]	[redacted]
MASC	[redacted]	[redacted]
Other	[redacted]	[redacted]

Attribution of screening costs

Attribution of costs of screening is dependent on the tumour type considered within the analysis and whether screening is conducted in current practice (Table 67). For tumours where NTRK fusions are already included in the Genomic Testing Directory (MASC) and whole genome sequencing is reimbursed for specific tumour types (paediatric tumours and sarcoma), costs of whole-genome sequencing have been included. Although predicted to reduce over time, these have previously been reported to be £800 per genome in standard UK practice (146). In these tumour types it is assumed NTRK-fusion positive patients will be identified through the established pathway and this cost is included for both entrectinib and comparator treatment.

In tumour types where one or more genetic test is conducted in standard clinical practice (colorectal cancer, thyroid cancer, non-small-cell lung cancer and breast cancer [secretory and non-secretory]), it was assumed that the cost of standard testing is £75.00 (147). This assumes that where more than one test is required, a panel test will be utilised and reimbursed to the same value.

In other tumour types where testing is not covered within the Genomic Testing Directory (pancreatic cancer, neuroendocrine tumours, other cancers), no costs are attributed to standard practice.

Table 67: Summary of tumour types covered within NHS Genomic Testing Directory

Tumour type	NTRK-fusion testing	Whole genome sequencing	Other biomarker screening	No molecular testing within directory
CRC			x	
NSCLC (squamous and non-squamous)			x	
Pancreatic				x
Non-secretory breast cancer			x	
Secretory Breast Carcinoma (0.02% HER2-)			x	
Thyroid (papillary/anaplastic)			x	
Neuroendocrine tumours				x
Sarcoma (non-paediatric)		x		
MASC	x			
Other				x
Paediatric cancers		x		

Costs of screening as shown in Table 68 are then applied for the number of patients screened for each of the tumour types. The frequency rate of NTRK mutations is used as an indicator of the number of patients identified as potentially NTRK-fusion positive, and this is further narrowed to identify true NTRK-fusion positive patients using NGS. Clinical data provided by an investigator involved in the entrectinib clinical development programme suggests that the IHC testing approach will remove 89% of NTRK-fusion negative samples, reducing the requirement for NGS confirmation to approximately 1 in 10 patients. No data are currently available on the sensitivity and specificity for the Ventana IHC assay, therefore these are assumed to be 100%. Sensitivity and specificity rates reported for a representative NGS assay (Oncomine Focus Assay) are 100% for gene fusions (148).

Table 68: Testing cost approaches

Diagnostic test	Unit cost	Reference
ETV6-NTRK3 testing via NHS Testing Directory*	£75.00	Assumption: pathologist input obtained during TA406
Whole genome sequencing via NHS Testing Directory	£800.00	SSAR 2019 (146)
IHC testing cost	£75.00	Assumption: pathologist input obtained during TA406
Next Generation Sequencing panel	[redacted]	[redacted]

In the base case scenario, 100% of incremental screening costs are applied to entrectinib due to the uncertainty of other tumour-agnostic medicines reaching the market at the anticipated time of approval of entrectinib. However, due to the simultaneous NICE appraisal of another NTRK fusion-targeting medicine, larotrectinib, a scenario analysis is performed whereby 50% of incremental screening costs are applied to entrectinib, to avoid double counting in the event that two NTRK fusion-targeting medicines are available. Scenarios in which the incremental screening cost is split four ways (to account for availability of further tumour-agnostic medicines) and in which it is excluded altogether are also explored (see Section B.3.8).

The costs of screening each tumour type to identify one entrectinib-eligible patient are shown in Table 69.

Table 69: Costs of screening by tumour type to identify one patient (base case analysis)

Tumour type	Base case: entrectinib	Base case: comparator
CRC	[redacted]	[redacted]
NSCLC (squamous and non-squamous)	[redacted]	[redacted]
Pancreatic	[redacted]	[redacted]
Non-secretory breast cancer	[redacted]	[redacted]
Secretory Breast cancer	[redacted]	[redacted]
Thyroid (papillary/anaplastic)	[redacted]	[redacted]
Neuroendocrine tumours	[redacted]	[redacted]
Sarcoma (non-paediatric)	[redacted]	[redacted]
MASC	[redacted]	[redacted]

Other	[redacted]	[redacted]
Weighted average within integrated analysis	[redacted]	[redacted]

As screening will be required to identify eligibility for treatment with entrectinib, all screening costs within the economic model are attributed as a weighted cost for the proportion of tumour types in the model in Cycle 1.

B.3.6 Summary of base-case analysis inputs and assumptions

B.3.6.1 Summary of base-case analysis inputs

Table 70: Summary of variables applied in the economic model

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
General model parameters			
Time horizon	30 years	Fixed	B.3.2
Discount rate - efficacy	3.5%	Fixed	
Discount rate - costs	3.5%	Fixed	
Cost year	2017 - 2018	Fixed	
Population parameters			
Age	57 years	Fixed	B.2.3
Body weight	73.6 kg	Fixed	
Height	168.9 cm	Fixed	
Body surface area	1.84 m^2	Fixed	
Clinical inputs			
Assessment of progression	RECIST v1.1	Fixed	B.3.3
Survival extrapolation			
PFS – entrectinib	Exponential	Multivariate normal	B.3.3
PFS – comparator	Exponential	Multivariate normal	
OS – entrectinib	Exponential	Multivariate normal	
OS – comparator	Exponential	Multivariate normal	
Utilities – base case			
Progression-free – entrectinib	[redacted]	[redacted]	B.3.4.1
Progressed disease – entrectinib	0.59	0.141	
PFS – comparator	0.73	0.147	
Progressed disease – comparator	0.59	0.141	
Technology acquisition costs per pack (unit costs at list price)			
Entrectinib	£5,160.00	Fixed	B.3.5.1

[redacted]	[redacted]	[redacted]	
Weighted comparator (with discounting)	£1,581.00	Fixed	
Administration costs: Intervention and Comparator – per month			
Entrectinib	£14.59	Normal	B.3.5.1
Weighted comparator	£233.00	Normal	
Administration costs: Subsequent therapies – per month			
Entrectinib	£233.00	Fixed	B.3.5.1
Comparator	£0.00	Fixed	
Supportive care costs – per month			
PFS (entrectinib)	£249.39	Normal	B.3.5.2
PFS (comparator)	£317.48	Normal	
PPS	£331.59	Normal	
Terminal care cost (last 3 months)			
Terminal care cost	£6,252.37	Normal	B.3.5.2
Adverse event management costs			
Anaemia	£505.00	Normal	B.3.5.3
Fatigue	£42.00	Normal	
Neutropenia	£273.00	Normal	
Weight increased	£0.00	Normal	
Subsequent treatment			
Entrectinib: patients receiving comparator therapy post-progression	[redacted]	Fixed	B.3.5.1
Comparator: patients receiving post-progression therapy	0%	Fixed	
Cost of NTRK test			
Average costs of screening with potential to identify 1 NTRK+ patient (entrectinib)	£55,556.55	Normal distribution	B.3.5.4
Average costs of screening with potential to identify 1 NTRK+ patient (comparator)	£39,718.06	Normal distribution	B.3.5.4

B.3.6.2 Assumptions

A summary of assumptions within the economic model is provided in Table 71.

Table 71: Assumptions within economic model

Area	Assumption	Justification
Time horizon	30 years	Sufficient to capture all changes in patient outcomes and costs for an advanced-stage patient

Comparator	Weighted chemotherapy comparator	Represents likely outcomes for standard practice according to NICE guidance. Represents conservative estimate with respect to incremental effectiveness of entrectinib.
Population	Population within integrated analysis is generalisable to a population within England and Wales	This is a key uncertainty in the model and is one of the points addressed in the proposed CDF data collection plan. Tumour types identified in England and Wales will be dictated to some extent by existing/planned screening programmes. This is also explored via scenario analyses by 100% weighting being applied to the most and least cost-effective tumour types.
Clinical effectiveness: PFS & OS (entrectinib)	Exponential	Best statistical fit to the entrectinib data, representing a conservative but clinically plausible assumption.
Clinical effectiveness: PFS & OS (comparator)	Exponential	NICE TSD14 highlights conversion of published medians to exponential mean as reasonable in the absence of patient-level data (133).
Clinical effectiveness: prognostic factors	No adjustment made for NTRK fusion positive status	Due to the limited data available, no adjustment is made for the prognostic implications of NTRK fusion positive status. This is tested within scenario analyses and comprises part of the CDF data collection proposal.
Clinical effectiveness: prognostic factors	No adjustment for CNS metastases	Although randomised trials typically do not recruit patients with baseline CNS metastases, variable levels of reporting made it infeasible to adjust the comparator outcomes in the base case. An indicative scenario analysis was conducted to test the influence of a matched proportion of CNS mets patients on the ICER.
Treatment duration	Entrectinib treatment duration is equivalent to PFS	As per the SmPC, entrectinib is administered until disease progression or unacceptable toxicity
PFS supportive care	Assumption of three levels of drug administration costs: oral, simple IV and complex IV	A simplifying assumption was required to standardise the numerous routes and costs of delivering comparator chemotherapy. This takes a conservative approach and likely underestimates the cost implications of delivering complex chemotherapy over multiple days (e.g. FOLFIRI)
PPS supportive care	Assumption that entrectinib and comparator patients receive equivalent levels of healthcare following progression	Clinical input suggests that this assumption is reasonable.
End of life cost	Based on previous NICE TAs	Applied as a one off cost for all patients who die to take into consideration the added expense of terminal care
HRQoL	Based on EQ-5D data collected in STARTRK-2	Consistent with previous appraisals
	Weighted average of data from previous NICE appraisals	Selected data were identified and accepted within previous NICE technology appraisals

	Post-progression utility is equivalent for comparator and entrectinib	Conservative assumption; limited data for entrectinib post-progression are significantly higher than weighted comparator (0.84) and may not represent the full progressed disease health state.
	Omission of AE disutilities in the base case analysis	The disutility associated with AEs was assumed to have been captured in the EQ-5D responses in STARTRK-2. This is in-line with the approach taken in past appraisals in oncology.
Safety	"Weight increased" is considered to be the only adverse event which has a higher frequency for entrectinib than comparator	Based on clinical expert feedback. Safety analysis also does not account for chemotherapy adverse events in the base case, which can have a significant impact on quality of life and costs.
Subsequent treatment	[redacted] of entrectinib patients assumed to receive post-progression comparator therapy	Based on trial data.
Screening	Cost-effective approach to screening with IHC and NGS panel is proposed	The proposed screening approach aims to minimise the cost of screening solid tumours while utilising current screening methods. It represents a conservative approach as it does not account for the benefits incurred outside of this evaluation (e.g. from identifying eligible patients for clinical trials)
	100% of incremental screening applied to entrectinib	Conservative approach; a second NTRK fusion-targeted medicine, larotrectinib, is being appraised on parallel timelines to entrectinib. This approach therefore risks double-counting screening costs in the event that both products are available to the NHS.

B.3.7 Base-case results

B.3.7.1 Base-case incremental cost-effectiveness analysis results

Base case results including the commercial-in-confidence commercial agreement are shown in Table 72.

Entrectinib is associated with a total cost of [redacted], resulting in a projected life-year gain of [redacted] years and quality-adjusted life year gain of [redacted] QALYs. The costs of established management are estimated to be £62,931, for a projected life year gain of 1.74 years and associated QALY gain of 1.12. Incrementally this results in an increased cost of [redacted], with a benefit of [redacted] life years and [redacted] QALYs.

Due to the confidential nature of comparator chemotherapy discounts, it has not been possible to account for any price reductions within the base case. However, this is explored *via* sensitivity analysis. Base case results for entrectinib at list price are shown in Table 73.

Table 72: Base-case results (with confidential PAS, includes screening costs)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£62,931	1.74	1.12	[redacted]	[redacted]	[redacted]	£54,646
Entrectinib	[redacted]	[redacted]	[redacted]				

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

Table 73: Base-case results (list price, includes screening costs)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£62,931	1.74	1.12	[redacted]	[redacted]	[redacted]	[redacted]
Entrectinib	[redacted]	[redacted]	[redacted]				

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

B.3.8 Sensitivity analyses

B.3.8.1 Probabilistic sensitivity analysis

To assess the uncertainty surrounding the variables included in the cost-effectiveness model, a probabilistic sensitivity analysis (PSA) was undertaken using 2,000 samples. The mean values, distributions around the means, and sources used to estimate the parameters are detailed in B.3.6.

Results of the PSA compared to deterministic results are presented in Table 74. The scatterplot in Figure 23 shows the iterations and the incremental cost-effectiveness plane is shown in Figure 24. A cost-effectiveness acceptability frontier is provided in Figure 25.

The analyses below are based on the proposed commercial discount for entrectinib.

Table 74: Comparison of deterministic and probabilistic results

	Deterministic				Probabilistic			
Technologies	**Total costs (£)**	**Total LYG**	**Total QALYs**	**ICER incremental (£/QALY)**	**Total costs (£)**	**Total LYG**	**Total QALYs**	**ICER incremental (£/QALY)**
Established management	£62,931	1.74	1.12	£54,646	[redacted]	[redacted]	[redacted]	£53,473
Entrectinib	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. 95% confidence intervals are shown in square brackets.

Figure 23: Cost-effectiveness plane (scatterplot)

Figure 24: Incremental cost-effectiveness plane

Figure 25: Cost-effectiveness acceptability frontier

B.3.8.2 Deterministic sensitivity analysis

Deterministic sensitivity analysis was conducted in the form of univariate sensitivity analysis to reflect uncertainty in a number of parameters related to both entrectinib and the comparator.

Selection of parameters for inclusion in the analysis was conducted *a priori.* Generally, parameters were selected due to uncertainty in their estimation, in particular the outcomes

and costs associated with established management, as well as a number of factors relevant to entrectinib.

Unless otherwise stated, base case values were adjusted across a +/- 20% range.

The selected parameters included in the univariate sensitivity analysis are shown in Table 75, while the tornado plot is shown in Figure 26.

Table 75: Parameter values for univariate sensitivity analysis

Parameter	Base case value	Lower value	Upper value
Weighted screening costs per patient (entrectinib)	£55,556.55	£44,445.24	£66,667.86
Comparator median PFS (months)	4.35	3.62	5.43
Comparator median OS (months)	17.23	14.36	21.54
Weekly cost of entrectinib (+/- 10%)	[redacted]	[redacted]	[redacted]
Monthly cost of comparator (weighted average)	£1,581.00	£1,264.80	£1,897.20
Utility: entrectinib PFS	0.81	0.72	0.91
Utility: comparator PFS	0.73	0.58	0.87
Utility: PPS	0.59	0.47	0.71
Monthly HCRU PFS - entrectinib	£249.39	£198.71	£299.27
Monthly HCRU PFS - comparator	£317.48	£253.98	£380.98
Monthly HCRU PPS	£331.59	£265.27	£397.91
End-of-life costs (one-off cost)	£7,247.14	£5,797.71	£8,696.57

Figure 26: Univariate sensitivity analysis for entrectinib vs comparator

B.3.8.3 Scenario analyses

Scenario analyses were conducted to assess uncertainty around structural assumptions of the model. Results with the commercial scheme applied for entrectinib are shown in Table 76. The following structural and methodological assumptions were considered:

- Alternative PFS extrapolations:
  - Traditional parameterisations (entrectinib)
- Alternative OS extrapolations:
  - Traditional parameterisations (entrectinib)
- Treatment duration:
  - Trial-observed vs marketing authorisation (to progression)
- Time horizon:
  - 5, 10, 15, 20, 25 and 30 years (30 years base case)
- Screening costs:
  - Shared attribution of screening costs across two *NTRK* fusion-targeted medicines
  - Shared attribution across four two *NTRK* fusion-targeted medicines
  - Exclusion of screening costs from model
- Prognosis of comparator:
  - Inclusion of central-nervous system outcomes for comparator
  - Adjustment for *NTRK* prognostic hazard ratio
- Post-progression therapy:

  - 50% post-progression chemotherapy use after entrectinib
  - 80% post-progression chemotherapy use after entrectinib
  - 50% post-progression chemotherapy use after entrectinib and comparator
- PFS utility
  - Matching entrectinib PFS utility to comparator PFS utility
- Tumour-weighting
  - 100% weighting applied to MASC
  - 100% weighting applied to pancreatic cancer

Table 76: Scenario analyses

		Entrectinib			Established Management			ICER
	Description	**LYs**	**QALYs**	**Costs**	**LYs**	**QALYs**	**Costs**	
PFS distribution (entrectinib)	Exponential: base case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,187
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55,850
	Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,116
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55,805
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	53,937
OS distribution (entrectinib)	Exponential: base case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	85,275
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	37,234
	Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	170,585
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	44,283
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	182,360
Treatment duration	Base case: according to label	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	Trial-observed	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55,004
Time horizon (years)	Base case: 30	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	72,011
	10	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	56,787
	15	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	55,003
	20	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,706
	25	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,655
Screening costs	Base case: 100% attributed to entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	50% attributed to entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	45,688
	25% attribution to entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	41,210
	Screening costs excluded	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	36,731
Prognosis of comparator	Base case: aggregated trial reported outcomes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646

	Adjustment to reflect poorer NTRK prognosis (HR= 2.33) (63)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	35,471
	Incorporation of CNS metastases (comparator) (69, 125)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	47,969
Post-progression therapy	Base case: 0% active treatment for comparator patients; 50% for entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	0% active treatment for comparator patients; 35% for entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	59,864
	0% active treatment for comparator patients; 80% for entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	70,301
	Equivalent post-progression treatment (50% each)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	56,689
PFS utility	Base case: Entrectinib PFS utility derived from trial data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	Entrectinib PFS utility reduced to match comparator PFS value	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	62,214
Tumour weighting	Base case – trial weighting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	54,646
	100% weight applied to MASC comparator outcomes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	32,373
	100% weight applied to pancreatic cancer comparator outcomes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	120,713

B.3.8.4 Summary of sensitivity analyses results

As shown in the scatter plot, entrectinib is anticipated to result in a significant clinical benefit relative to established management. Although there is a moderate amount of variation in outcomes dependent on choice of extrapolation method, each approach suggests that entrectinib is likely to be associated with an incremental gain in quality-adjusted life years. Based on assessment of statistical fit and clinical expert validation (Section B.3.3.6), the exponential fit provides a conservative but clinically-plausible estimate until more mature confirmatory data are available.

Key drivers of the economic evaluation reflect the uncertainty associated with the current paucity of data regarding *NTRK* fusion-positive disease area. The cost and survival outcomes associated with established management have a significant influence on the ICER, to a similar extent of uncertainties associated with entrectinib itself. The method for consideration of screening also clearly has a significant impact on the decision outcome, however it is clear that as additional medicines become available to spread the cost of genomic screening, the cost-effectiveness will continue to improve. This is discussed in more detail in Section B.3.11.

This submission has taken a conservative approach in many respects and has aimed not to introduce structural assumptions unnecessarily. At least two important prognostic factors have not been incorporated within the base case: the prognostic implications of an *NTRK* fusion, and similarly the prognostic implications and costs of patients with CNS metastases (comprising 22% of the entrectinib-treated population). Conducting a naïve adjustment of the comparator data for these factors individually highlights a significant improvement in the cost-effectiveness of entrectinib (£35,471 and £47,969, respectively). Lastly, the model assumes equivalent adverse event profiles, which is likely to be a conservative assumption with regard to entrectinib.

B.3.9 Subgroup analysis

No subgroups have been included within this submission. This is primarily due to the limited evidence available number of patients recruited to the entrectinib integrated efficacy analysis, meaning any clinically-defined subgroup (e.g., by tumour type or line of therapy) will have too small a sample to draw any meaningful conclusions. However, for indicative purposes a sensitivity analysis was conducted on two tumour types, whereby 100% weighting of costs and outcomes was applied to MASC and pancreatic cancer. These were chosen as they were the most and least cost-effective tumour types when compared with the

integrated efficacy analysis population (as a proxy of entrectinib's performance in any one tumour type). This analysis resulted in ICERs of £32,373 for MASC and £120,713 for pancreatic cancer. The primary driver for this difference appeared to be vastly differing incremental screening costs for these tumour types (£138,437.50 for pancreatic cancer and £0 for MASC).

It should be noted that the same approach using neuroendocrine and thyroid tumours produced implausible negative ICERs due to their long overall survival outcomes.

B.3.10 Validation

B.3.10.1 Validation of cost-effectiveness analysis

Selection of the appropriate distributions has been driven by statistical fit to the data and clinical plausibility of the extrapolated outcomes. Due to the limited availability of evidence relevant to an *NTRK*-fusion positive cohort, clinical plausibility of the curves was conducted with UK investigators from the STARTRK-2 study based on their knowledge of entrectinib and clinical interest in solid malignancy research and experimental medicine.

The economic model was constructed specifically from the UK-NHS perspective. The structure is consistent with other oncology models in utilising a partitioned survival approach and has utilised sources derived from previous relevant NICE technology appraisals.

Internal quality control and validation of the model was conducted by an external consultancy. Cell by cell validation was conducted which included formula checking, cell references and all aspects of model functionality. A number of 'pressure tests' were conducted, often using extreme values. The results of the model using these values were then compared to expected outputs to assess functional accuracy.

B.3.11 Interpretation and conclusions of economic evidence

The appraisal of entrectinib for *NTRK*-fusion positive solid tumours represents one of the first attempts to evaluate the cost-effectiveness of a "tumour-agnostic" medicine, in the UK or worldwide. As such, no previous economic evaluations have been published.

This evaluation has made use of data from the integrated clinical analysis of entrectinib, based on three non-randomised basket studies pooling data for 54 *NTRK*-fusion positive patients. In the absence of a comparator arm, it has been necessary to create an indicative comparator cohort based on treatments recommended for relevant solid tumours in England and Wales.

The demographics of the individual tumour types within the integrated analysis are likely to be generalisable to an English population. However, due to the challenges of patient identification for this rare, genomically-defined population, the approach to screening will have an influence which patients are ultimately treated with entrectinib within the NHS.

Extrapolation of the current survival outcomes for entrectinib resulted in a projected life-year gain of [redacted], versus [redacted] years for comparator chemotherapy. Quality-adjusted life years showed a similar level of patient benefit, with a gain of [redacted] QALYs for entrectinib and [redacted] QALYs for the comparator. Although subject to uncertainty due to the limited follow-up available, these conservative estimates and a range of scenario analyses suggest that entrectinib is likely to result in favourable outcomes for *NTRK*-fusion positive solid tumours patients compared with established management.

Although the costs of established management are often subject to confidential discounts and are therefore uncertain, [redacted] In fact, when cost components are considered, the cost of post-progression therapy contributes significantly to absolute incremental costs vs entrectinib drug costs. This is potentially due to the assumption that 35% entrectinib patients will be treated with comparator chemotherapy following progression.

The base case ICER is estimated to be £54,646, a figure which deems to be plausibly cost-effective, particularly when taking into account screening costs. This base case includes consideration of screening costs for the identification of eligible patients; when these costs are excluded the ICER reduces to £36,731. While Roche recognises the need to attribute an appropriate portion of costs to screening, it is important to highlight that there are important factors related to the economic evaluation of genomic sequencing which are not possible to capture within this appraisal. For example, comprehensive genomic profiling may identify multiple different actionable targets (e.g. ALK/ROS1) even where *NTRK*-fusion negative, or result in spillover health benefits for family members by identifying hereditary risk factors (e.g. BRCA mutations) (149). It may also identify eligibility for clinical trials such as the National Lung Matrix Trial (NCT02664935) (150), which could result in changes in patient outcomes but also result in cost-savings for the NHS. We therefore believe that the base case ICER does not fully reflect the patient and economic benefits associated with genomic screening, and would urge the Committee to consider this during their appraisal of entrectinib.

The key strengths with this economic evaluation relate to the evidence sources utilised and the conservative approach taken:

- Entrectinib data are obtained from a pooled analysis of over 50 patients with rare genomic fusions, and will increase in sample size during the course of CDF data collection
- Utility data for entrectinib are aligned with the NICE reference case and valued according to UK societal preferences
- Resource use and costs data are obtained from the preferred sources of previous NICE appraisals and therefore attempt to represent the most likely cost estimates for the NHS

Given the novelty of this appraisal, there are a number of limitations to the analysis. Some are typical of oncology appraisals, such as the requirement for extrapolation of survival outcomes. Others are specific to tumour-agnostic indications: most notably, the current outcomes achieved for an *NTRK*-fusion positive cohort are highly uncertain. For this reason, conservative approaches have been taken in the base case when developing a comparator cohort; scenario analyses have provided insights to the potential impact of NTRK-fusion positive status. Alternative approaches such as the use of a landmark analysis, which in theory could utilise data for trial-based non-responders to create a comparator population, was not considered appropriate as it may provide an overly-optimistic estimate of incremental effectiveness and introduce unnecessary uncertainty (151). In addition, the sample size of non-responders is too small to provide a meaningful comparator sample.

The process of developing this submission has identified a number of challenges for health technology assessment, including generalisability of the trial population, understanding baseline clinical outcomes, and the assessment of a highly heterogeneous population, which are likely to be common issues for other multiple tumour-agnostic indications nearing marketing authorisation within the next 3 – 5 years. For this reason, a proposal has been put forward for entrectinib to enter the Cancer Drugs Fund under a commercial access agreement to collect further data and reduce key uncertainties arising during the course of the appraisal. Alongside the benefits of better understanding the cost-effectiveness of entrectinib, this proposal will also enable the NHS, Industry and other stakeholders to gain experience in a manageable population, and develop or refine frameworks for the appropriate introduction of future tumour-agnostic therapies into the healthcare system.

B.4 References

1. NHS England. Improving Outcomes Through Personalised Medicine. 2016.
2. Buchanan J, Wordsworth S, Schuh A. Issues surrounding the health economic evaluation of genomic technologies. Pharmacogenomics. 2013;14(15):1833-47.
3. Chief Medical Officer. Generation Genome: Annual Report of the Chief Medical Officer 2016.
4. Phillips KA, Deverka PA, Marshall DA, Wordsworth S, Regier DA, Christensen KD, et al. Methodological Issues in Assessing the Economic Value of Next-Generation Sequencing Tests: Many Challenges and Not Enough Solutions. Value Health. 2018;21(9):1033-42.
5. Rolfo C, Ruiz R, Giovannetti E, Gil-Bazo I, Russo A, Passiglia F, et al. Entrectinib: a potent new TRK, ROS1, and ALK inhibitor. Expert opinion on investigational drugs. 2015;24(11):1493-500.
6. Khotskaya YB, Holla VR, Farago AF, Mills Shaw KR, Meric-Bernstam F, Hong DS. Targeting TRK family proteins in cancer. Pharmacology & therapeutics. 2017;173:58-66.
7. Snider WD. Functions of the neurotrophins during nervous system development: what the knockouts are teaching us. Cell. 1994;77(5):627-38.
8. Rubin JB, Segal RA. Growth, survival and migration: the Trk to cancer. Cancer treatment and research. 2003;115:1-18.
9. Gatalica Z, Xiu J, Swensen J, Vranic S. Molecular characterization of cancers with NTRK gene fusions. Modern Pathology. 2019;32(1):147-53.
10. Ardini E, Menichincheri M, Banfi P, Bosotti R, De Ponti C, Pulci R, et al. Entrectinib, a Pan-TRK, ROS1, and ALK Inhibitor with Activity in Multiple Molecularly Defined Cancer Indications. Mol Cancer Ther. 2016;15(4):628-39.
11. F. Hoffmann-La Roche Ltd. Entrectinib Summary of Product Characteristics [DRAFT]. 2018.
12. Liu D, Offin M, Harnicar S, Li BT, Drilon A. Entrectinib: an orally available, selective tyrosine kinase inhibitor for the treatment of NTRK, ROS1, and ALK fusion-positive solid tumors. Therapeutics and clinical risk management. 2018;14:1247-52.
13. National Research Council (US). Committee on A Framework for Developing a New Taxonomy of Disease. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington (DC): National Academies Press (US); 2011.
14. Seo MK, Cairns J. Do cancer biomarkers make targeted therapies cost-effective? A systematic review in metastatic colorectal cancer. PloS one. 2018;13(9):e0204496.
15. NHS England. Personalised Medicine 2018 [Accessed October 2018]. Available from: https://www.england.nhs.uk/healthcare-science/personalisedmedicine/.
16. NHS England. 2018/2019 final draft National Genomic Test Directory FAQ. 2018.
17. NHS England. NHS Genomic Medicine Service 2018 [Accessed October 2018]. Available from: https://www.england.nhs.uk/genomics/nhs-genomic-med-service/.
18. Grant M, Maytum JP. What will follow the first hundred thousand genomes in the NHS? Personalized medicine. 2018;15(4):239-41.
19. NHS England. The NHS Long Term Plan. 2019.
20. Cancer Research UK. Policy statement: patient access to molecular diagnostics and targeted medicines in England. 2018.
21. National Institute for Health and Care Excellence. Guideline - Lung cancer: diagnosis and management, Draft for consultation 2018 [Date Accessed 8th April 2019]. Available from: https://www.nice.org.uk/guidance/ng122/documents/short-version-of-draft-guideline.
22. Lemery S, Keegan P, Pazdur R. First FDA Approval Agnostic of Cancer Site - When a Biomarker Defines the Indication. The New England journal of medicine. 2017;377(15):1409-12.
23. Schwaederle M, Zhao M, Lee JJ, Eggermont AM, Schilsky RL, Mendelsohn J, et al. Impact of Precision Medicine in Diverse Cancers: A Meta-Analysis of Phase II Clinical Trials.

Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(32):3817-25.
24. Schwaederle M, Zhao M, Lee JJ, Lazar V, Leyland-Jones B, Schilsky RL, et al. Association of Biomarker-Based Treatment Strategies With Response Rates and Progression-Free Survival in Refractory Malignant Neoplasms: A Meta-analysis. JAMA oncology. 2016;2(11):1452-9.
25. FDA. Press release: FDA approves first cancer treatment for any solid tumor with a specific genetic feature 2017 [Date Accessed 10th October 2018]. Available from: https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm560167.htm.
26. FDA. Press release: FDA approves larotrectinib for solid tumors with NTRK gene fusions 2018 [Date Accessed 8th April 2019]. Available from: https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm626720.htm.
27. Garber K. Tissue-agnostic cancer drug pipeline grows, despite doubts. Nature reviews Drug discovery. 2018;17(4):227-9.
28. American Association for Cancer Research. Harnessing the Power of Precision Medicine – Treating Cancers with Tissue-agnostic Therapies 2018 [Accessed February 2019]. Available from: https://blog.aacr.org/harnessing-the-power-of-precision-medicine-treating-cancers-with-tissue-agnostic-therapies/.
29. Business Wire. FDA Grants Priority Review to Genentech's Personalized Medicine Entrectinib 2019 [Date Accessed 29th April 2019]. Available from: https://www.businesswire.com/news/home/20190218005443/en/FDA-Grants-Priority-Review-Genentech%E2%80%99s-Personalized-Medicine.
30. Amatu A, Sartore-Bianchi A, Siena S. NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. ESMO open. 2016;1(2):e000023.
31. Coppola V, Barrick CA, Southon EA, Celeste A, Wang K, Chen B, et al. Ablation of TrkA function in the immune system causes B cell abnormalities. Development (Cambridge, England). 2004;131(20):5185-95.
32. Vaishnavi A, Le AT, Doebele RC. TRKing down an old oncogene in a new era of targeted therapy. Cancer discovery. 2015;5(1):25-34.
33. Passiglia F, Caparica R, Giovannetti E, Giallombardo M, Listi A, Diana P, et al. The potential of neurotrophic tyrosine kinase (NTRK) inhibitors for treating lung cancer. Expert opinion on investigational drugs. 2016;25(4):385-92.
34. Balanza R, Arrangoiz R, Cordera F, Munoz M, Luque-de-Leon E, Moreno E, et al. Mammary analog secretory carcinoma of the parotid gland: A case report and literature review. International journal of surgery case reports. 2015;16:187-91.
35. Bishop JA, Yonescu R, Batista DA, Westra WH, Ali SZ. Cytopathologic features of mammary analogue secretory carcinoma. Cancer cytopathology. 2013;121(5):228-33.
36. Tognon C, Knezevich SR, Huntsman D, Roskelley CD, Melnyk N, Mathers JA, et al. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. Cancer Cell. 2002;2(5):367-76.
37. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nature genetics. 2013;45(8):927-32.
38. Rubin BP, Chen CJ, Morgan TW, Xiao S, Grier HE, Kozakewich HP, et al. Congenital mesoblastic nephroma t(12;15) is associated with ETV6-NTRK3 gene fusion: cytogenetic and molecular relationship to congenital (infantile) fibrosarcoma. The American journal of pathology. 1998;153(5):1451-8.
39. Knezevich SR, McFadden DE, Tao W, Lim JF, Sorensen PH. A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. Nature genetics. 1998;18(2):184-7.
40. Wu G, Diaz AK, Paugh BS, Rankin SL, Ju B, Li Y, et al. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nature genetics. 2014;46(5):444-50.

41. Wiesner T, He J, Yelensky R, Esteve-Puig R, Botton T, Yeh I, et al. Kinase fusions are frequent in Spitz tumours and spitzoid melanomas. Nature communications. 2014;5:3116.
42. Greco A, Miranda C, Pierotti MA. Rearrangements of NTRK1 gene in papillary thyroid carcinoma. Molecular and cellular endocrinology. 2010;321(1):44-9.
43. Leeman-Neill RJ, Kelly LM, Liu P, Brenner AV, Little MP, Bogdanova TI, et al. ETV6-NTRK3 is a common chromosomal rearrangement in radiation-associated thyroid cancer. Cancer. 2014;120(6):799-807.
44. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. Cancer. 2016;122(7):1097-107.
45. Ricarte-Filho JC, Li S, Garcia-Rendueles ME, Montero-Conde C, Voza F, Knauf JA, et al. Identification of kinase fusion oncogenes in post-Chernobyl radiation-induced thyroid cancers. J Clin Invest. 2013;123(11):4935-44.
46. Vaishnavi A, Capelletti M, Le AT, Kako S, Butaney M, Ercan D, et al. Oncogenic and drug-sensitive NTRK1 rearrangements in lung cancer. Nat Med. 2013;19(11):1469-72.
47. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. Nature communications. 2014;5:4846.
48. Suh JH, Johnson A, Albacker L, Wang K, Chmielecki J, Frampton G, et al. Comprehensive Genomic Profiling Facilitates Implementation of the National Comprehensive Cancer Network Guidelines for Lung Cancer Biomarker Testing and Identifies Patients Who May Benefit From Enrollment in Mechanism-Driven Clinical Trials. Oncologist. 2016;21(6):684-91.
49. Ardini E, Bosotti R, Borgia AL, De Ponti C, Somaschini A, Cammarota R, et al. The TPM3-NTRK1 rearrangement is a recurring event in colorectal carcinoma and is associated with tumor sensitivity to TRKA kinase inhibition. Molecular oncology. 2014;8(8):1495-507.
50. Sigal D, Tartar M, Xavier M, Bao F, Foley P, Luo D, et al. Activity of Entrectinib in a Patient With the First Reported NTRK Fusion in Neuroendocrine Cancer. J Natl Compr Canc Netw. 2017;15(11):1317-22.
51. Frattini V, Trifonov V, Chan JM, Castano A, Lia M, Abate F, et al. The integrated landscape of driver genomic alterations in glioblastoma. Nature genetics. 2013;45(10):1141-9.
52. Kim J, Lee Y, Cho HJ, Lee YE, An J, Cho GH, et al. NTRK1 fusion in glioblastoma multiforme. PloS one. 2014;9(3):e91940.
53. Subramaniam DS, Xiu J, Mehta S, Gatalica Z, Swensen J, N S. RNA-Seq analysis of glioma tumors to reveal targetable gene fusions. Journal of Clinical Oncology. 2017;35(15 (suppl)):2019.
54. Yamamoto H, Yoshida A, Taguchi K, Kohashi K, Hatanaka Y, Yamashita A, et al. ALK, ROS1 and NTRK3 gene rearrangements in inflammatory myofibroblastic tumours. Histopathology. 2016;69(1):72-83.
55. Ross JS, Wang K, Gay L, Al-Rohil R, Rand JV, Jones DM, et al. New routes to targeted therapy of intrahepatic cholangiocarcinomas revealed by next-generation sequencing. Oncologist. 2014;19(3):235-42.
56. Kralik JM, Kranewitter W, Boesmueller H, Marschon R, Tschurtschenthaler G, Rumpold H, et al. Characterization of a newly identified ETV6-NTRK3 fusion transcript in acute myeloid leukemia. Diagnostic pathology. 2011;6:19.
57. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature. 2016;531(7592):47-52.
58. F. Hoffmann-La Roche Ltd. Integrated analysis CSR - Clinical overview (entrectinib). 2018.
59. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med. 2017;23(6):703-13.

60. Pavlick D, Schrock AB, Malicki D, Stephens PJ, Kuo DJ, Ahn H, et al. Identification of NTRK fusions in pediatric mesenchymal tumors. Pediatric blood & cancer. 2017;64(8).
61. Chmielecki J, Bailey M, He J, Elvin J, Vergilio JA, Ramkissoon S, et al. Genomic Profiling of a Large Set of Diverse Pediatric Cancers Identifies Known and Novel Mutations across Tumor Spectra. Cancer Res. 2017;77(2):509-19.
62. Astsaturov IA, Ellis P, Swensen J, Arguello D, Reddy SK, Gatalica Z, et al. Fusion analysis of solid tumors to reveal novel rearrangements in breast carcinomas. Journal of Clinical Oncology. 2016;34(15_suppl):11504-.
63. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A, et al. ALK, ROS1, and NTRK Rearrangements in Metastatic Colorectal Cancer. J Natl Cancer Inst. 2017;109(12).
64. Musholt TJ, Musholt PB, Khaladj N, Schulz D, Scheumann GF, Klempnauer J. Prognostic significance of RET and NTRK1 rearrangements in sporadic papillary thyroid carcinoma. Surgery. 2000;128(6):984-93.
65. Pillai RK, Jayasree K. Rare cancers: Challenges & issues. The Indian journal of medical research. 2017;145(1):17-27.
66. Preusser M, Winkler F, Valiente M, Manegold C, Moyal E, Widhalm G, et al. Recent advances in the biology and treatment of brain metastases of non-small cell lung cancer: summary of a multidisciplinary roundtable discussion. ESMO open. 2018;3(1):e000262.
67. Christensen TD, Spindler KL, Palshof JA, Nielsen DL. Systematic review: brain metastases from colorectal cancer--Incidence and patient characteristics. BMC cancer. 2016;16:260.
68. Jin J, Gao Y, Zhang J, Wang L, Wang B, Cao J, et al. Incidence, pattern and prognosis of brain metastases in patients with metastatic triple negative breast cancer. BMC cancer. 2018;18(1):446.
69. Roughley A, Damonte E, Taylor-Stokes G, Rider A, Munk VC. Impact of Brain Metastases on Quality of Life and Estimated Life Expectancy in Patients with Advanced Non-Small Cell Lung Cancer. Value Health. 2014;17(7):A650.
70. Ni W, Chen W, Lu Y. Emerging findings into molecular mechanism of brain metastasis. Cancer medicine. 2018;7(8):3820-33.
71. Lockney NA, Wu AJ. Alectinib for the management of ALK-positive non-small cell lung cancer brain metastases. Journal of thoracic disease. 2017;9(2):E152-E4.
72. Baik CS, Chamberlain MC, Chow LQ. Targeted Therapy for Brain Metastases in EGFR-Mutated and ALK-Rearranged Non-Small-Cell Lung Cancer. Journal of Thoracic Oncology. 2015;10(9):1268-78.
73. National Institute for Health and Care Excellence. Brain tumours (primary) and brain metastases in adults NICE guideline [NG99] 2018 [Date Accessed July 2018]. Available from: https://www.nice.org.uk/guidance/ng99/chapter/Recommendations.
74. Moraes FY, Shultz DB, Murphy ES, Khuntia D, Chao ST, JH S. Strategies to mitigate the effects of whole-brain radiation therapy on neurocognitive function in patients with brain metastases. Appl Rad Oncol 2017;6(1):6-10.
75. Kristensen A, Solheim TS, Fløtten Ø, Grønberg BH. Associations between hematologic toxicity and health-related quality of life during first-line chemotherapy in advanced non-small-cell lung cancer: a pooled analysis of two randomized trials. Acta Oncologica. 2018;57(11):1574-9.
76. Botchkarev VA. Molecular mechanisms of chemotherapy-induced hair loss. The journal of investigative dermatology Symposium proceedings. 2003;8(1):72-5.
77. Karczmarek-Borowska B, Pelc AUM, Rabiej AUE, Grądalska-Lampart AUM. The quality of life of non-small cell lung cancer patients treated with chemotherapy. The quality of life of non-small cell lung cancer patients treated with chemotherapy. 2014;82(4):349-57--57.
78. Teixeira RJ, Applebaum AJ, Bhatia S, Brandao T. The impact of coping strategies of cancer caregivers on psychophysiological outcomes: an integrative review. Psychol Res Behav Manag. 2018;11:207-15.

79. Fenn KM, Evans SB, McCorkle R, DiGiovanna MP, Pusztai L, Sanft T, et al. Impact of Financial Burden of Cancer on Survivors' Quality of Life. Journal of Oncology Practice. 2014;10(5):332-8.
80. MacMillan Cancer Support. Cancer's hidden price tag 2013 [Date Accessed 30th April 2019]. Available from: https://www.macmillan.org.uk/_images/Cancers-Hidden-Price-Tag-report-England_tcm9-270862.pdf.
81. Lange AM, Lo HW. Inhibiting TRK Proteins in Clinical Cancer Therapy. Cancers. 2018;10(4).
82. Drugs.com. Vitrakvi Approval History 2018 [Date Accessed 30th April 2019]. Available from: https://www.drugs.com/history/vitrakvi.html.
83. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. The New England journal of medicine. 2018;378(8):731-9.
84. Cook PJ, Thomas R, Kannan R, de Leon ES, Drilon A, Rosenblum MK, et al. Somatic chromosomal engineering identifies BCAN-NTRK1 as a potent glioma driver and therapeutic target. Nature communications. 2017;8:15987.
85. Renfro LA, Sargent DJ. Statistical controversies in clinical research: basket trials, umbrella trials, and other master protocols: a review and examples. Annals of oncology : official journal of the European Society for Medical Oncology. 2017;28(1):34-43.
86. Yan L, Zhang W. Precision medicine becomes reality-tumor type-agnostic therapy. Cancer communications (London, England). 2018;38(1):6.
87. European Medicines Agency. Personalised Medicines - report from a workshop on personalised medicines held by EMA on 14 March 2017. 2017.
88. American Cancer Society. Chemotherapy Side Effects 2016 [Date Accessed 30th April 2019]. Available from: https://www.cancer.org/content/cancer/en/treatment/treatments-and-side-effects/treatment-types/chemotherapy/chemotherapy-side-effects/.
89. Skalova A, Vanecek T, Majewska H, Laco J, Grossmann P, Simpson RH, et al. Mammary analogue secretory carcinoma of salivary glands with high-grade transformation: report of 3 cases with the ETV6-NTRK3 gene fusion and analysis of TP53, beta-catenin, EGFR, and CCND1 genes. The American journal of surgical pathology. 2014;38(1):23-33.
90. Spiro RH. Management of malignant tumors of the salivary glands. Oncology (Williston Park, NY). 1998;12(5):671-80; discussion 83.
91. Cancer Research UK. Prevalence of STS 2019 [Accessed February 2019]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/soft-tissue-sarcoma/incidence#heading-Four.
92. Italiano A, Mathoulin-Pelissier S, Cesne AL, Terrier P, Bonvalot S, Collin F, et al. Trends in survival for patients with metastatic soft-tissue sarcoma. Cancer. 2011;117(5):1049-54.
93. Eli Lilly and Company. Press Release 2019 [Accessed February 2019]. Available from: https://investor.lilly.com/news-releases/news-release-details/lilly-reports-results-phase-3-soft-tissue-sarcoma-study.
94. Cancer Research UK. UK Lung Cancer Statistics. 2019.
95. Neel DS, Bivona TG. Resistance is futile: overcoming resistance to targeted therapies in lung adenocarcinoma. NPJ precision oncology. 2017;1.
96. Goldberg SB, Oxnard GR, Digumarthy S, Muzikansky A, Jackman DM, Lennes IT, et al. Chemotherapy with Erlotinib or chemotherapy alone in advanced non-small cell lung cancer with acquired resistance to EGFR tyrosine kinase inhibitors. Oncologist. 2013;18(11):1214-20.
97. F. Hoffmann-La Roche Ltd. A Phase 1, Dose Escalation Study of Entrectinib (ALKA-372-001) In Adult Patients With Advanced/Metastatic Solid Tumors 2018.
98. F. Hoffmann-La Roche Ltd. Interim Clinical Study Report GO29365 - Report No. 1078954 2018.
99. F. Hoffmann-La Roche Ltd. GO40782 (STARTRK-2): An Open-Label, Multicenter, Global Phase II Basket Study of Entrectinib for the Treatment of Patients with Locally

Advanced or Metastatic Solid Tumors that Harbor NTRK1/2/3, ROS1, or ALK Gene Rearrangements. 2018.
100. F. Hoffmann-La Roche Ltd. Supplementary results report for study STARTRK-NG (CO40778) in pediatric patients. 2018.
101. Demetri GD, Paz-Ares L, Farago AF, Liu SV, Chawla SP, Tosi D, et al., editors. Efficacy and safety of entrectinib in patients with NTRK fusion-positive tumours: pooled analysis of STARTRK-2, STARTRK-1 and ALKA-372-001. European Society of Medical Oncology (ESMO); 2018.
102. F. Hoffmann-La Roche Ltd. Integrated analysis CSR - Summary of clinical safety (entrectinib). 2018.
103. F. Hoffmann-La Roche Ltd. A Phase 1, Multicenter, Open-Label Study of Oral Entrectinib (RXDX-101) in Adult Patients With Locally Advanced or Metastatic Cancer Confirmed to be Positive for NTRK1, NTRK2, NTRK3, ROS1, or ALK Molecular Alterations (STARTRK-1). 2018.
104. F. Hoffmann-La Roche Ltd. Primary Clinical Study Report (STARTRK-NG)– Phase 1/1b, Open-Label, Dose-Escalation and Expansion Study of Entrectinib (RXDX-101) in Children and Adolescents with Recurrent or Refractory Solid Tumors and Primary CNS Tumors, with or without TRK, ROS1, or ALK Fusions. Report No. 1089445.; 2018.
105. F. Hoffmann-La Roche Ltd. Integrated analysis CSR - Summary of clinical efficacy_NTRK (entrectinib). 2018.
106. Kazandjian D, Blumenthal GM, Luo L, He K, Fran I, Lemery S, et al. Benefit-Risk Summary of Crizotinib for the Treatment of Patients With ROS1 Alteration-Positive, Metastatic Non-Small Cell Lung Cancer. Oncologist. 2016;21(8):974-80.
107. F. Hoffmann-La Roche Ltd. Statistical analysis plan - Evaluation of the Efficacy and Safety of Oral Entrectinib (RXDX-101) for the Treatment of Patients With Locally Advanced or Metastatic Solid Tumors That Harbor NTRK Gene Fusions. 2018.
108. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. The extent of population exposure to assess clinical safety for drugs intended for long-term treatment of non-life-threatening conditions E1.; 1994.
109. De Braud FG, Niger M, Damian S, Bardazza B, Martinetti A, Pelosi G. Alka-372-001: first-in-human, phase I study of entrectinib—an oral pan-trk, ROS1, and ALK inhibitor—in patients with advanced solid tumors with relevant molecular alterations. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(suppl; abstr 2517).
110. Desai AV, Brodeur GM, Foster J, Shusterman S, Sabnis AJ, Macy M, et al. STARTRK-NG: A phase 1/1b study of entrectinib in children and adolescents with advanced solid tumors and primary CNS tumors, with or without TRK, ROS1, or ALK fusions. Proceedings: AACR Annual Meeting 2017; April 1-5, 2017; Washington, DC. 2017.
111. Drilon A, Sankhala KK, Liu SV, Cho BC, Blakely C, Chee CE, et al. STARTRK-2: A global phase 2, open-label, basket study of entrectinib in patients with locally advanced or metastatic solid tumors harboring TRK, ROS1, or ALK gene fusions. Proceedings: AACR Annual Meeting 2017; April 1-5, 2017; Washington, DC. 2017.
112. Drilon A, Siena S, Ou SI, Patel M, Ahn MJ, Lee J, et al. Safety and Antitumor Activity of the Multitargeted Pan-TRK, ROS1, and ALK Inhibitor Entrectinib: Combined Results from Two Phase I Trials (ALKA-372-001 and STARTRK-1). Cancer discovery. 2017;7(4):400-9.
113. Patel MR, Bauer TM, Liu SV, Drilon AE, Wheler JJ, Shaw AT. STARTRK-1: phase 1/2a study of entrectinib, an oral Pan-Trk, ROS1, and ALK inhibitor, in patients with advanced solid tumors with relevant molecular alterations. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(suppl; abstr 2596).
114. Desai AV, Brodeur GM, Foster J, Berg SL, Basu EM, Shusterman S, et al. Phase 1 study of entrectinib (RXDX-101), a TRK, ROS1, and ALK inhibitor, in children, adolescents, and young adults with recurrent or refractory solid tumors. Journal of Clinical Oncology. 2018;36(15_suppl):10536-.

115. F. Hoffmann-La Roche Ltd. Data Summary (31 May 2018 compared with D75 data) - entrectinib integrated efficacy evaluable population 2019.
116. F. Hoffmann-La Roche Ltd. Clinical Expert Opinion (Data on File). 2019.
117. FDA. Ignyta Granted Breakthrough Therapy Designation for Entrectinib by U.S. Food and Drug Administration 2017 [Date Accessed 2nd May 2019]. Available from: https://www.drugs.com/nda/entrectinib_170515.html.
118. European Medicines Agency. EMA issues PRIME designation for entrectinib in NTRK fusion solid tumours 2017 [Date Accessed 2nd May 2019]. Available from: https://ecancer.org/news/12591-ema-issues-prime-designation-for-entrectinib-in-ntrk-fusion-solid-tumours.php.
119. Hanna N, Shepherd FA, Fossella FV, Pereira JR, De Marinis F, von Pawel J, et al. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2004;22(9):1589-97.
120. Herbst RS, Baas P, Kim DW, Felip E, Perez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet (London, England). 2016;387(10027):1540-50.
121. Shepherd FA, Dancey J, Ramlau R, Mattson K, Gralla R, O'Rourke M, et al. Prospective randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinum-based chemotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2000;18(10):2095-103.
122. Tabernero J, Yoshino T, Cohn AL, Obermannova R, Bodoky G, Garcia-Carbonero R, et al. Ramucirumab versus placebo in combination with second-line FOLFIRI in patients with metastatic colorectal carcinoma that progressed during or after first-line therapy with bevacizumab, oxaliplatin, and a fluoropyrimidine (RAISE): a randomised, double-blind, multicentre, phase 3 study. The Lancet Oncology. 2015;16(5):499-508.
123. Demetri GD, von Mehren M, Jones RL, Hensley ML, Schuetze SM, Staddon A, et al. Efficacy and Safety of Trabectedin or Dacarbazine for Metastatic Liposarcoma or Leiomyosarcoma After Failure of Conventional Chemotherapy: Results of a Phase III Randomized Multicenter Clinical Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2016;34(8):786-93.
124. Laurie SA, Siu LL, Winquist E, Maksymiuk A, Harnett EL, Walsh W, et al. A phase 2 study of platinum and gemcitabine in patients with advanced salivary gland cancer: a trial of the NCIC Clinical Trials Group. Cancer. 2010;116(2):362-8.
125. Ali A, Goffin JR, Arnold A, Ellis PM. Survival of patients with non-small-cell lung cancer after a diagnosis of brain metastases. Current oncology (Toronto, Ont). 2013;20(4):e300-e6.
126. Eichler AF, Loeffler JS. Multidisciplinary management of brain metastases. Oncologist. 2007;12(7):884-98.
127. Inno A, Di Noia V, D'Argento E, Modena A, Gori S. State of the art of chemotherapy for the treatment of central nervous system metastases from non-small cell lung cancer. Transl Lung Cancer Res. 2016;5(6):599-609.
128. FDA. FDA Announces First Approval of Targeted Therapy Based on Basket Study 2017 [Date Accessed 2nd May 2019]. Available from: https://www.mskcc.org/trending-topics/fda-announces-first-approval-targeted-therapy-based-basket-study.
129. ESMO. ESMO 2016 Press Release: Single-Arm Trials Improve Early Access to Rare Cancer Drugs 2016 [Date Accessed 2nd May 2019]. Available from: https://www.esmo.org/Conferences/Past-Conferences/ESMO-2016-Congress/Press-Media/Single-Arm-Trials-Improve-Early-Access-to-Rare-Cancer-Drugs.
130. Bertrand T, Kothe M, Liu J, Dupuy A, Rak A, Berne PF, et al. The crystal structures of TrkA and TrkB suggest key regions for achieving selective inhibition. Journal of molecular biology. 2012;423(3):439-53.

131. Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU TSD 19: Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. Avaliable at: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf; accessed January 2019. 2017.
132. Schwartz LH, Litière S, de Vries E, Ford R, Gwyther S, Mandrekar S, et al. RECIST 1.1-Update and clarification: From the RECIST committee. European journal of cancer (Oxford, England : 1990). 2016;62:132-7.
133. Latimer N. NICE DSU Technical Support Document 16: Adjusting Survival Time Estimates in the Presence of Treatment Switching 2014 [Available from: http://www.nicedsu.org.uk/TSD16_Treatment_Switching.pdf.
134. Latimer N. NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA. NICE Decision Support Unit. 2013.
135. Dolan P, Gudex C, Kind P, A W. A Social Tariff for EuroQoL: Results from a UK General Population Survey. University of York, Discussion Paper 138. 1995.
136. Lane S, Levy AR, Mukherjee J, Sambrook J, Tildesley H. The impact on utilities of differences in body weight among Canadian patients with type 2 diabetes. Current medical research and opinion. 2014;30(7):1267-73.
137. National Institute for Health and Care Excellence. Eribulin for treating locally advanced or metastatic breast cancer after 1 chemotherapy regimen [TA515] 2018 [Date Accessed 15th May 2019]. Available from: https://www.nice.org.uk/guidance/ta515.
138. British National Formulary (BNF). Date updated January 2019. Availabe at: https://www.medicinescomplete.com/mc/bnf/current/ 2019 [
139. National Institute for Health and Care Excellence. Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy [TA520] 2018 [Date Accessed 30th January 2019]. Available from: https://www.nice.org.uk/guidance/ta520.
140. National Institute for Health and Care Excellence. Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer [TA476] 2017 [Date Accessed 31st January 2019]. Available from: https://www.nice.org.uk/guidance/ta476/resources/paclitaxel-as-albuminbound-nanoparticles-with-gemcitabine-for-untreated-metastatic-pancreatic-cancer-pdf-82604969382085.
141. Department of Health. 2017/18 reference cost data 2017 [Date Accessed 31st January 2019]. Available from: https://improvement.nhs.uk/documents/1973/2_-_National_schedule_of_reference_costs_v2.xlsx.
142. Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care 2018 [Date Accessed 31st January 2019]. Available from: https://kar.kent.ac.uk/70995/1/Unit%20Costs%202018%20-%20FINAL%20with%20bookmarks%20and%20covers.pdf.
143. Curtis L, Burns A. PSSRU Unit Costs of Health and Social Care 2018. Avalabe at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/; accesed January 2019.: University of Kent, Canterbury, Unit PSSR; 2018.
144. Georghiou TB, M. Exploring the cost of care at the end of life. Nuffield Trust. 2014.
145. F. Hoffmann-La Roche Ltd. Data on file. 2019.
146. Scottish Science Advisory Council. Informing the Future of Genomic Medicine in Scotland. 2019.
147. National Institute for Health and Care Excellence (NICE). TA406: Crizotinib for untreated anaplastic lymphoma kinase-positive advanced nonsmall-cell lung cancer 2016 [Date Accessed 17th May 2019]. Available from: https://www.nice.org.uk/guidance/ta406/resources/crizotinib-for-untreated-anaplastic-lymphoma-kinasepositive-advanced-nonsmallcell-lung-cancer-pdf-82604546118853.

148. ThermoFisher Scientific. Oncomine Focus Assay https://www.thermofisher.com/ch/en/home/clinical/preclinical-companion-diagnostic-development/oncomine-oncology/oncomine-focus-assay.html2019 [
149. Chief Medical Officer. Generation Genome: Annual Report of the Chief Medical Officer 2016. 2017.
150. Clinical Trials.Gov. National Lung Matrix Trial: Multi-drug Phase II Trial in Non-Small Cell Lung Cancer. 2018;NCT02664935.
151. National Institute for Health and Care Excellence. Vismodegib for treating basal cell carcinoma. Technology Appraisal Guidance. 2017;TA489.

B.5 Appendices

Appendices are provided as a separate document this Document B.

- Appendix C: Summary of product characteristics (SmPC) and European public assessment report (EPAR)
- Appendix D: Identification, selection and synthesis of clinical evidence
- Appendix E: Subgroup analysis
- Appendix F: Adverse reactions
- Appendix G: Published cost-effectiveness studies
- Appendix H: Health-related quality-of-life studies
- Appendix I: Cost and healthcare resource identification, measurement and valuation
- Appendix J: Clinical outcomes and disaggregated results from the model
- Appendix K: Checklist of confidential information
- Appendix L: Identification of clinical outcomes for economic analysis
- Appendix M: Inclusion and exclusion criteria in the ALKA, STARTK-1, STARTRK-2, and STARTRK-NG trials

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Clarification questions

July 2019

File name	Version	Contains confidential information	Date
ID1512 entrectinib clarification questions_company response_4Jul2019 [ACIC]_v2.0	V2	Yes	04.07.2019

Section A: Clarification on effectiveness data

Entrectinib trials

A1. Priority question: Please provide individual participant data for the latest data cut of the integrated efficacy analysis population, including the following patients groups: primary CNS tumours, paediatric patients, and previous TKI use. The following variables are requested: Tumour type (please be as specific as possible), line of therapy, response (ORR), time to progression, mortality events and censoring time.

Individual patient response data by line of therapy and tumour type for the integrated analysis population, adult primary CNS tumours and paediatric populations available to date are provided below. However, these data should be interpreted with caution due to the differing methods by which response was assessed across the three patient groups. As discussed at the clarification teleconference on 17th June, there is no previous tyrosine kinase inhibitor subgroup in the NTRK fusion-positive population of the STARTRK trial programme. Unfortunately, due to legal and governance reasons, Roche is not able to provide any further patient-level data beyond that given below. However, if still relevant in light of the detailed response data provided here and where feasible, Roche may be able to conduct further prospective analyses requested by the ERG.

The independent review committee (IRC)-assessed response data by tumour type and line of therapy are based on the more recent [redacted] clinical cut-off date (see Figure 1 and Table 1**)**. Caution should be exercised in the interpretation of these data as response for CNS tumours is measured according to different criteria (Response Assessment in Neuro-Oncology Criteria, RANO) than systemic solid tumours, which are measured according to RECIST v1.1 (Response Evaluation Criteria In Solid Tumours).

In the adult primary CNS tumour population, investigator-assessed response data are available for five patients. However, the IRC data only include one primary CNS tumour patient, as IRC data from the four STARTRK-2 adult primary CNS patients are not available. Investigator-assessed response rates are provided in Figure 2 and Table 2. Investigator-assessed response data are available for all five adult primary

CNS tumour patients. Also of note, in the company submission (CS), reference is made to six adult patients with primary CNS tumours; however, one patient was excluded from analysis due to a protocol deviation (the patient was ECOG PS3).

In the paediatric population, only investigator-assessed response data are available. In addition, one patient did not have measurable disease at baseline, and therefore is not represented in the waterfall plot.

Figure 1: IRC-assessed response data available to date by tumour type and line of therapy – Best % change from baseline

Table 1: IRC-assessed response data available to date by tumour type and line of therapy

Efficacy Evaluable Population including CNS Primary and Paediatric (N=66)			
		IRC Assessed	
#lines of prior systemic therapy	Tumour	Response status	Maximum %Change in Sum Lesion Diameter
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]	[redacted] [redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Efficacy Evaluable Population including CNS Primary and Paediatric (N=66)			
		IRC Assessed	
#lines of prior systemic therapy	Tumour	Response status	Maximum %Change in Sum Lesion Diameter
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		[redacted]

Efficacy Evaluable Population including CNS Primary and Paediatric (N=66)			
		IRC Assessed	
#lines of prior systemic therapy	Tumour	Response status	Maximum %Change in Sum Lesion Diameter
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		[redacted]
[redacted]			

Figure 2: Investigator-assessed response rates by tumour type and line of therapy – Best % change from baseline

Table 2: Investigator-assessed response rates by tumour type and line of therapy

Efficacy Evaluable Population including CNS Primary and Paediatric (N=66)			
		Investigator Assessed	
#lines of prior systemic therapy	Tumour	Response status	Maximum %Change in Sum Lesion Diameter
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Efficacy Evaluable Population including CNS Primary and Paediatric (N=66)			
		Investigator Assessed	
#lines of prior systemic therapy	Tumour	Response status	Maximum %Change in Sum Lesion Diameter
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Efficacy Evaluable Population including CNS Primary and Paediatric (N=66)			
		Investigator Assessed	
#lines of prior systemic therapy	Tumour	Response status	Maximum %Change in Sum Lesion Diameter
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]		[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			

A2a. Please detail the testing regime used to identify patients with NTRK fusions in the ALKA and STARTRK trials and confirm whether NTRK fusions were confirmed by FISH or NGS either locally or centrally.

For the patients in the efficacy evaluable cohort, the sponsor included patients with a confirmed NTRK fusion either from local or central testing. The testing utilised was a nucleic acid based technology, which includes both fluorescence in situ hybridisation (FISH) and next-generation sequencing (NGS). For the 54 efficacy evaluable cohort, 52 were enrolled by NGS testing, 1 by polymerase chain reaction (PCR) and 1 by NanoString. If patients had a local result, tissue (if available) was required to be sent for central confirmation. In total, 44/54 efficacy evaluable patients had tissue available for testing and a central positive NTRK fusion result. For the non-fusion patients in ALKA and STARTRK-1, we did not routinely collect molecular reports, so we cannot summarise that information. All patients in the efficacy evaluable population had molecular confirmation of fusion status.

A2b. Please comment on whether there is a risk that a small proportion of trial patients who were observed to be NTRK positive may have been false positives.

There is currently no accepted gold standard for the determination of NTRK fusion status. Given the differences in technologies and potential cutoffs there may be a small risk of false positives.

A3. Please comment on why only one patient with an NTRK 2 fusion was recruited to the ALKA and STARTRK trials, and on the biological plausibility of differences in response to entrectinib in this fusion type.

The low patient recruitment number seen for patients with a confirmed NTRK2 fusion is reflective of the prevalence of this fusion within the wider population. Based on data from the Foundation Medicine cancer genomics database NTRK2 fusions constitute a small minority ([redacted]) of all NTRK fusion types. The absolute prevalence of NTRK2 across all tumour types based on Foundation Medicine data is [redacted]. For further details on the Foundation Medicine Data, please refer to the response to question A8. In the larotrectinib efficacy analysis in TRK fusion positive adults and children, only 2% of patients had NTRK2 fusions (1). In terms of differences in potential responses to entrectinib based on NTRK fusion type, there are currently insufficient data available to comment on this.

A4. Figure 4 on Page 40 of the main submission mentions a 'natural history follow-up cohort' of patients who are NTRK positive but did not enrol in the STARTRK-2 study. Is there any evidence available for this cohort, including demographics as per Table 9 of the main submission, and outcomes as listed in Figure 4.

As indicated within the schematic on page 40 of the CS, there were a group of patients which had tested positive for NTRK1/2/3, ROS1 or ALK fusions and consented but were not subsequently enrolled into the interventional studies. To date, only two NTRK fusion positive patients fell into this category, one with metastatic thyroid cancer and another with a metastatic salivary gland tumour. Both these patients died, although time to death information is not available and the data would be of very limited use given the very small sample size.

Population

A5. Priority question: Please clarify the positioning of entrectinib in terms of lines of therapy for each tumour type, and what current practice is for that line. Table 6 on page 30 of the main submission mentions 'second line and beyond'

but it needs to be clear if this refers to 2nd line, 3rd line, etc. as this may impact upon the choice of comparator data.

The anticipated indication for entrectinib is "" [redacted] " The proposed lines of therapy by tumour type are provided in the comparator table provided in Appendix L of the CS. A complete version of this table is provided in response to question A15 (Table 3).

A6. Priority question: Table 6 on page 30 of the main submission proposes the use of entrectinib for 10 cancer sites, but Table 4 on page 23 indicates at least 22 potential cancers with NTRK fusions would be covered by the licensed indication, with even more identified in the literature. Please can the company clarify the population they intend this technology to be used in on the NHS – as per the licensed indication, or for only those tumour sites represented in the entrectinib trials?

The anticipated indication for entrectinib is " [redacted] " Roche intends for entrectinib to be used as per the licensed indication within the NHS. This is initially anticipated to be via the Cancer Drugs Fund (CDF) route, with a comprehensive data collection plan in place (see section B.2.11.1 of CS for provisional details of this) to address clinical and cost-effectiveness uncertainties, such as those associated with tumour types not included in the integrated analysis.

A7. Section B1.3.5 states that the company are carrying out a review of the prognosis of patients with NTRK fusions. Has any relevant data arisen from this review? Can you please provide references included in this review?

This literature review is currently being conducted, with completion due at the end of July/early August. Roche commits to providing the findings of this review to NICE at the earliest opportunity.

A8. On page 24 of the main submission, reference is made to a cohort of 116,698 patients tested for NTRK fusions. Can you please provide further details of this cohort and a breakdown of the prevalence of NRTK 1,2, and 3 fusions by tumour type (please be as specific as possible regarding the tumour type)?

The patient cohort mentioned on page 24 of the CS refers to NGS profiling of 116,698 adult and paediatric tumour samples using the Foundation Medicine Inc. (FMI) NGS platform. An updated analysis was conducted in Q4 2018 consisting of a total of 166,067 patients; the overall prevalence of NTRK gene fusions remained at [redacted] in this updated analysis. Absolute prevalence of NTRK1, 2 and 3 fusion types was [redacted] and [redacted], respectively. The results of this analysis, broken down by NTRK1, 2 and 3 fusions and tumour type are provided in appendix A.

Testing

A9. Priority question: The proposed testing regime suggests using immunohistochemistry (IHC) tests to initially screen patients. Given the broad anticipated marketing authorisation this is likely to imply a substantial increase in the number of IHC tests being processed nationally. Please comment on the plausibility of implementing this, given current testing infrastructure.

Roche acknowledges the challenges associated with the introduction of the pan-tumour indication proposed for entrectinib. Particular consideration has been given to the testing strategy detailed within the CS. As the NHS continues to move towards increased uptake of NGS-based genomic profiling, it has been necessary for Roche to propose a hybrid strategy to testing in order to facilitate patient access to Entrectinib at this current time. This hybrid approach, as the ERG identifies, is based on the implementation of an immunohistochemistry (IHC) pre-screen. Given the potential number of tumour types covered by the marketing authorisation, Roche also acknowledges that there is likely to be an associated increase in demand for IHC testing to identify NTRK-positive patients. However, Roche sees this hybrid approach as an interim step in the short-to-medium term to support initial patient access. In the longer term, we envisage the majority of testing being undertaken via NGS in line with the NHS's vision for genomic testing. Existing diagnostic testing

infrastructure already supports large scale IHC-based testing with large numbers of tumour types listed in the National Genomic Test Directory for Cancer (2). Since IHC is an established testing method, Roche anticipates that disruption due to implementation of NTRK testing will be manageable, as NTRK is added to existing testing directories. Finally, Roche intends to work in partnership with the NHS to support NTRK test implementation and ongoing service needs as testing is embedded.

A10. For mammary analogue secretory carcinomas (MASC), the National Genomic Test Directory lists FISH or RT-PCR as the test administered on the NHS. We understand that these tests only identify ETV6-NTRK3 fusions. Please confirm this is the case and state which NTRK fusions were displayed in the MASC patients recruited to the ALKA and STARTRK trials?

We can confirm that for the clinical indication of Secretory Carcinoma (Salivary Gland), the current version of the National Genomic Testing Directory lists the test technology as ETV6-NTRK3 FISH/RTPCR. The indicated gene target for these tests is listed as only ETV6-NTRK3. Within the efficacy evaluable population (n=54) for entrectinib, there were a total of 7 MASC patients who had ETV6-NTRK3 fusions.

A11a. Please provide any data on file regarding the sensitivity and specificity of the IHC pan-TRK Ventana test[1].

Pan-TRK IHC has been shown to be a resource efficient method that may serve as an adjunct to genetic testing for the assessment of NTRK fusions. The VENTANA pan-TRK assay (EPR17341) has not been optimised to delineate between TRK wild-type and chimeric-fusion proteins. As a result, sensitivity and specificity data are limited to the data presented within the package insert. Please see appendix B for a copy of the package insert.

A11b. Please confirm the cost of this pan-TRK Ventana test.

The current list price of the VENTANA Pan-Trk (EPR17341) assay is [redacted] per test (Catalog Number: 790-7026). Within the economic modelling the price associated with the VENTANA assay is [redacted] (PPS, Price Per Slide) this is due to the incorporation of additional slide preparation costs ([redacted]).

[1] http://reagent-catalog.roche.com/product/1909?type=2442

A12. Please provide further details on the referenced 89% specificity of the IHC approach described on page 126 of the main submission, and the data supporting this figure.

As detailed within the CS, a UK clinical trial site was consulted on the IHC positivity rate for NTRK fusion positive tumours. Based on the real-world experience of the Christie Hospital trial facility, investigator feedback on patient screening observed a positivity rate of 11%. These patient samples were subsequently passed on for centralised confirmation as per study protocol.

It should be noted that this 11% figure refers to any degree of positivity by IHC; the investigator reported that the "true" IHC positivity rate may be lower. Consequently, the 11% figure used represents a conservative estimate, in that the true IHC screening costs may in fact be lower.

Comparators

A13. Do you envisage that entrectinib and larotrectinib are likely to be used in the same populations? Do you consider the published larotrectinib studies to provide alternative distributions of patients across tumour sites.

Roche cannot comment on Bayer's anticipated positioning or licence for larotrectinib. However, in the event that a similar tumour-agnostic licence is granted to larotrectinib and it is subsequently used as per this licence, then it is likely that entrectinib and larotrectinib will be used in similar patient populations.

The entrectinib integrated analysis population was the result of an extensive screening programme of approximately [redacted] patients, with patients screened by over 150 sites across 15 countries; consequently, Roche believes that the integrated analysis population is likely to be reflective of clinical practice. Roche cannot comment on the conduct of the tumour-agnostic NTRK fusion-positive larotrectinib studies, or the screening methods used. However, it is possible given the similar profiles of the two drugs that the larotrectinib trial populations could be an alternative distribution of patients.

For the ERG's information, a study has been conducted at the Memorial Sloan Kettering Cancer Center in the United states to identify a cohort of NTRK fusion-positive patients, which shows a tumour-type distribution. The results of this study

are presented in appendix C. While a formal comparison with the entrectinib integrated analysis population has not been made, in general it is similar.

A14. Priority question: The anticipated marketing authorisation for entrectinib states that patients will be eligible for treatment with entrectinib as an "[redacted]." Can you please comment on what therapies are likely to fall in this category of unacceptability?

As per the decision criteria defined in section B.2.9.1 of the CS, we have defined chemotherapy, hormone therapy and best supportive care as the therapeutic classes informing comparator choice for entrectinib according to the anticipated licence and consequent position of entrectinib in the treatment pathway. These criteria were the result of clinician input and discussion during consultations with NHSE and NICE (please refer to Data on File document in Appendix D). During an advisory board in December 2018, clinicians suggested that entrectinib should be used "in patients in which chemotherapy is not acceptable due to intolerance or lack of efficacy" (see Data on File document in Appendix D. It is difficult define an "unacceptable" therapy in absolute terms due to the heterogeneity of available treatment options across multiple different rumour types; for example, an "acceptable" PFS or ORR outcome may be different between non-small cell lung cancer and thyroid cancer. To reiterate the anticipated position of entrectinib in the care pathway, Roche anticipates use in later lines of treatment in the majority of cases, at the point where treatment options are very limited or exhausted altogether. For details of the comparators used in the economic analysis, and therefore what may be considered "unacceptable", please refer to the table provided in response to question A15.

A15. Priority question: Table 30 in Appendix L of the company submission does not appear to be complete (for example non-small-cell lung cancer (NSCLC) is not included). Please provide details on the sources of PFS and OS evidence for tumour sites missing from this table.

A section of the table was omitted from the original CS in error. The complete table is shown below (Table 3).

Table 3: Overview of NICE-recommended comparators

NICE TA	Year	Population	Intervention(s)	Line of therapy	Clinical outcomes		
					ORR (%)	Median PFS (m)	Median OS (m)
Breast cancer							
TA515	2018	-Locally advanced or metastatic -1 prior chemotherapy regimen	Capecitabine	2L	11.5	4.1	14.5
TA423	2016	-Locally advanced or metastatic disease ≥2 prior chemotherapy regimen	Eribulin	3L+	12.2	3.6	13.2
TA423 (aggregated "physician's choice" comparators)	2016	-Locally advanced or metastatic disease ≥2 prior chemotherapy regimen	Vinorelbine	3L+	4.7	2.2	10.5
	2016	-Locally advanced or metastatic disease ≥2 prior chemotherapy regimen	Gemcitabine + paclitaxel		4.7	2.2	10.5
				Average of medians		3.0	12.2
				Average of exponential means		4.4	17.6
Non-small-cell lung cancer							
TA520 (mixed histology)	2018	Locally advanced or metastatic disease	Docetaxel	2L+	13.4	3.4	9.6

TA428 (mixed histology)	2017	≥ 1 previous chemotherapy regimen			9.0	4	8.5
TA483 (squamous histology)	2017				9.0	2.8	6
TA484 (non-squamous histology)	2017				12.0	4.2	9.4
TA403 (mixed histology)	2016				13.6	3	9.1
TA347 (non-squamous histology)	2015				3.6	2.8	10.3
TA124 (non-squamous histology)	2007				8.8	2.9	7.9
					Docetaxel average of medians	3.3	8.7
TA347 (non-squamous histology)	2015	Locally advanced or metastatic disease ≥ 1 previous chemotherapy regimen	Nintedanib + docetaxel	2L+	4.7	4.2	12.6
					Average of medians	3.8	10.7
					Average of exponential means	5.4	15.4
Colorectal cancer							

TA307	2014	Advanced or metastatic disease Progression following oxaliplatin-based therapy	FOLFIRI	2L	11.1	4.7	12.1
TA242	2012	Advanced or metastatic disease Following first line chemotherapy	Irinotecan	2L	34.8	6.2	15.6
TA405	2016	Advanced or metastatic disease Following previous treatment with available therapies	Trifluridine-tipiracil	3L	0.9	2	9
			Trifluridine-tipiracil		1.6	2	7.2
			Best supportive care		0.0	1	6.6
			Best supportive care		0.0	1.7	5.2
					Average of medians	2.6	9.1
					Average of exponential means	3.8	13.1
Neuroendocrine tumours (refractory/unsuitable for lutetium therapy)							
TA449 and TA539	2017 and 2018	Unresectable or metastatic neuroendocrine tumours	Everolimus (pancreatic NET)	1L	4.8	11	44.0
			Best supportive care (pancreatic NET)		2	4.6	37.7*
			Everolimus (GI/Lung NET)		2	11	37.2

			Best supportive care (GI/Lung NET)		1	3.9	39.6*
					Average of medians	8.0	39.6
					Average of exponential means	11.6	57.1
Pancreatic tumours							
TA476	2017	Metastatic disease	Gemcitabine + nab-paclitaxel	1L	23	5.5	8.7
			Gemcitabine monotherapy		7	3.7	6.6
NICE Guideline NG85	2018	Metastatic disease	FOLFIRINOX	1L	31.6	6.4	11.1
					Average of medians	5.2	8.8
					Average of exponential means	7.5	12.7
Papillary and anaplastic thyroid cancer (unsuitable/progressed following radioactive iodine)							
TA535	2018	Locally advanced or metastatic disease Unresponsive to radioactive iodine	Best supportive care	2L+	1.5	3.7	19.1 (after cross-over adjustment)
			Best supportive care	2L+	0.5	5.4	42.8*
					Average of medians	4.6	31.0

					Average of exponential means	6.6	44.7
Soft tissue sarcoma							
TA465	2017	Advanced disease Unsuitable for curative surgery or unresponsive to radiotherapy	Doxorubicin	1L+	7.5	4.1	14.7
TA185	2010	Locally advanced or metastatic disease Relapsed/refractory following one anthracycline and ifosfamide	Trabectedin	2L+	5.1	3.3	13.9
					Average of medians	3.9	14.3
					Average of exponential means	5.6	20.6

A16. Priority question: Please provide a detailed description of the methods used to identify and select comparator data:

As discussed in section B.2.9 of the CS, a standard systematic review to identify comparator data was not feasible due to the broad indication covered by the decision problem, since a scoping search identified in excess of 1,000,000 publications. As a result, comparators were identified using the NICE Pathways website as a starting point.

a. The search terms used to identify appropriate guidance.

The search terms used on the NICE Pathways website were limited to the tumour type under review: "lung cancer", "breast cancer", "colorectal cancer", etc.

b. The criteria used in the selection process.

Where search terms resulted in multiple possible pathways on the NICE Pathways website, for example "breast cancer, advanced" and "breast conditions", the pathway relevant to the decision problem was chosen, for example the pathway referring to management of advanced/metastatic patients. Therapies were selected from the pathways on the basis of the decision criteria described in section B.2.9.1 of the CS (described in further detail in the data on file document provided in response to question A18).

c. Documentation of selection decisions and reasons for excluding potentially relevant guidance.

No documentation was created regarding selection and exclusion decisions since the decision criteria were adhered to.

d. Documentation on decision regarding which median values to extract e.g. where multiple values are presented.

No documentation was created regarding selection of which median value was extracted where multiple values were presented. Generally, technology appraisals were informed by one randomised controlled trial, and there was only one median value provided for each outcome that was relevant to the decision problem or the scope of the technology appraisal for the given comparator. Subgroup values were not used unless they were pertinent to the scope. Where multiple median values

were presented as a consequence of multiple follow-up analyses, the value from the primary analysis was used. The clinical data used for the selected comparators was validated with clinical experts in each field, who agreed that the data was as per their expectations, taking into the consideration that the median values are derived from clinical trial populations.

A17. Priority question: Please justify the decision to draw values from multiple sources of guidance for each comparator, given that entrectinib will only likely be used in one position in the pathway and that more recent guidance will supersede older guidance. See also question A5.

Comparator efficacy data was drawn from multiple technology appraisals for individual comparators in the same line of therapy where available (for example, docetaxel in NSCLC). This decision was taken to increase the robustness of the comparator data, by taking a mean of multiple values, and to ensure that an outlying or extreme value was not inadvertently used. In most cases, the majority of technology appraisals for individual comparators were conducted within a few years of each other.

A18. Priority question: The submission quotes reference 116 F. Hoffmann-La Roche Ltd. Clinical Expert Opinion (Data on File). 2019. Please provide this reference and any additional information as appropriate.

Please see appendix D for this Data on File document. This Data on File describes how the decision criteria were arrived at, the clinical opinion that informed them, and the clinical expert validation of comparator choice.

End-of-life

A19. Priority question: Please comment on whether the company expect end-of-life criteria to be met across all or a subset of the patients potentially eligible to receive entrectinib.

As per section B.2.13.3, Roche anticipates entrectinib to meet end-of-life criteria across all patients potentially eligible for entrectinib, on the basis of the results of the comparative analysis of the integrated population. It should be noted that the anticipated positioning of entrectinib is in later lines of treatment for most tumour types, where there is either no available therapy or outcomes of existing therapy are

poor; consequently, many patients will be nearing the end of their treatment pathway and available therapeutic options. Therefore, this is consistent with an "end-of-life" position.

According to the comparator data sourced for tumour types included in the integrated analysis population, the only tumour type which may have a survival prognosis of more than two years is neuroendocrine; however, the best supportive care data used to inform the survival outcome for this tumour type was confounded by cross-over to active treatment. It is also important to consider that, even in the assessment of an intervention for a single tumour type with a poor prognosis such as lung cancer, where EoL criteria are clearly met, a proportion of patients may be expected to live well in excess of two years (see survival data in TA520 (3), TA484 (4), TA483 (5), TA428 (6)).

The prognostic implications of NTRK gene fusions are also a consideration for end-of-life criteria. Limited data available to date suggest that patients with NTRK gene fusions perform less well on current standard of care than patients without (7). In summary, Roche expects entrectinib to meet end-of-life criteria for the population as defined by the anticipated licence and the appraisal final scope.

In summary, Roche expects entrectinib to meet end-of-life criteria for the population defined by the anticipated and licence and the appraisal scope.

Section B: Clarification on cost-effectiveness data

B1. Priority question: Please add the latest trial data cut-off for progression-free survival and overall survival to the economic model and present these results.

The latest trial data cut-off for progression-free survival and overall survival for the integrated analysis population has been added to the economic model. This data is from the [redacted] clinical cut-off date, the details of which were provided in part B.2 of the original CS. The update was implemented after other requested model corrections were made. As a result, the new base case incremental cost-effectiveness ratio (ICER) based on the updated data cut is £52,609. The original scenario and sensitivity analyses have been re-run and the results are presented in appendix E. This base case includes consideration of screening costs for the

identification of eligible patients; when these costs are excluded the ICER decreases to £36,914.

Additional key scenarios include limiting the duration of subsequent therapy to 3 months and 6 months, reducing the ICER to £39,849 and £40,093, respectively. Further tumour weighting sensitivity analyses have been conducted on all tumour types, whereby 100% weighting of costs and outcomes were applied. Plausible ICERs reported are versus breast cancer (£34,854), CRC (£38,303), lung cancer (£46,240) and sarcomas (£36,927).

B2. Priority question: The integrated efficacy analysis data set currently excludes the following patients: Patients with primary CNS tumours, previous TKI use, and paediatric patients. Please provide further justification for the exclusion of these patients and confirm whether you expect any NICE recommendation to include these patient groups? Please present a scenario analysis which includes these patients in the economic model.

The integrated analysis population consists of 54 adult solid tumour patients, excluding primary CNS tumours. The reasons that this data set excludes both primary CNS tumour and paediatric patients are:

- At the time of the clinical cut-off date (CCOD) for the primary analysis (31st May 2018), sufficient survival follow-up was only available for the integrated analysis patient population
- For primary CNS tumour patients, in addition to the lack of follow-up, response is measured according to different criteria for tumours in the CNS (Response Assessment in Neuro-Oncology Criteria, RANO); this makes comparison or combination with patients who have systemic solid tumours (measured according to Response Evaluation Criteria In Solid Tumours, RECIST v1.1), challenging
- For both the adult primary CNS and paediatric patient populations, response and PFS were assessed by the investigator, as opposed to IRC in the integrated analysis population; these inconsistent methods mean any interpretation of combined data should be conducted with caution

- Regarding the economic analysis, as discussed on the checkpoint teleconference on 10th April, inclusion of paediatric tumours is very challenging due the absence of any robust comparator data for these patients, particularly those with infantile fibrosarcoma (IFS)

As discussed at the ERG clarification teleconference on 17th June, there is no previous tyrosine kinase inhibitor subgroup in the NTRK fusion-positive population of the STARTRK trial programme.

The anticipated indication for entrectinib is "[redacted]." Roche anticipates a NICE recommendation in accordance with the proposed licence, which at the present time includes paediatric patients and adult patients with primary CNS tumours.

A scenario analysis has been conducted whereby the five efficacy-evaluable adult primary CNS tumour patients and the seven paediatric patients have been added to the model. It should be noted that, for the purposes of the model, the paediatric primary CNS patients have been grouped with the adult primary CNS patients for the weighted comparator costs and outcomes, since common comparators are assumed for these patients. In addition, one of the four primary CNS paediatric patients was a CNS embryonal tumour rather than glioma, however for the purpose of the model this patient has been grouped with the glioma patients.

Screening data, comparator outcomes and costs, and utilities have been sourced for each tumour type where possible. These are summarised below.

Screening data

The National Genomic Test Directory indicates that paediatric solid tumours should undergo whole genome sequencing, whilst adult gliomas are subject to testing for other specific biomarkers. The frequency of NTRK fusions in the two additional populations are shown in Table 4.

Table 4: Frequency of NTRK fusions in adult primary CNS and paediatric tumours

Tumour type ((8), Appendix A)	NTRK Fusion Rate	Number needed to screen
Glioma	[redacted]	[redacted]
Infantile Fibrosarcoma	100%	1
Melanoma	[redacted]	[redacted]

Comparator data

Comparator data for the new tumour types has been sourced using the same decision criteria and methods as for the integrated analysis population, with the NICE Pathways website informing comparator choice. For glioma, chemotherapy has been chosen as the standard of care for recurrent glioma, since treatment-naïve glioma is treated using surgery and/or radiotherapy with or without chemotherapy. For temozolomide, outcome data was extracted from NICE TA23. However, no outcome data was found in NICE literature for procarbazine, CCNU (lomustine) and vincristine (PCV) or single-agent CCNU; therefore, studies referenced in the ESMO clinical practice guideline for high grade glioma (9) were used. Dosing information to derive costs for PCV chemotherapy was derived from the literature (10). For utility data, a pragmatic search was conducted to identify a value for recurrent glioblastoma patients. For PFS utility, a value of 0.731, derived from a study by the Peninsula Technology Assessment Group on carmustine implants in glioma, was selected (11). For PPS, no utility values were found in the search and therefore an average of the known PPS utilities in the economic analysis was used.

For paediatric melanoma, single-agent dacarbazine was chosen as the comparator for patients in whom targeted therapy or immunotherapy is unsuitable. Outcome data was sourced from a pivotal phase III study referenced in NICE Guideline 14 for the assessment and management of melanoma. Utility data were sourced from NICE TA357 - Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab; health states approximating the timing of PFS and PPS were chosen.

As previously discussed, the identification of a comparator for IFS is challenging due to the rare occurrence of this tumour type and the lack of standardised therapy. In addition, the literature is limited to case reports and case series. The literature

suggests that chemotherapy in combination with surgery (which may include amputation) is a common treatment approach in newly-diagnosed IFS. Chemotherapies used include vincristine, adriamycin, and cyclophosphamide (adria-VAC); vincristine, actinomycin-D, and cyclophosphamide (actino-VAC); and etoposide and ifosfamide (12). Survival rates are reported to be between 89–94%, therefore minimisation of toxic chemotherapy is described as the present challenge (13). An older case series reports recurrence rates of 32–33% (14). In order to select an appropriate comparator for IFS, a pragmatic approach was taken whereby best supportive care for recurrent disease is used (in accordance with the defined decision criteria which excludes surgery as a comparator), with outcomes and utilities being an average of the known comparator data for other tumour types.

The comparator data used in the scenario analysis is shown in Table 5.

Table 5: Overview of NICE-recommended comparators for adult primary CNS and paediatric tumours

Source	Year	Tumour type	Treatment	Line of therapy	ORR	PFS	OS	PFS Utility (source)	PPS Utility (source)
High grade glioma (after surgery/radiotherapy)									
TA23	2016	Recurrent grade III or IV glioma	Temozolomide	2L	5.4	2.89	7.34	0.73 (11)	0.60 (average of known)
Brada M *et al*, 2010 (15)	2010	Recurrent grade III or IV glioma	Procarbazine, CCNU (lomustine) and vincristine	2L	NR	3.6	6.7		
Batchelor T *et al*, 2013 (16)	2013	Recurrent grade III or IV glioma	Single agent CCNU (lomustine)	2L	14.4	3.0	9.8		
Infantile Fibrosarcoma (after surgery/chemotherapy)									
NA	NA	Recurrent infantile fibrosarcoma	Best supportive care	2L	NA	4.1 (average of known)	15.8 (average of known)	0.73 (average of known)	0.60 (average of known)
Malignant melanoma									
NICE guideline NG14 (Middleton MR *et al*, 2000) (17)	2000	Previously treated stage IV melanoma	Dacarbazine	2L+	12.1	1.5	6.4	0.75 (NICE TA357 - ≥180 days to death value)	0.60 (NICE TA357 – 90-180 days to death used as proxy for PPS value)

Scenario analysis results

For the reasons described above, primarily the limited follow-up and differing response/PFS measurements in the two new populations as well as the lack of reliable comparator data for IFS, the results of this scenario analysis should be interpreted with caution.

In this scenario, the ICER is £49,358 (see Table 6 for details). Excluding screening costs, the ICER drops to £35,770.

Table 6: Scenario analysis results (with confidential PAS, includes screening costs)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£61,228	1.61	1.04	[redacted]	[redacted]	[redacted]	£49,358
Entrectinib	[redacted]	[redacted]	[redacted]				

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

B3. Priority question: The model currently includes testing costs in the comparator arm. Please justify why this is the case, given that the majority of patients in NHS practice are not tested for NTRK fusions and as such any testing costs for NRTK fusions would be in addition to those already in place to identify other mutations.

As per table 67 in section B.3.5.4 of the CS, screening costs are applied to the comparator arm for mammary analogue secretory carcinoma (MASC) and sarcoma since NTRK fusion testing is already included in the genomic testing directory for MASC, and whole genome sequencing is reimbursed for sarcoma. The incremental testing cost in these tumour types is therefore zero.

In the cases of colorectal cancer (CRC), Non-small-cell lung cancer (NSCLC), breast cancer and thyroid cancer, testing is routinely conducted for other biomarkers, for example for ALK, EGFR and ROS1 in NSCLC. Consequently, for CRC, NSCLC, breast cancer and thyroid cancer, existing screening costs are applied. However, these costs consist solely of IHC assay costs since genome sequencing is not routinely conducted to detect the existing biomarkers. There is therefore an

incremental testing cost in these tumour types: an additional IHC assay to detect the presence of NTRK protein, and a NGS test in 11% of these patients to confirm the presence of an NTRK gene fusion.

For pancreatic, neuroendocrine and other cancers, no molecular testing is conducted according to the testing directory. Therefore, there are significant incremental screening costs for these patients. In addition, it should also be noted that NGS is likely to detect the presence of other actionable driver mutations where an NTRK gene fusion is not detected.

B4. Priority question: Please justify the selection of an exponential distribution for all comparators. In doing so please make reference to the model prediction that post-progression survival is substantially longer that pre-progression survival, and state whether predicted life years gained were validated against those predicted in the relevant source guidance.

The exponential distribution was the optimal statistical fit based on Akaike information criterion (AIC) and Bayesian information criterion (BIC) on the entrectinib data. In addition, the comparator data limitations encourage the use of the most simplistic (in terms of data requirements) parametric model, such as the exponential distribution.

For the standard of care (SoC), exponential distribution was used for consistency on the comparison with entrectinib, in order to avoid more complicated parametric models that often overestimate long term benefits. In addition, given the fact that median overall survival (OS) and progression-free survival (PFS) are most commonly reported consistently in the literature for SoC regimens and given the mathematical properties of the density function of the exponential distribution, using an exponential model is an appropriate and transparent way to calculate the mean and area under the curve. Any other parametric model would have required substantial assumptions in order to calculate the mean via a reported median. In addition, as discussed in the CS, NICE Technical Support Document 14 (18) states that where only published summary statistics are available, the conversion of median to exponential mean represents a reasonable approach where individual patient data are lacking.

Longer post-progression survival was a consistent observation we have seen from reported survival estimates in the literature and these longer estimates were proportionally similar to what we have observed in entrectinib. We are aware that post-progression survival can be influenced by various factors and therefore validation of these estimates in an unadjusted population on a non-randomised trial setting is very difficult. For that reason, sensitivity analysis was undertaken in order to account for this uncertainty.

Spot checks against relevant source guidance for the two most common tumour types in the integrated analysis (NSCLC and sarcoma) have been conducted to compare life-years gained (LYG). These reveal that the figures in the source guidance are broadly consistent with the calculated LYG in the economic model. For example, NICE TA347 (19) (nintedanib and docetaxel in NSCLC) quote ERG-estimated LYG of 1.49 and 1.23 for nintedanib + docetaxel and docetaxel alone, respectively; this compares with 1.23 in the entrectinib model. NICE TA185 (20) and TA465 (21) indicated LYGs of 1.61 for trabectedin and 2.06 for doxorubicin, respectively; this averages 1.84 LYG, compared with 1.62 in the entrectinib model. These comparisons should of course be interpreted with caution, since the LYG figures in the guidance documents are derived directly from patient-level data while, as described above, the figures in the entrectinib model are derived from conversion of median survival outcomes.

B5. Priority question: The economic model appears to apply discounting twice to the cost of post-progression second-line treatment costs in the entrectinib arm: firstly, when the costs are estimated in the SoC engine sheet (SOC_NTRK+, column AB), and secondly, when the costs are applied in the entrectinib engine sheet (Entrectinib_NTRK+, column BA). It also does not appear to consider the cost of adverse events and administration. Please confirm whether this is the case, and if appropriate, provide a model with corrections to these issues.

The discounting error has been corrected. Formula in column BA in Entrectinib_NTRK tab has been updated. Adverse events are applied in the model as a one-off cost so are therefore not applied a second time in the post-progression health state. As an administration cost should be applied to second-line treatment, this has been corrected in the model in cells K80 and K81 in the Cost_Inputs tab.

B6. Priority question: On the "Settings" Sheet Cell F58 implies there are 7.024 days in a week. This appears to be a minor calculation error. Please confirm whether this is the case, and if appropriate, provide a corrected model.

The number of days in a year is 365.25, hence 7.024 days in a week is calculated by dividing 365.25 by 52. The value has therefore been retained in the model.

B7. In the analysis of the quality of life data from the trials, please confirm the number of patients who were available for analysis at each time point, and the proportion of patients who responded to the questionnaire.

Please refer to Table 7 below for the number of patients who provided EQ-5D responses at each time point. Please see the response to question B8 for clarification as to which patients were included in the analysis.

Table 7: The number of patients completing EQ-5D by visit

Name Visit	n	Percentage, % compared to baseline
Cycle 1 Day 1	48.0	100.0
Cycle 2 Day 1	43.0	89.6
Cycle 3 Day 1	41.0	85.4
Cycle 4 Day 1	36.0	75.0
Cycle 5 Day 1	36.0	75.0
Cycle 6 Day 1	35.0	72.9
Cycle 7 Day 1	31.0	64.6
Cycle 8 Day 1	31.0	64.6
Cycle 9 Day 1	26.0	54.2
Cycle 10 Day 1	26.0	54.2
Cycle 11 Day 1	22.0	45.8
Cycle 12 Day 1	16.0	33.3
Cycle 13 Day 1	15.0	31.3
Cycle 14 Day 1	14.0	29.2
Cycle 15 Day 1	13.0	27.1
Cycle 16 Day 1	12.0	25.0
Cycle 17 Day 1	9.0	18.8
Cycle 18 Day 1	7.0	14.6
Cycle 19 Day 1	5.0	10.4
Cycle 20 Day 1	6.0	12.5

Cycle 21 Day 1	4.0	8.3
Cycle 22 Day 1	3.0	6.3
Cycle 23 Day 1	3.0	6.3
Cycle 24 Day 1	2.0	4.2
Cycle 25 Day 1	1.0	2.1
Cycle 26 Day 1	1.0	2.1
End Of Treatment	11.0	22.9

B8. Priority question: In the analysis of the quality of life data from the trials please confirm which patients were included in the analysis. Did this include the 54 patients in the survival analysis, all entrectinib NTRK+ patients, or all entrectinib patients including those with ALK and ROS mutations?

EQ-5D data were only collected in the STARTRK-2 trial, from which 51 of the 54 integrated analysis population patients came. EQ-5D assessment were collected from 44 of the 51 STARTRK-2 patients during the PFS period. These 44 patients across 9 tumour types contributed 409 observations to the mixed model that was used to estimate the utility. Please refer to the response to question A8 for details of the proportion of patients who responded to the questionnaire at each time point.

It is important to note that a random effects model was used to account for the fact there are repeated observations per subject, with a nested random effect by patient within tumour type included in the modelling.

B9. Priority question: The economic model assumes that a proportion of patients receiving entrectinib will go on to receive second-line therapy, but assumes that these patients continue to receive this until death. Can you comment on the clinical plausibility of this, and consider presenting an alternative scenario in which patients receive second-line therapy for only a proportion of the post-progression period?

The assumption that subsequent therapy is continued until death represents a conservative approach and is a simplification to allow for multiple subsequent therapies. However, the clinical plausibility of subsequent therapy until death is low, since it is likely only to be administered until a second progression event (which may coincide with death for a proportion of patients).

Consequently, two alternative scenarios have been built into the model. On the Cost_Inputs tab, the user can select the subsequent therapy duration from a drop down list of three options: 3 months, 6 months or until death. Additional cut-off times can be added as required. If the base-case option of 'until death' is selected, the ICER is £52,609. If the user chooses to limit the duration of subsequent therapy to 6 months the ICER decreases to £40,093, and if 3 months is selected, the ICER further decreases to £39,849.

B10. Priority question: Please confirm which patients in the efficacy data set received subsequent treatment (by tumour site) after treatment with entrectinib, and where available what they received, and the duration of treatment. Please present a scenario analysis in the economic model which includes the costs of these therapies.

Available subsequent therapy data is provided in Appendix F; this shows the proportion of patients receiving subsequent therapy after entrectinib and lists the therapies used. However, the specific data as requested are not available.

A scenario analysis using this data has not been provided due to the following reasons:

- Several subsequent therapies used within the trials are not recommended by NICE (for example, bevacizumab, olaratumab, pazopanib), therefore the scenario is not relevant to the NHS
- There is limited follow-up of subsequent therapy from the trials, meaning the true duration of subsequent therapy will not be captured, leading to a high degree of uncertainty
- Data of sufficient detail is not available to inform the analysis
  - Several subsequent therapies are list as generic classes such as "investigational drug", "monoclonal antibodies", "protein kinase inhibitors" and "other antineoplastic agents"; as such it is not possible to model these therapies
  - There is insufficient information on the combinations of the listed treatments
  - There is insufficient information on their dosing and administration frequency

- The scenario would likely have a limited impact on the analysis

B11. Please clarify the source of standard error of utilities for all tumour types as presented in Table 51 on page 111 of the main submission.

Standard deviations for utilities were not reported in most of the source documents; where they were, these have been used in the model as standard errors. Where no standard error or deviation was reported, a common standard error of 0.14 was used for all these estimates in order to overcome this issue. The value of 0.14 was chosen as it was broadly consistent with uncertainty around published utility values, and wide enough to cover uncertainty for a Probabilistic Sensitivity Analysis.

B12. Please clarify the distributions used and measures of uncertainty for all non-fixed parameters in Table 70 on page 128 of the main submission. Please clarify whether all distributions were checked for sensible values (e.g. does the Normal distribution for median PFS have a non-zero probability of negative values?)

According to standard modelling guidelines, for costs, we used a log-normal distribution to reflect the skew often found in cost data. For utilities, we use the Gamma instead of Beta, in order to allow for negative values that often can be found in oncology setting. For the variation of survival parameters for entrectinib we use the multivariate normal based on the Cholesky decomposition of the covariance matrix.

For the variation of published survival estimates, due to lack of covariance matrices and correlations reported and the use of an exponential model for the extrapolation, the extrapolated mean is varied around a normal distribution. This was to avoid any assumptions on skewness and allow for a normal range of assessments of the uncertainty around these estimates.

B13. Priority question: Please repeat the tumour weighting sensitivity analysis as presented in Section B.3.8.3 (Table 76) for all other tumour sites, as was done for MASC and pancreatic cancers.

The additional tumour weighting sensitivity analyses have been conducted and are presented with the re-run scenario analyses conducted in response to question B1,

in appendix E. Due to the model updates, the ICERs vs MASC and pancreatic cancers have slightly decreased from £32,373 and £120,713 to £31,064 and £114,524, respectively.

B14. Please clarify how the figure of 34% of patients with advanced or metastatic disease was determined in the budget impact model (hidden BIM sheet, cells I13:M13).

This figure was sourced from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/all-cancers-combined#heading-Two. The data was derived from the Excel spreadsheet available at this page by adding the stage III and stage IV percentages together to arrive at a figure of 34%. Please note that an incorrect figure is referenced in the Budget Impact document itself (46%), and that all calculations stem from the figure of 34%.

Section C: Textual clarification and additional points

C1. Please clarify the last sentence on data extraction in Section L.1.3 on page 60 in Appendix L.

This sentence refers to cases where chemotherapies are listed on the NICE pathways website, but specific technology appraisals for the regimen in the given setting do not exist, and so clinical evidence is not available from that source. An example is the case of FOLFIRINOX in the treatment of pancreatic cancer.

However, we later elected to include these data in the table and source relevant data from elsewhere. Again, in the example of FOLFIRINOX, a citation provided by a clinical expert was used, which is also referred to in NICE Guideline NG85. As a result, this sentence only applies to MASC, where no directly relevant source data was found.

C2. Appendix L page 64: What sources were used to calculate an average progression-free survival for platinum and gemcitabine in MASC?

Since PFS was not available from the referenced study, the PFS value used for MASC was an average of the PFS outcomes sourced for the other tumour types, which resulted in a clinically plausible PFS of 4.35 months.

C3. In Table 16 in Appendix E.1.1, there are [redacted] patients categorised as 'No' for having received 'any prior systemic therapy'. However, the same table states that [redacted] patients had zero prior therapies. Why don't these numbers match, and which value is correct?

The six patients categorised as "No" in Appendix E.1.1 table 16 refers to the number of patients who have never received any prior systemic therapy at all for any stage of disease, including an earlier occurrence. The 20 patients in the same table refer to patients who have never received systemic therapy for this occurrence of advanced/metastatic disease. Therefore, 14 of the 20 first-line patients had received prior therapy for an earlier occurrence of the disease, for example in the neo-adjuvant/adjuvant setting, while 6 had not.

C4. On page 86 of the main submission it is noted that entrectinib was granted a promising innovative medicine (PIM) designation by the MHRA. Can further information be given about when an early access to medicines scheme (EAMS) decision will be made?

A PIM application was made in order to obtain designation from the MHRA and allow for the possibility of an EAMS. However, due to the accelerated timelines associated with the entrectinib regulatory process, the rarity of the mutation and the lack of current testing infrastructure, an EAMS was subsequently deemed to be infeasible.

C5. In table 2 on page 18 of the main submission, the recommended dose of entrectinib for paediatric patients who have the ability to swallow whole capsules is given. Please explain the method of administration and dosage for people who cannot swallow.

In the STARTRK-NG study, patients who were unable to swallow intact capsules were administered with an experimental formulation which could be sprinkled over food. Roche is currently testing GI tube administration of the commercial formulation and is developing a new age appropriate formulation. [redacted].

References

1. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. The New England journal of medicine. 2018;378(8):731-9.
2. NHS England. Genomic Test Directory 2019 [Date Accessed 15th May 2019]. Available from: https://www.england.nhs.uk/publication/national-genomic-test-directories/.
3. National Institute for Health and Care Excellence. Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy [TA520] 2018 [Date Accessed 30th January 2019]. Available from: https://www.nice.org.uk/guidance/ta520.
4. National Institute for health and Care Excellence. Nivolumab for previously treated non-squamous non-small-cell lung cancer [TA484] 2017 [Date Accessed 30th January 2019]. Available from: https://www.nice.org.uk/guidance/ta484.
5. National Institute for Health and Care Excellence. Nivolumab for previously treated squamous non-small-cell lung cancer 2017 [Date Accessed 29th March 2019]. Available from: https://www.nice.org.uk/guidance/ta483.
6. National institute for health and Care Excellence. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy [TA428] 2017 [Date Accessed 30th January 2019]. Available from: https://www.nice.org.uk/guidance/ta428.
7. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A, et al. ALK, ROS1, and NTRK Rearrangements in Metastatic Colorectal Cancer. J Natl Cancer Inst. 2017;109(12).
8. Knezevich SR, McFadden DE, Tao W, Lim JF, Sorensen PH. A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. Nature genetics. 1998;18(2):184-7.
9. Stupp R, Brada M, van den Bent MJ, Tonn JC, Pentheroudakis G. High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology : official journal of the European Society for Medical Oncology. 2014;25 Suppl 3:iii93-101.
10. Solimando DA, Jr., Waddell JA. Procarbazine, Lomustine, and Vincristine (PCV) Regimen for Central Nervous System Tumors. Hosp Pharm. 2017;52(2):98-104.
11. Rogers G, Garside R, Mealing S, Pitt M, Anderson R, Dyer M, et al. Carmustine implants for the treatment of newly diagnosed high-grade gliomas: a cost-utility analysis. Pharmacoeconomics. 2008;26(1):33-44.
12. Loh ML, Ahn P, Perez-Atayde AR, Gebhardt MC, Shamberger RC, Grier HE. Treatment of infantile fibrosarcoma with chemotherapy and surgery: results from the Dana-Farber Cancer Institute and Children's Hospital, Boston. Journal of pediatric hematology/oncology. 2002;24(9):722-6.
13. Yoshihara H, Yoshimoto Y, Hosoya Y, Hasegawa D, Kawano T, Sakoda A, et al. Infantile fibrosarcoma treated with postoperative vincristine and dactinomycin. Pediatrics international : official journal of the Japan Pediatric Society. 2017;59(3):371-4.
14. Blocker S, Koenig J, Ternberg J. Congenital fibrosarcoma. Journal of pediatric surgery. 1987;22(7):665-70.
15. Brada M, Stenning S, Gabe R, Thompson LC, Levy D, Rampling R, et al. Temozolomide versus procarbazine, lomustine, and vincristine in recurrent high-

grade glioma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(30):4601-8.
16. Batchelor TT, Mulholland P, Neyns B, Nabors LB, Campone M, Wick A, et al. Phase III randomized trial comparing the efficacy of cediranib as monotherapy, and in combination with lomustine, versus lomustine alone in patients with recurrent glioblastoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(26):3212-8.
17. Middleton MR, Grob JJ, Aaronson N, Fierlbeck G, Tilgen W, Seiter S, et al. Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2000;18(1):158-66.
18. Latimer N. NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA. NICE Decision Support Unit. 2013.
19. National Institute for Health and Care Excellence (NICE). TA347: Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non-small-cell lung cancer 2016 [Date Accessed 26th June 2019]. Available from: https://www.nice.org.uk/guidance/ta347/resources/nintedanib-for-previously-treated-locally-advanced-metastatic-or-locally-recurrent-nonsmallcell-lung-cancer-pdf-82602612880837.
20. National Institute for Health and Care Excellence (NICE). TA185: Trabectedin for the treatment of advanced soft tissue sarcoma 2016 [Date Accessed 26th June 2019]. Available from: https://www.nice.org.uk/guidance/ta185/resources/trabectedin-for-the-treatment-of-advanced-soft-tissue-sarcoma-pdf-82598497821637.
21. National Institute for Health and Care Excellence (NICE). TA465: Olaratumab in combination with doxorubicin for treating advanced soft tissue sarcoma 2016 [Date Accessed 26th June 2019]. Available from: https://www.nice.org.uk/guidance/ta465/resources/olaratumab-in-combination-with-doxorubicin-for-treating-advanced-soft-tissue-sarcoma-pdf-82604907236293.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Clarification questions follow-up

11th July 2019

File name	Version	Contains confidential information	Date
ID1512_Follow-up clarification question responses_11Jul2019	1.0	Yes	11th July 2019

1. **No justification was provided for the use in the economic model of the exponential models using the new data cut (clarification question B1). Please provide fit statistics (AIC and BIC) for each survival model presented. Please ensure that these results are available for the population excluding patients with CNS primary tumours and children, as in the originally presented analysis, and including these patients, as requested by the ERG in clarification question B2.**

The AIC and BIC statistics for the updated integrated analysis population relating to PFS and OS are shown in Table 1 and Table 2, respectively. The AIC and BIC statistics for the updated integrated analysis population including primary CNS tumour and paediatric patients relating to PFS and OS are shown in Table 3 and Table 4, respectively.

As per the results of the base-case integrated analysis population, the exponential distribution is consistently the best statistical fit according to both AIC and BIC in all scenarios. This therefore provides further justification for the choice of an exponential distribution to extrapolate both PFS and OS data for entrectinib.

Table 1: Statistical goodness-of-fit and hazard trend for PFS – updated integrated analysis population (B1 scenario)

Parametric distribution	AIC	BIC	Hazard trend
Exponential	**255.7**	**257.7**	**Stable**
Weibull	257.4	261.4	Increasing
Log-normal	257.2	261.2	Decreasing
Generalised gamma	258.8	264.7	Decreasing
Log-logistic	257	261	Decreasing
Gompertz	257.7	261.7	Stable

Table 2: Statistical goodness-of-fit and hazard trend for OS – updated integrated analysis population (B1 scenario)

Parametric distribution	AIC	BIC	Hazard trend
Exponential	**180.8**	**182.8**	**Stable**
Weibull	182.2	186.1	Increasing
Log-normal	183.6	187.5	Decreasing
Generalised gamma	184.1	190.1	Increasing
Log-logistic	182.5	186.5	Stable/Decreasing
Gompertz	182.2	186.2	Increasing

Table 3: Statistical goodness-of-fit and hazard trend for PFS – updated integrated analysis population plus primary CNS and paediatric patients (B2 scenario)

Parametric distribution	AIC	BIC	Hazard trend
Exponential	**281.6**	**283.8**	**Stable**
Weibull	282.5	286.9	Increasing
Log-normal	282.8	287.2	Decreasing
Generalised gamma	284.0	290.6	Decreasing
Log-logistic	282.0	286.4	Decreasing
Gompertz	283.4	287.8	Increasing

Table 4: Statistical goodness-of-fit and hazard trend for OS – updated integrated analysis population plus primary CNS and paediatric patients (B2 scenario)

Parametric distribution	AIC	BIC	Hazard trend
Exponential	**207.2**	**209.4**	**Stable**
Weibull	208.0	212.4	Increasing
Log-normal	209.0	213.4	Decreasing
Generalised gamma	210.0	216.5	Increasing
Log-logistic	208.0	212.4	Stable/Decreasing
Gompertz	208.6	212.9	Increasing

2. **Please provide a model that has the functionality to provide results by tumour type, to allow the ERG to verify the results of the analysis presented for question B13 in the clarification document.**

These analyses were originally conducted by the manual removal of comparator data for all but one tumour type in the SOC_NTRK+ tab, resulting in a comparison of the entrectinib total data versus the single tumour type comparator costs and outcomes.

However, as requested by the ERG, a new version of the model provided in response to clarification question B1 is provided here, which includes a drop-down functionality to more easily conduct these analyses. The drop-down menu can be found in cell C23 of the Results Table tab. This functionality has been tested and the results match the scenarios provided in response to question B1.

3. **Given that it is not possible to provide the individual patient data requested in A1, please provide us with the Kaplan-Meier data for PFS and OS for patients who did not respond to entrectinib. Please provide this a) for the population excluding patients with CNS primary tumours and children, and b) including these patients. A tabular rather than graphical presentation would be preferred, including the total number of events, and number at risk over time.**

The requested data are provided below. PFS and OS Kaplan-Meier data for responders and non-responders from the updated data cut of the original integrated analysis population are provided in Figure 1, Table 5, Figure 2 and Table 6. PFS and OS Kaplan-Meier data for responders and non-responders from this population including the additional primary CNS and paediatric patients are provided in Figure 3, Table 7, Figure 4 and Table 8.

Figure 1: PFS Kaplan-Meier curves for responders versus non-responders - updated integrated analysis population (B1 scenario)

[redacted]

Table 5: PFS Kaplan-Meier data for responders and non-responders - updated integrated analysis population (B1 scenario)

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 2: OS Kaplan-Meier curves for responders versus non-responders - updated integrated analysis population (B1 scenario)

Table 6: OS Kaplan-Meier data for responders and non-responders - updated integrated analysis population (B1 scenario)

strata	time_ months	surv	lower	upper	n.risk	n.cum. events

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 3: PFS Kaplan-Meier curves for responders vs non-responders - updated integrated analysis population plus primary CNS and paediatric patients (B2 scenario)

[redacted]

Table 7: PFS Kaplan-Meier data for responders and non-responders - updated integrated analysis population plus primary CNS and paediatric patients (B2 scenario)

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 4: OS Kaplan-Meier curves for responders vs non-responders - updated integrated analysis population plus primary CNS and paediatric patients (B2 scenario)

[redacted]

Table 8: OS Kaplan-Meier data for responders and non-responders - updated integrated analysis population plus primary CNS and paediatric patients (B2 scenario)

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

strata	time_ months	surv	lower	upper	n.risk	n.cum. events
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

These data should be interpreted with caution due to the heterogeneity of the patient populations, particularly in the analysis that includes primary CNS and paediatric patients for the reasons stated in response to clarification question B2, primarily relating to the differing methods by which response was assessed.

In the event that these data are used to conduct a landmark analysis, it should be noted that Roche does not consider this to be appropriate methodology for the following reasons:

- In a previous appraisal (NICE TA489), this approach was not considered appropriate due to the following concerns:

    - The appropriate point in time at which to define non-responders is arbitrary and uncertain, therefore the non-responder control arm could include patients who subsequently respond
    - The number of patients with metastatic disease was very small (n=96); this is considerably more patients than the 54 included in the integrated analysis population
    - There was concern that Gorlin Syndrome and other baseline covariates were not adjusted for; defining relevant covariates in such a heterogeneous group as the integrated analysis population may not be possible
    - Overall survival data was immature; this is also a consideration for the entrectinib data
- The sample size of non-responders is too small to provide a meaningful or robust comparator sample
- The heterogeneity of the population means the two groups are likely to be highly unbalanced in terms of baseline characteristics
- Conducting a landmark analysis pre-selects for non-responders who are likely to perform better than the original baseline population
- It cannot be assumed that entrectinib has no activity in non-responding patients; it is possible that entrectinib may temporarily halt or slow the progression of tumours in non-responding patients (i.e. those with stable disease), thereby improving the outcome of the non-responder group
- The choice of the landmark point is not prospective, leaving the analysis open to bias
- For the primary CNS and paediatric tumours, there is limited follow-up
- It introduces further uncertainty into the model

NICE National Institute for Health and Care Excellence

Patient organisation submission

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	**Jayne Bressington**

2. Name of organisation	**GIST Support UK**
3. Job title or position	**Vice Chair & Trustee GIST Support UK** **&** **Patient Director PAWS-GIST** *(Paediatric Adolescent and Wild-type GIST) a subdivision of GIST Support UK focusing on the forms of GIST affecting younger patients and all with wild-type GIST for whom there are currently no effective treatments where surgery is not possible.*
4a. Brief description of the organisation (including who funds it). How many members does it have?	**GIST Support UK (GSUK) is a registered charity (No. 1129219) formed in April 2009.** **We are a network of GIST cancer patients & carers working with top GIST specialists & National/International groups, to promote best practice. We exist to help GIST patients and their families come to terms with living with GIST cancer and we raise funds to:** • **Stimulate and fund GIST research.** • **Support Patients living with GIST cancer** • **Provide Information for GIST patients and their clinicians** • **Raise awareness of GIST cancer** **We receive no government funding and are run by a board of, currently ten, volunteer trustees who have a close association and experience of GIST cancer. All of our research is funded through donations and fundraising by our supporters. We also receive some funds from pharmaceutical companies (Novartis, Bayer, Bristol Myers Squibb) to assist with hosting patient meetings and provision of patient information literature.** **GSUK is not a membership organisation. Each year we engage with over a thousand GIST patients and carers, both newly diagnosed and longer-term survivors, via:**

	• **our telephone helpline,** • **regional patient carer meetings,** • **PAWS-GIST clinics,** • **our private online patient forum** • **social media Facebook & twitter platforms** **This amounts to many thousands of patient and carer experiences since the charity was formed in 2009.**
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
5. How did you gather information about the experiences of patients and carers to include in your submission?	**GSUK has gathered information about the experiences of patients since it became a charity in 2009.** **GSUK engages with GIST patients, clinicians and researchers both in the UK and internationally to further our understanding of GIST cancer and develop new treatment options. The charity has played a key role in the development and implementation of infrastructure in the UK to support GIST patients, including the National GIST Guidelines & Natonal GIST Tissue Bank. Through our specialist PAWS-GIST initiative we have created the PAWS-GIST clinic at Addenbrookes hospital in Cambridge, for rarer subsets of GIST patients such as those with NTRK fusions who currently do not have effective treatment options.** **Through our work to support GIST patients we gain valuable information about patient experiences. GSUK engages directly with patients in a variety of ways; our private listserve (email forum community) for patients and carers, patient and carer meetings (held 3 times per year in**

	locations throughout the UK), PAWS-GIST clinics (held 3 times each year) and via our telephone helpline (available 24/7). **We have seen the evidence presented at numerous GIST conferences in mainland Europe, USA & UK recommending that Quadruple negative GIST patients should all be tested to see if they have an *NTRK* fusion.** **Quadruple negative GIST patients represent the second largest group of patients who have attended the PAWS-GIST clinic to date. Tests have commenced to see which ones carry an NTRK fusion.**
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	**Many GIST cancer patients manage, with effective treatment, to live relatively normal lives, continuing to work and play as best they can while manging the side effects of treatment. There are some who are fortunate that their GIST cancer is found early and before it has spread, they have it removed while still small and it does not return. This is as close to a cure as currently exists.** **Depending on the extent of disease, surgery can involve quite drastic interventions such as removal of the stomach. Often the disease has reached an advanced stage prior to diagnosis, limiting the potential for surgery to totally remove the cancer. Toxic side effects are also encountered from anticancer therapies, and tolerance of these side effects varies significantly. Side effects to the drug therapies currently available via NHS include hypertension, hypothyroidism, debilitating hand foot syndrome, diarrhoea, fatigue, nausea, skin rashes and so on. The list of side effects is quite extensive but with advice from oncologists, cancer nurse specialists and fellow patients we observe that these can be managed and tolerated by many patients, providing the chance to live longer and live a normal life. However, some patients do not tolerate these drug side effects and are forced to cease treatment. Additionally, existing therapies are often ineffective in halting disease progression for certain sub-groups of patients.** **Living with GIST cancer as a patient and a carer is possible but every day that you wake up you hope that it was a bad dream and that it isn't real. This is a standard defence mechanism for**

	cancer patients and their families. Learning to cope is something that you have to do and the last thing that you want to do as a carer is to give the impression that things will not be OK. You have to give your loved one hope. **The traumas and horrors of living with a type of GIST cancer that does not have a treatment that works can shatter family's lives. Carers take many forms, parents, partners, siblings, children and friends, all desperate to help and save the person that they love.**
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	**Patients and carers are very grateful for the treatments that are available via the NHS.** **Currently for GIST patients this consists of:** • **Surgery** • **Imatinib** • **Sunitinib** • **Regorafenib** **Unfortunately, not all GIST cancers are the same and there are many for whom the above treatments are not effective because either their primary mutation is not targeted by the above treatments or their disease metastasizes beyond the control of the above treatments.** **All GIST patients are currently given the above options. With the advent of knowledge such as the existence of NTRK fusion driven GIST and a specific targeted treatment such as Entrectinib this may in future change the standard treatment pathway.** **We are very grateful that there is some research happening in the world and that a treatment has been discovered for those GIST patients who have an *NTRK* fusion mutation.** **Currently such a treatment is not available via the NHS but we hope that further to this appraisal that it will be available for patients with GIST caused by *NTRK* fusion mutations along with other NTRK fusion treatments so that these drugs can be used in sequence to overcome recurrent resistance mutations should they arise.**

8. Is there an unmet need for patients with this condition?	**Yes.** ***NTRK* mutated GIST patients do not currently have access to a targeted treatment for their type of GIST mutation via the NHS.** **For some patients with particular types of GIST, the anticancer drugs that are currently available are less / not effective. This includes PAWS-GIST patients (which includes those with NTRK fusion). A key reason for this is due to the lack of existing available therapies targeting specific mutations that drive these cancers, demonstrating a significant un-met need for targeted therapies such as Entrectinib.** **N.B. Both the FDA and the EMA highlighted the need for further development of targeted approaches in this NTRK fusion positive solid tumour patient population of unmet need.** **Drilon et al (13) showed that sarcoma, including soft tissue, infantile fibrosarcoma and GIST comprises the largest cohort of cancer patients to harbour NTRK fusions in their study and two other studies identified one patient each with ETV6-NTRK3 fusion GIST. Both patients exhibited wild type KIT/PDGFR/BRAF disease, which are a group of patients forming part of our PAWS-GIST group.**
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	**The advantages of this technology are that Entrectinib:** • **exhibits potent anti-proliferative activity in all NTRK fusion mutations, which are the cause of the cancer.** • **is administered orally** • **is well tolerated** • **is suitable for both adults and children** • **use in studies has resulted in deep and durable systemic responses in NTRK fusion positive**

	patients with solid tumours. Drugs of this type are exactly what rare cancer patients are desperate to find and use to shrink and stop their tumours and get their life back on track. **They are specifically what our PAWS-GIST patients who harbour NTRK fusions need to arrest their cancer.**
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	**The only disadvantage that we can see is being an NTRK fusion cancer patient and not being able to access Entrectinib.** **As with any drug there are side effects but those listed are tolerable and can be managed.**
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	**Our understanding is that the patients currently classified as Quadruple negative GIST patients are the target group where NTRK fusion GIST's can be found. They have been named quadruple Wildtype GIST as they lack abnormalities in the four signalling pathways KIT, PDGFRA, SDH or RAS. Entrectinib specifically targets the NTRK fusion mutation, and has shown clinically meaningful, deep and durable systemic responses in NTRK fusion positive patients and that it is a tolerable drug with a manageable safety profile.** **We understand that it will become standard practice in NHS England this year for all cancer patients undergoing surgery to have their tumours sequenced. The incidence of finding patients for whom NTRK fusion inhibitors such as Entrectinib will be a suitable treatment will increase. We are very excited that this technology is becoming standard practice within the NHS.** **We are already screening all quadruple negative GIST patients who attend the PAWS-GIST clinic at Addenbrookes Hospital, Cambridge to find those with NTRK fusions. We hope they will then have access to Entrectinib.**

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	**The only inequality we can see currently is that other countries are fast tracking tumour agnostic NTRK fusion inhibitors such as Entrectinib to be available to patients with NTRK fusions; Entrectinib received breakthrough designation and priority medicines designation from the US Food and Drug administration (FDA) and European Medicines Agency (EMA) and Sakigake designation by the Japanese health authorities for treating both adult and paediatric patients with NTRK positive solid tumours.** **Currently Entrectinib is not available to patients in the UK.**
Other issues	
13. Are there any other issues that you would like the committee to consider?	**No**
Key messages	
14. In up to 5 bullet points, please summarise the key messages of your submission: • **NTRK fusions are the root cause of some GIST cancers.** • **Entrectinib is a tumour agnostic precision medicine that targets NTRK fusion mutations.** • **Trials have resulted in dramatic results for GIST patients with NTRK fusions.** • **The NHS whole genome sequencing being launched this year will identify the patients with NTRK fusions.**	

- **Using Entrectinib in GIST patients with an NTRK fusion will reduce unnecessary expenditure on other ineffective therapies that are very expensive for the NHS.**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Patient organisation submission

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Roy Castle Lung Cancer Foundation
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Roy Castle Lung Cancer Foundation is a UK wide lung cancer charity. We fund lung cancer research, tobacco control initiatives and work in lung cancer patient care (information, support and advocacy activity). Our funding base is a broad mixture including community, retail, corporate, legacies and charitable trusts. Clearly, our patient group members and contacts are a self-selected group, who have taken the step to seek out information or have accessed specialist support services. As most lung cancer sufferers tend to be older, from lower social class groups and with the five year survival being around 15%, less physically well, we acknowledge that our patients are perhaps not representative of the vast majority of lung cancer patients, who are not so well informed. It is, however, important that the opinions expressed to us, be passed on to NICE, as it considers the place of this product in the management of solid tumours, such as lung cancer
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	None
5. How did you gather information about the experiences of patients and carers to include in your submission?	The Foundation has contact with patients/carers through its UK wide network of over 55 monthly Lung Cancer Patient Support Groups, patient/carer panel, online forums and its Lung Cancer Information Helpline

Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	We are only able to provide information on experience of lung cancer patients. This appraisal is, however, looking at a wider group of patients, with other solid tumours. As we understand it, there is little information available on the specific characteristics of patients with NTRK positive non small cell lung cancer (nsclc). Thus, our comments are for nsclc in general. According to the National Lung Cancer Audit, the one year survival for lung cancer is 37%. Thus, this group of lung cancer patients have a particularly poor outlook. with an obvious impact on family and carers. Symptoms such as breathlessness, cough and weight loss are difficult to treat, without active anti-cancer therapy. Furthermore, these are symptoms which can be distressing for loved ones to observe.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	In recent years, we have seen new therapy options for some patients with Non Small Cell Lung Cancer – Target Therapies and Immunotherapies. There is, however, a need to identify further new targets and therapies for these groups.
8. Is there an unmet need for patients with this condition?	Most definitely

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Entrectinib is the first therapy specifically targeted at NTRK fusion positive disease. Data presented shows ORR of 70% in in NTRK fusion positive nsclc and shows good intracranial response in patients with baseline brain metastasis.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	The side effects associated with the therapy. We understand that, in the main, these were Grade 1 and 2. Most were managed by dose reduction/interruption.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	
Other issues	
13. Are there any other issues that you would like the committee to consider?	
Key messages	

14. In up to 5 bullet points, please summarise the key messages of your submission:

- First targeted therapy being assessed for NTRK positive disease
- Oral therapy
- Data presented shows systemic and intracranial response
-

•

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Patient organisation submission

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Sarcoma UK
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Sarcoma UK is the only cancer charity in the UK focusing on all types of sarcoma. Sarcoma UK works with patients, carers, supporters, health professionals and researchers to drive awareness of sarcoma, promote early diagnosis and improve patient experience. The charity is funded by voluntary donations from supporters who predominantly have a personal connection with the cause. Sarcoma UK is not a membership organisation but has a database of over 8000 active and engaged supporters. We receive no funding from government or other statutory sources.

4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	The Sarcoma UK Support Line has spoken to over 4250 individuals affected by sarcoma, giving us a unique understanding of the impact of living with the condition. The split of our contacts using the Support Line is 50% and 50% carers. This gives us a balanced view of how sarcoma affects all ages and demographics. We also speak directly to patients at our support groups to harness their views about lack of treatment options when surgical resection is not possible.
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Sarcoma is a rare disease with around 100 different subtypes and approximately 4,800 people diagnosed with a soft tissue sarcoma a year. Sarcoma is one of the hardest cancers to diagnose. People visit their GP more times than those with any other form of cancer before being diagnosed with sarcoma. Since setting up the Sarcoma UK Support line in February 2016, we have heard this confirmed from both patients and carers. This is also backed up by respondents to our National Sarcoma Survey on patient experience, published in 2016. The uncertainty of sarcoma is described by our callers. We have patients who call in a cycle aligned to their follow up appointment. Recurrence of local disease is common and not unheard of after 5 or even 10 years. Commonly, it is the patient who picks up a local recurrence, whilst metastatic disease is usually picked up routinely on chest X-ray without any symptomatic suggestion to the patient that something is wrong. The constant fear of recurrence combined with the fear of the unknown is often described by callers, alongside their fears around prognosis and the limited treatment options available to sarcoma

	patients. They tell us that the rare nature of sarcoma means that they have to become experts and the source of further information around their disease. We hear a lot from carers who reflect that a lack of public awareness about sarcoma. They don't know anything about the condition and fail to understand what and why this happening to their loved one. Sarcoma affects all ages, from paediatric patients to the elderly and this is hard on family life, especially for carers who may not be involved in the early stages of diagnosis and treatment. Gough (2011) comments that soft tissue sarcoma patients maintain a good quality of life with moderate symptoms until a rapid decline in the final weeks of their life. We believe this is unique to sarcoma, and is a contrast to other cancers like non-small cell lung cancer where there is a slow deterioration. This has implications for the patient and their families as home life and financial situations can change suddenly. The end of treatment and the introduction of best supportive care is made on average only 3.4 weeks before the end of life, perhaps because of the good quality of life maintained until the end of life. Callers to our support line often report fatigue, pain, limitation to their mobility, impact of treatment to their quality of life and anxiety. The heterogeneity of the disease means that sarcoma patients can have a wide spread of symptoms dependent on the location of the primary tumour and / or the metastatic disease.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	• Limited treatment options overall. • Dated current treatments. • New treatments for sarcoma are not emerging as fast as for other cancer groups. • Lack of clinical trials- National Sarcoma UK Survey 2015 (of 650 sarcoma patients in England and Scotland) found only a third of patients were offered a clinical trial and of these, only 20% took part. This clearly indicates that options are limited for access to new treatments and technologies once the small number of standard treatments have been exhausted. • No personalisation of treatments

8. Is there an unmet need for patients with this condition?	There is a high unmet need for sarcoma patients. Many patients are diagnosed late stage and curative surgery is not an option. Patients have: • Very limited treatment options. • No adjuvant treatment for most sub types, so predominately patient receive surgery with radiotherapy. Local recurrence is common but little option except further surgery. • Very few sub-type specific treatments • No curative intent treatment if there is metastatic disease.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	• Entrectinib is a drug which affects tumours which are confirmed NTRK-fusion positive and it is a step towards personalised medicine for sarcoma (and other patients). • It may reduce soft tissue sarcoma size to allow surgical removal / resection of the tumour, which would previously be untreatable, giving both longer life and quality of life to patients. • It will only be given to patients who have confirmed NTRK-fusion positive tumours. Uptake is likely to be high in the eligible population as patients who receive treatments knowing their tumour will respond feel less of a gamble and risk. • Sarcoma patients are listed on the NHS England directory to have Whole Genomic Sequencing as standard when the service is rolled out in ~July 2019 (current planned date given by Mark Caulfield). They will already have the confirmatory test as routine standard of care. • Oral delivery of treatment will have a huge benefit to patients, will mean less time in hospital and more “daily living”. It will also have less economic impact on both the patients and NHS, requiring fewer visits, with less time away from work, travel to treatment centre, less planning life around appointments. Oral treatment will require less nursing and medical staff time, fewer clinic spaces. • It will give sarcoma patients who are eligible more time with tolerable side effects with the families. Sarcoma patients tend to have a good quality of life.

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	NONE
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Yes – sarcoma patients who have locally advanced or metastatic solid tumours who have already used other therapies and have no other treatment options available except palliative care.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Sarcoma is a rare cancer and unique to its make up is the heterogeneity. We know that in principle, sarcoma patients are younger and able to remain actively engaged in work and family life until very close to the end of their life. For many the time from primary diagnosis to local recurrence or metastatic disease can be years of productive life. It is important that these small numbers of people are not discriminated against because they are unfortunate enough to be diagnosed with a rare cancer. They should have equal access to treatments.

Other issues	
13. Are there any other issues that you would like the committee to consider?	

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- Sarcoma is a less common cancer which has low public awareness
- Patients frequently experience difficult and late diagnosis leading to limited treatment options
- Entrectinib may reduce tumour size to enable effective surgery
- The oral medication regimen is low burden on patients and NHS service
- We fully support the approval of Entrectinib for NTRK-fusion positive patients

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

NICE National Institute for Health and Care Excellence

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☒ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Professional organisation submission

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	NCRI-ACP-RCP-RCR

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	NCRI-ACP-RCP-RCR
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	The main aims of the treatment are to control disease through tumour response/ delaying time to tumour progression and to prolong survival.

or prevent progression or disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Our experts would consider a reduction in tumour size by more than 30% as being clinically significant.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. TRK fusion positive disease is rare in common cancers (~1% patients) and very common in specific rare cancers such as secretory breast carcinoma, mammary analogue secretory carcinoma (MASC) and infantile fibrosarcoma (>90% patients positive). There are currently no targeted therapies available in the clinic to target this oncogenic driver of disease. There is a clear precedent, for example, EGFR inhibitors in NSCLC or BRAF inhibitors in melanoma that inhibiting an oncogenic driver with a selective drug can have significant tumour response and overall survival advantage.
What is the expected place of the technology in current practice?	

9. How is the condition currently treated in the NHS?	Entrectinib is applicable to multiple different solid tumour types that may harbour an NTRK fusion. It isn't possible to cover standard of care treatments across every disease type. In general, our experts are not aware of specific standard of care treatments that are available for the rare tumours listed above. For the more common cancer types, for example, lung, colorectal, breast there are a range of standard-of-care therapies available. For less common cancer types such as pancreatic, cholangiocarcinoma, thyroid, sarcomas and others, there are more limited lines of standard of care treatments available. To date, patients have not routinely been screened for the TRK fusion so the true prevalence in the UK population isn't clear and natural history of these patients in terms of response to standard treatment is uncertain. However, once standard-of-care treatments are exhausted patients would only have the option of best supportive care and if TRK fusion is present they would potentially stand to gain significant benefit from a TRK inhibitor.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	There are currently no guidelines for the management of TRK fusion positive cancers
• Is the pathway of care well defined? Does it vary or are there differences of opinion	It is challenging to describe pathways for every potential disease group this treatment may benefit. More broadly our experts think the issues will be 1) incorporating genomic profiling into pathways of care where (in some circumstances) it may not be currently routine (see further comments below) and 2) defining the

between professionals across the NHS? (Please state if your experience is from outside England.)	optimal line of treatment where entrectinib should be administered. The line of treatment may vary according to disease type - it may be beneficial early in rare tumours where there are few or no standard-of-care options available and for more common cancers where treatment pathways are well defined, proposed as a later line of therapy. Even in the latter scenario there may be a case to administering entrectinib earlier in the treatment journey rather than later, extrapolating from data for EGFR inhibitors in NSCLC, BRAF inhibitors in melanoma etc.(data not yet available for TRK to support either way) In relation to molecular testing there will be disease groups where this is commonplace with minimal impact on pathways of care and other disease groups where it is less familiar with greater impact. Nonetheless the oncology community is familiar and supportive of the concept of precision medicine and can adapt to this change.
• What impact would the technology have on the current pathway of care?	The main impact will be in relation to molecular testing. As above, some disease groups such as lung and colorectal cancer already routinely screen for a number of genomic alterations as part of the current pathway of care prior to making initial treatment decisions. The impact of including another molecular test in this setting should be minimal (although technologies would need to be considered carefully to maximise the use of tissue samples for parallel testing). In other disease groups where molecular testing doesn't currently form part of the diagnostic/treatment pathway, the impact may be slightly greater but this aligns with the wider NHS ambition to expand the

	amount of molecular testing to be offered to cancer patients and would hopefully lead to an effective use of entrectinib in the right (TRK fusion positive) population for cost-efficient use. The other consideration (as commented above) will be in determining the most appropriate line of therapy for entrectinib in the various different disease types.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	Currently TRK fusion testing does not form part of current care therefore it is unknown which patients harbour this alteration. Patients would therefore be treated with standard of care treatment, if available, in an unselected way. With entrectinib, patients would need to be screened in the first instance for the genomic alteration then the treatment selected on the basis of positive TRK fusion testing. This precision medicine approach would permit for selective use (and therefore associated costs) of the drug in those most patients likely to benefit.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Entrectinib should only be prescribed by oncologists in the secondary care setting.

• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Molecular testing would need to be available across multiple different disease types. The introduction of the cancer test directory and re-designation of the Genomic Laboratory Hubs has to some extent made TRK testing available although is currently limited to NTRK1 and NTRK3 in specific indications. The cost of screening all cancer patients for TRK fusion through nucleic acid based testing may currently be prohibitively expensive. An alternative route would be to consider an IHC test (which is cheaper and applicable across all pathology labs), as a first step for broad screening. If positive by IHC only these would go on to a confirmatory DNA/RNA-based test. The test directory would need to be broadened to all NTRK fusions and all relevant disease types if IHC positive. The only exception to this approach would be the rare cancer types with TRK fusion positivity >90% where an initial IHC pre-screen may not be necessary and immediate nucleic-based testing would be appropriate. Education would also be needed for oncology health care professionals on the meaning of TRK fusion, interpretation of genomic results and use of entrectinib.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes, The data available would support the use of entrectinib in TRK fusion positive disease.

• Do you expect the technology to increase length of life more than current care?	Yes. There is no current standard-of-care for patients with TRK fusion positive cancer. By extrapolation from other settings with the use of a targeted therapy in the presence of a genomic driver there are clear benefits in terms of disease response and survival. This is supported by available single arm trial data for entrectinib. It isn't feasible to undertake a randomised trial in this setting due to relatively small numbers of patients with the genomic alteration.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes. There is no current standard-of-care for patients with TRK fusion positive cancer. If tumours respond to entrectinib then in general disease related symptoms will improve.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Yes, the treatment will only be effective in patients with a TRK (or ROS1/ALK) fusion. The approximate prevalence of TRK fusion in different malignancies (taken from Cocco et al, NTRK fusion-positive cancers and TRK inhibitor therapy, Nature Reviews Clinical Oncology 15, 731-747, 2018) is as follows: >90% frequency - MASC, Secretory breast carcinoma, infantile fibrosarcoma 5-25% frequency – thyroid, GIST, spitzoid tumours, <5% - lung, breast, colorectal, sarcoma, cholangiocarcinoma, melanoma, haematological, head & neck, high-grade glioma
The use of the technology	

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	The treatment is similar to other small molecule targeted therapies so should not be challenging in terms of treatment administration. It is an oral therapy administered once daily. There are no unusual additional clinical requirements outside of the genomic pre-screening required prior to administration. The side effect profile is easily manageable.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Molecular pre-screening is required, as discussed, to identify patients with TRK fusion for the therapy. An initial pre-screen step by IHC is feasible with only positive cases going on to DNA/RNA-based testing for confirmation. Stopping rules would be according to standard practice in terms of radiological or clinical progression or unacceptable toxicity.

15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes. This is one of the first drugs to be applicable in a pan-disease setting where the presence of the genomic alteration drives potential benefit rather than disease type. The data available support the pan-disease benefit which is novel compared with other targeted therapies where the benefit is often limited to a single disease setting and where use of the drug in another disease type (even though the same molecular alteration may be present) doesn't result in the same degree of benefit.
• Is the technology a 'step-change' in the management of the condition?	There are currently no standard-of-care therapies available for TRK fusion positive disease therefore, the technology does represent a step-change in the management of patients with this genomic alteration and sets a precedent for the applicability of pan-disease treatment in the right genomically selected population and fits with NHS ambition for precision medicine based on genetic testing.

• Does the use of the technology address any particular unmet need of the patient population?	It is notable that the drug has CNS penetration; therefore, patients with brain metastases also stand to benefit from this therapy. This is an important consideration as often patients with brain mets have poor outcomes with few available treatments having CNS activity.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The side effect profile is tolerable. Some patients require a dose reduction or interruption but rarely need to discontinue due to toxicity (3.9% patients in the clinical trials).
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Three clinical trials have evaluated entrectinib in patients with TRK/ROS1 or ALK fusion positive disease: ALKA (EudraCT 2012-000148-88), STARTRK-1 (NCT02097810), and STARTRK-2 (NCT02568267). These studies have enrolled patients across 150 sites in 15 countries including the UK. All were single arm studies. There are no control arms to compare with UK practice but patient groups recruited were representative of the UK population (in terms of disease types recruited).
• If not, how could the results be extrapolated to the UK setting?	

• What, in your view, are the most important outcomes, and were they measured in the trials?	Response rate, progression free survival, intracranial activity and overall survival. All these parameters were measured in the trials.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not as far as we are aware.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. How do data on real-world experience compare with the trial data?	To date entrectinib has not been used outside of the clinical trial setting.

	It may be feasible to review real-world data on historical TRK fusion-positive patients (likely in the US where molecular profiling is common) to assess outcomes on standard-of-care therapy as a potential comparator for outcomes on entrectinib
Equality	
21a. Are there any potential equality issues that should be taken into account when considering this treatment?	Access to molecular pre-screening (whether by IHC or nucleic acid-based approach) would need to be considered.
21b. Consider whether these issues are different from issues with current care and why.	
Key messages	

22. In up to 5 bullet points, please summarise the key messages of your submission.

- Novel pan-disease drug indication
- Requirement for routine screening across the NHS for TRK fusions
- Consider pre-screen approach carefully - perhaps IHC as first step then confirmatory DNA/RNA testing
- Line of treatment indication will vary according to disease type
- Randomised data are not available to compare with standard-of-care. Comparisons with real-world data will be required.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

Clinical expert statement

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	Dr Debashis Sarker
2. Name of organisation	King’s College London

3. Job title or position	Senior Lecturer and Consultant in Medical Oncology; Cancer Lead, London South Genomic Laboratory Hub
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☒ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the	☐ yes

rest of this form will be deleted after submission.)	
The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aim of treatment is to improve survival in patients with NTRK fusion positive solid tumours by delaying tumour related progression.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	For patients with NTRK fusion positive solid tumours, a clinically significant treatment response would be defined according to RECIST 1.1 criteria, ie partial response (at least a 30% decrease in size of the longest diameter of target lesions) or complete response (disappearance of all target lesions).
9. In your view, is there an unmet need for patients and	There is an unmet need with regards to treatment of NTRK fusion positive solid tumours. There are no molecularly targeted therapies currently approved for these malignancies. NTRK fusions occur at a high incidence (>90%) in certain rare cancers (e.g. infantile fibrosarcoma, secretory breast carcinoma), whilst occurring at low incidence (commonly <1%) in many common cancers (e.g. lung, colorectal, breast and

healthcare professionals in this condition?	pancreatic cancer). Many of these malignancies are associated with limited treatment options and poor prognosis for these patients.
What is the expected place of the technology in current practice?	
10. How is the condition currently treated in the NHS?	NTRK fusions are found across a range of adult and paediatric cancers: • Cancers enriched for NTRK fusions (frequency >90%, e.g. secretory breast carcinoma, infantile fibrosarcoma, mammary analogue secretory carcinoma) • Cancers harbouring NTRK fusions at intermediate frequency (5-25% e.g. spitzoid melanoma, gastrointestinal stromal tumours, papillary thyroid carcinomas) • Cancers harbouring NTRK fusions at low frequencies (<5%, e.g. lung adenocarcinoma, colorectal carcinoma, high grade glioma) These malignancies are currently treated in the NHS with a variety of standard therapies including surgery, radiotherapy and systemic therapies specific to each tumour type. For many of these malignancies, there are limited lines of systemic therapy for advanced disease (eg pancreatic adenocarcinoma, sarcomas, cholangiocarcinoma). However, there are no current NTRK targeted drugs available for NTRK fusion positive cancers.

• Are any clinical guidelines used in the treatment of the condition, and if so, which?	The following guidelines have been published with regards screening for NTRK fusions: • Marchio C, Scaltriti M, Ladanyi M et al. ESMO recommendations on the standard methods to detect NTRK fusions in daily practice and clinical research. Ann Oncol. 2019 Sep 1;30(9):1417-1427 • Penault-Llorca F, Rudzinski ER, Sepulveda AR. Testing algorithm for identification of patients with TRK fusion cancer. J Clin Pathol. 2019 Jul;72(7):460-467. Specific recommendations for screening TRK fusion cancers in children: • Albert CM, Davis JL, Federman N et al. TRK Fusion Cancers in Children: A Clinical Review and Recommendations for Screening. J Clin Oncol 2019; 37: 513-524. There are currently no clinical guidelines for treatment of NTRK fusion cancers.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Given the very diverse and heterogenous group of malignancies with NTRK fusions (adult+paediatric; common+rare malignancies), a single unifying pathway of care will be very difficult to define. However, common issues across all malignancies are primarily related to the appropriate screening for NTRK fusions to determine eligibility for entrectinib. The National Genomic Test Directory specifies which genomic tests are commissioned by the NHS in England, the technology platform by which the testing will be delivered and clinical requirements for access to the test. This process thereby seeks to address any variation in quality and access to genetic testing across the country and to standardise the commissioning and contracting model for genomics in England. Currently seven Genomic Laboratory Hubs (GLHs) are responsible for delivering the new genomic testing service, working with a network of local genomic laboratory partners, for defined geographic regions. NHS England currently defines a national approach to next generation sequencing (NGS) gene panels, with the majority of testing proposed through pan-solid tumour large NGS panels (eg 500 genes with ability

	to detect NTRK and other fusions). In addition, the National Genomic Test Directory currently supports whole genome sequencing for all paediatric cancers and sarcomas. On the current Natrional Genomic Test Directory, NTRK fusions are currently covered for the following indications: • *ETV6-NTRK3* (by FISH or RT-PCR) for secretory carcinoma, inflammatory myofibroblastic tumour, spindle cell soft tissue tumour, infantile fibrosarcoma • *ETV-NTRK3* congenital mesonephric blastoma Therefore, whilst there is no testing currently supported for common cancers (eg colorectal carcinoma), it is envisaged that this will be incorporated within the National Genomic Test Directory within the next 6-12 months for all solid tumours.
• What impact would the technology have on the current pathway of care? How would this differ between different tumour types, please refer to questions 23 and 24 below.	The main impact on current pathways of care relate to both the changes in molecular testing for solid tumours required for detection of NTRK fusions (as outlined above) and determination of where in the treatment pathways would patients be treated with entrectinib. In the combined analysis of the entrectinib phase I trials (STARTRK-1 and ALKA-372-001, Drilon et al, Cancer Discovery 2017), 83% of patients had received more than 3 lines of prior systemic therapy. However, from the limited data available there does not appear to be any relationship between response to entrectinib and line of therapy. Increasingly, oncologists are recognising that earlier use of molecularly therapy targeting 'trunk' alterations for patients with advanced solid tumours (even when other 'standard' lines of therapy are available) is potentially associated with improvements in clinical outcome. This could be due to patients having better performance status with potential improvement in drug tolerance allowing higher dose intensity, and lower chance of developing resistant subclones.

11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	As described above, there is currently no testing for NTRK fusions for solid tumours in the National Genomic Test Directory (aside from the rare cancers listed). Therefore, routine NTRK screening for advanced solid tumours is required to be in place across the seven GLHs to identify the cohort of patients suitable for treatment with entrectinib.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Entrectinib should only be used in secondary/tertiary care settings by oncologists specifically trained and accredited in use of systemic anticancer therapy.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	For entrectinib to be introduced, routine screening of NTRK fusions is required for patients with advanced solid tumours. Broadly, current recommendations from both Europe (ESMO guidelines: Marchio et al, Ann Oncol 2019) and the US (Penault-Llorca F et al, J Clin Pathol 2019), are that in tumours with high frequency of NTRK fusions (eg infantile fibrosarcoma, secretory breast carcinoma) any technique could work in principle, however the best options as confirmatory techniques are FISH, RT-PCR or RNA-based targeted panels. Alternatively, for tumours with lower incidence of NTRK fusions, a “two-step” approach could be considered, which includes immunohistochemistry (IHC) first and confirmation of any positivity detected with IHC by subsequent NGS panels. Therefore, investment is likely to be required both for pathologists to perform IHC for TRK proteins and for NGS panels to incorporate NTRK fusion testing. The latter is in part being addressed by NHS England and

	NHS Improvement through the National Genomic Test Directory and delivery through the GLHs, although there is currently some variation in capability of NGS large gene panels across the GLHs.
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Treatment with entrectinib across three phase 1/2 clinical trials (STARTRK-1, ALKA-372-001 and STARTRK-2) was associated with an objective response rate of 57.4% [95% CI, 43.2–70.8] by blinded independent central review in 54 patients with cancers with NTRK gene fusions after a median follow-up of 15.5 months (Demetri GD et al, Annals Oncol 2018). Median progression-free survival was 11.2 months (95% CI, 8.0–14.9 months) and median overall survival was 20.9 months (95% CI, 14.9–not estimable). In patients with CNS metastases (n = 12), the overall response rate was 50.0% and the median progression-free survival was 14.3 months (95% CI 5.1 months–not estimable). Based on this data, entrectinib is expected to provide clinically significant benefits over current standard of care therapies.
• Do you expect the technology to increase length of life more than current care?	Yes. There have been no randomised trials of entrectinib in NTRK fusion positive solid cancers, due to both the rarity of NTRK fusions in common cancers and the very rare malignancies associated with high frequency of NTRK fusions. However, based on the survival data from the single arm phase 1/2 clinical trials described above, I would anticipate significant survival benefit over standard of care therapies for all malignancies associated with NTRK fusions.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes. There have been no quality of life data from the phase 1/2 trials of entrectinib. From the phase 1/2 trials, entrectinib is generally well tolerated with predominately Grade 1 or 2 adverse events that were reversible with dose modification. Dose reduction occurred in 15% of patients in the phase I studies

	(STARTRK-1 and ALKA-372-001). The most common treatment related adverse events of any grade were fatigue/asthenia (46%), dysgeusia (42%), paraesthesias (29%), nausea (28%) and myalgias (23%). Based on the high response rates and overall acceptable tolerability, I would anticipate entrectinib would be associated with improvements in health related quality of life over current standard of care therapies.
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Entrectinib will only be effective for patients with NTRK fusion positive solid cancers, as previously defined in section 10.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors	Entrectinib is given as a once daily oral administration. As defined above, entrectinib is generally well tolerated and oncologists should be familiar with management of toxicities associated with the drug. There are no other additional clinical requirements or additional safety or toxicity monitoring required. From a patient perspective I do not anticipate any specific issues affecting ease of use or acceptability.

affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	All patients receiving entrectinib are required to have advanced solid tumours with evidence of NTRK fusions using the technologies described above. Other clinical factors of relevance required to start entrectinib will be adequate performance status (PS 0-2) and organ function. If the patient has brain metastases then these must be asymptomatic and stable. Discontinuation of entrectinib will be due to clinical or radiological disease progression or intolerance despite dose reduction.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No, any health related benefits related to entrectinib will be included in the QALY calculation.
17. Do you consider the technology to be innovative in its potential to make a significant and substantial	Yes, I consider entrectinib to be highly innovative with the potential to make a substantial impact on patient outcomes based on data showing durable and robust responses in NTRK fusion positive solid tumours.

impact on health-related benefits and how might it improve the way that current need is met?	Entrectinib is one of the very first 'tumour agnostic' drugs with activity based on a specific genomic aberration rather than a tumour histology-specific subtype.
• Is the technology a 'step-change' in the management of the condition?	Entrectinib does represent a 'step-change' in the management of NTRK fusion positive solid tumours, given that there have previously been no drugs available for management of this indication. As discussed above, entrectinib also represents a broader more fundamental step-change in our approach to precision medicine in cancer, being one of the very first 'tumour agnostic' drugs.
• Does the use of the technology address any particular unmet need of the patient population?	Entrectinib has demonstrated equivalent anti-tumour activity for patients with CNS metastases, which are usually associated with worse prognosis and limited benefit from existing systemic anticancer therapies.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	From the phase 1/2 trials, entrectinib is generally well tolerated with predominately Grade 1 or 2 adverse events that were reversible with dose modification. Data from 355 patients treated across three phase 1/2 trials the majority of treatment-related adverse events (AEs) were grade 1–2 and managed with dose reduction (27.3%); discontinuation rate due to treatment-related AEs was 3.9%. Impact of toxicities is therefore anticipated to be moderate.
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	Yes. The three phase 1/2 clinical trials of entrectinib (STARTRK-1, STARTRK-2 and ALKA-372-001) were conducted in 15 countries including the UK. The distribution of patients and underlying histological tumour types covered are representative of current UK clinical practice.
• If not, how could the results be extrapolated to the UK setting?	N/A
• What, in your view, are the most important outcomes, and were they measured in the trials?	Outcomes measured in these trials of clinical relevance are response rate and duration of response (the primary outcome measures); in addition, progression free survival, overall survival in patients with and without CNS disease and safety (secondary outcomes).
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Longer term safety data on entrectinib has not yet been published. On target toxicity associated with NTRK inhibition and reported in the clinical trials include paraesthesias, weight gain, cognitive disturbance and dizziness; however, the frequency of these moderate-severe adverse events is low.
20. Are you aware of any relevant evidence that might	No

not be found by a systematic review of the trial evidence?	
21. How do data on real-world experience compare with the trial data?	I am not aware of any 'real-world' data of entrectinib in patients with NTRK fusion positive solid tumours, therefore comparison with trial data is not yet possible.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	The main equality issue relates to equity of access to NTRK testing within the 7 GLHs, as described in detail in the sections above.
22b. Consider whether these issues are different from issues with current care and why.	
Topic-specific questions	
23. Is it appropriate to consider entrectinib as a first-line treatment option for the	MASC: entrectinib is an appropriate first line therapy option, due to very limited standard therapy options available for this disease

following locally advanced or metastatic solid tumours: MASC, soft-tissue sarcoma, pancreatic cancer, cholangiocarcinoma, gynaecological cancer? Please consider this question for people ineligible for curative surgery or radiotherapy with no immunotherapy or targeted therapy options.	Soft-tissue sarcoma: this is a very diverse and heterogenous group of neoplasms. Although some soft tissue sarcomas are associated with resistance to standard of care cytotoxics (in which case entrectinib could potentially be considered as 1st line), other soft tissue sarcomas are more chemosensitive and second line use of entrectinib is likely to be preferable. Pancreatic cancer: given the very poor prognosis associated with standard of care cytotoxics, entrectinib could be considered as an appropriate first-line therapy. Cholangiocarcinoma: similar to pancreatic cancer, given the poor prognosis associated with standard of care cytotoxic therapy, entrectinib could be considered as an appropriate first line therapy. Gynaecological cancer: similar to soft tissue sarcomas, this is a wide-ranging group of malignancies with differing responses to standard of care therapies. Under this 'catch-all' term, entrectinib is unlikely to be an appropriate treatment option for all patients with gynaecological cancer in the first line setting and would be more appropriate as second line therapy or beyond.
24. Is it appropriate to consider entrectinib as a treatment option for locally advanced or metastatic solid tumours at second-line or beyond for the following tumour types:	NSCLC: entrectinib is appropriate to be considered in the 2nd line or beyond setting (after 1st line platinum chemotherapy or immunotherapy) Breast Cancer: in patients with triple negative breast cancer, entrectinib is appropriate to be considered in the 2nd line setting due to its poorer prognosis. However, for patients with hormone receptor positive breast

NSCLC, breast, thyroid cancer, colorectal cancer, neuroendocrine tumours?	cancer, it may be more appropriate to consider entrectinib after failure of hormone therapy in addition to at least one line of palliative chemotherapy. Colorectal cancer: entrectinib is appropriate to be considered in the 2nd line or beyond setting Neuroendocrine tumours: this is a diverse group of malignancies with therapy based on histological grade. However, given the relative lack of treatments and poor prognosis associated with many neuroendocrine tumours, entrectinib is appropriate to be considered in the 2nd line or beyond setting.
25. Is it appropriate to consider entrectinib as a first-line treatment option for locally advanced or metastatic solid tumours for the following tumour types if targeted or immunotherapies are not appropriate: NSCLC, breast, thyroid cancer, colorectal cancer, neuroendocrine tumours?	In all of these cancer types, given the minimal data associated with entrectinib in these tumour types, it would be appropriate to consider entrectinib only after 1 line of standard of care systemic therapy.

26. What other locally advanced or metastatic solid tumour types have NTRK fusions? At what point in the treatment pathway would it be appropriate to consider entrectinib as a treatment option for each of these solid tumour types?	Broadly considering other cancers with low frequencies (renal cell cancer, melanoma, gastrointestinal stromal tumours, head and neck cancers, high grade gliomas), it would be appropriate to consider entrectinib as second line therapy (ie after 1 line of standard of care systemic therapy).
Key messages	

27. In up to 5 bullet points, please summarise the key messages of your statement.

- Entrectinib demonstrates clinically significant anti-tumour activity across a range of solid tumours with NTRK fusions, including high response rates in patients with CNS metastases
- Requirement for NTRK screening – likely combination of IHC and NGS delivered through National Genomic Test Directory
- Entrectinib appears to be well tolerated overall but longer term safety data (especially given duration of responses) is required
- One of the very first 'tumour agnostic' cancer drugs directed against specific genomic aberrations, likely to represent a new wave in precision medicine for cancer
- Uncertainties remain around which line of therapy entrectinib should be used in; although this is likely to vary depending on tumour type, broad current consensus suggests tumour agnostic therapies such as entrectinib should be administered as earlier lines of therapy.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

NICE National Institute for Health and Care Excellence

National Institute for Health and Care Excellence

Cancer Drugs Fund Clinical Lead statement

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

Background

Tumour agnostic drugs

1. NTRK inhibitors are the first tumour agnostic drugs which are expected to be licensed in Europe but others are likely to follow in the next few years. There is evidence of benefit for anti PD-L1 immunotherapy in cancer patients whose tumours exhibit microsatellite instability-high or mismatch repair deficiency or high tumour mutational burden. There are clinical trials in other drugs targeting NTRK gene fusion cancers and also resistance to 1st generation NTRK inhibitors. A number of basket clinical trials are running in cancer patients with other mutations or gene fusions (e.g. RET, FAP etc).

Incidence of NTRK gene fusions

2. There is an emerging evidence base as to the incidence of NTRK gene fusions. Some very rare cancers have high (80-100%) proportions with NTRK gene fusions (e.g. the mammary analogue secretory variant of salivary gland cancer, the secretory variant of breast cancer, paediatric mesoblastic nephroma, infantile fibrosarcoma). Some rare cancers have modest (20-40%) proportions of NTRK gene fusions (e.g. paediatric non-brain stem glioblastoma, spitzoid melanoma) or low (2-12%) incidences (e.g. papillary thyroid cancer, some brain malignancies, cholangiocarcinoma, gastrointestinal stromal tumours). Most cancers and all the commoner

cancers have very low proportions of NTRK gene fusions of 1% or less.

3. NHS England and NHS Improvement notes that in the entrectinib submission from Roche, evidence is presented which shows NTRK gene fusion to be evident in about 0.5% of various surveys of unselected patients and in [redacted] of the patients screened for the entrectinib clinical studies. The latter [redacted] figure could be biased by the inclusion of those patients with rare cancers in whom NTRK gene fusion is much more common.

4. NHS England and NHS Improvement concludes that on the current evidence and when all solid tumours are considered, it is reasonable to assume an incident proportion of between 0.5 and 1% with NTRK gene fusions. NHS England and NHS Improvement therefore considers that a base case figure of 0.5% should be used in this appraisal and a scenario analysis be done at a 1% incidence.

Natural history of cancers with NTRK gene fusions

5. Little is known as to the natural history of NTRK gene fusion positive varieties of solid tumours. Roche in its submission presents some preliminary evidence which, for example, suggests that the outlook for metastatic colorectal cancer patients with NTRK/ROS1/ALK genetic changes (n=27) is worse than those without such changes (n=319). However, this is not a pure NTRK gene fusion group and the incidence of NTRK gene fusion in colorectal cancer is thought to be <1%. The contribution of the ALK and ROS1 patients to this adverse outcome could explain much of this apparent difference. NHS England and NHS Improvement recognises that there may be a difference in outlook for incurable patients with metastatic cancer who have NTRK gene fusions but there is no robust evidence to support this at present.

Draft marketing authorisation

6. [redacted] Similar wording was used in the main phase II study of entrectinib which delivered 51 of the 54 analysed patients. The definitions of ‘standard prior therapies’ and ‘no acceptable standard therapies’ are very important (these issues are explored in detail in the following paragraphs in terms of individual tumours). The phrase ‘no acceptable standard therapies’ is particularly open to potentially variable interpretation.

Generalisability of the trial population as regards clinical benefit

7. NHS England and NHS Improvement notes that 37% of the 54 patients were treatment naïve for chemotherapy and the median time since cancer diagnosis was 21 months. NHS England and NHS Improvement therefore considers that there is a potentially considerable bias in these early entrectinib studies. This is as a result of the inclusion of patients who knew they had a NTRK fusion cancer and wished to have the opportunity of receiving entrectinib whilst the trial was open and they were eligible for treatment. It may be therefore that ‘standard therapies’ had not been fully explored. In addition, the entrectinib studies have patients which are biased in terms of rare cancers figuring significantly e.g. sarcomas 24%, salivary gland cancers 13%, thyroid cancers 9%. There is therefore uncertainty as to the generalisability of the entrectinib clinical data.

8. NHS England and NHS Improvement notes that 22% of the 54 entrectinib patients had cerebral metastases. Although the involvement of the central nervous system was either untreated and asymptomatic or previously treated and controlled, the presence of such metastases confers an adverse prognosis to patients having

systemic therapy. NHS England and NHS Improvement welcomes the inclusion of such patients in these clinical studies as this makes the results more generalisable to clinical practice.

Activity and toxicity of entrectinib

9. Entrectinib is clearly a very active drug in NTRK fusion positive malignancy. It cannot be directly compared with larotrectinib for response rate, progression-free survival and overall survival in view of the differing case and age mix in the respective pooled analyses e.g. in terms of proportions of tumour types treated with entrectinib versus larotrectinib, non-small cell lung cancer 19% vs 8%, breast cancer 11% vs 1%, infantile fibrosarcoma 0% vs 14%, melanoma 0% vs 8% etc. Roche reports only 5 paediatric patients treated with entrectinib in a separate study.

10. The clinical impact of entrectinib is striking but the median duration of follow-up is only 12.9 months, the number treated and evaluable is small and the numbers of patients with specific cancers are very small. Of note too is that entrectinib is clearly active in patients with metastases in the central nervous system with a similar response rate in the brain to that observed systemically in all patients in the pooled analysis. Any conclusions as to the durability of response in patients with brain metastases have to be even more guarded in view of the very small patient numbers and the short duration of follow-up.

11. Entrectinib was reasonably well tolerated with a discontinuation rate of [redacted] in its safety population (for reasons other than for progressive disease or death). Of these [redacted], a half had treatment-related discontinuations.

12. Currently, systemic therapy is organised around tumour site-specific teams as knowledge and experience of the natural history of individual cancers is very important in the optimal care of patients.

The rarity of NTRK gene fusions in most cancers means that individual oncologist experience in the use of entrectinib will be very small. Consideration will therefore have to be given within cancer centres of sharing experience of entrectinib use in order to assist in the best management of side-effects.

The treatment pathway and comparators

13. The issue of where in the treatment pathways patients would be treated with entrectinib is an important one, partly as it determines what the comparator costs should be but mainly as it resolves what the comparator durations of survival should be. This is because Roche has submitted a naïve weighted comparison of outcomes with entrectinib versus what it believes to be the correct comparator albeit in populations of patients with unknown NTRK gene fusion status. Roche considers that entrectinib would be used as 1st line systemic therapy for incurable patients with –

    - mammary analogue secretory carcinoma of the salivary gland. NHS England and NHS Improvement agrees as 1st line chemotherapy is not very effective and has many side-effects. NHS England and NHS Improvement notes the great rarity of this type of salivary gland tumour.

    - soft tissue sarcoma. NHS England and NHS Improvement disagrees with the use of enrectinib as 1st line therapy for sarcomas as a whole as there are many different types of soft tissue sarcoma and the data on the efficacy of entrectinib is limited to only 13 patients. Whilst entrectinib would be considered 1st line systemic therapy in some sarcomas which are chemo-resistant (e.g. malignant peripheral nerve sheath tumour), other types of sarcoma are more chemo-sensitive to standard and NICE-recommended therapy and thus it is more likely that promising but unproven

entrectinib would be used 2nd line in such cases. Trabectedin (as a NICE-recommended 2nd line treatment option for some types of sarcoma) and best supportive care would then be the most appropriate comparators.

- pancreatic cancer and cholangiocarcinoma. There were only 3 and 1 patients with these 2 cancers, respectively, treated in the entrectinib pooled studies. The evidence base is therefore extremely small and NHS England and NHS Improvement is uncertain as to whether clinicians and patients would jointly opt for entrectinib as 1st line systemic therapy in these two cancers despite the poor survival outcomes associated with standard therapies. If entrectinib is used as 2nd line therapy, then best supportive care is the comparator.

- gynaecological cancer. There was only one patient treated with 'gynaecological cancer' in the entrectinib clinical studies and this term embraces a number of very different cancers. NHS England and NHS Improvement therefore considers it very unlikely that entrectinib would displace any current 1st line standard treatments. The correct comparator depends on which gynaecological cancer is meant by this term.

14. Roche in its submission considers that the following NTRK fusion cancers would receive entrectinib as 2nd or further line treatment:

- non-small cell lung cancer. NHS England and NHS Improvement considers this is reasonable after any immunotherapy and 1st line cytotoxic chemotherapy as the efficacy of docetaxel ± nintedanib is low and toxicity is substantial. The survival of such patients could easily be less than the median figure of 10.7 months used in the company submission for lung cancer patients. If Roche however

wishes to use the cost of docetaxel and nintedanib, then it has to use the survival outcomes associated with such treatments.

- breast cancer. NHS England and NHS Improvement regards that entrectinib would be used but since there were only 6 patients with breast cancer in the entrectinib studies (and it is not known how many of these had the rare secretory breast cancer variant which expresses NTRK gene fusion very highly), treatment with entrectinib would be after the failure of 2 lines of palliative chemotherapy.

- colorectal cancer. NHS England and NHS Improvement regards that entrectinib would be used but since there were only 4 patients with colorectal cancer in the entrectinib studies, treatment with entrectinib would be after the failure of 2 lines of palliative chemotherapy (oxaliplatin- and irinotecan-based treatments are both NICE-recommended therapies). Trifluridine/tipiracil is NICE recommended as 3rd line treatment but is not very effective and hence best supportive care is also an option as a comparator in colorectal cancer. Survival duration after 2 lines of palliative chemotherapy could be less than the median figure of 9.1 months used in the company submission.

- thyroid cancer. Roche correctly assumes that 1st line systemic therapy (other than radio-iodine) is with '-inib' therapy (5 patients treated with entrectinib in the clinical studies) and hence best supportive care is the comparator for entrectinib.

- neuroendocrine carcinoma. NHS England and NHS Improvement does not consider that everolimus will be displaced by entrectinib and hence the cost of everolimus

should be discarded and the best supportive care survival data after everolimus used in the comparator to entrectinib analysis.

15. NHS England and NHS Improvement has set out all this detail as a weighted cost, progression free (PFS) and overall survival (OS) analyses have been used in the comparison with entrectinib. Whilst the costs of the comparator arm should be reduced to reflect best supportive care where appropriate, so should the survival outcomes be used for best supportive care where appropriate. Since overall survival is a key determinant of the cost effectiveness of entrectinib in NTRK gene fusion whereas the cost of comparator therapies is not, NHS England and NHS Improvement considers that the company's analysis (as regards this issue of place in the treatment pathway and accompanying outcomes) may be overestimating comparator survival and thus overestimating the ICER.

Pooling of the entrectinib studies

16. NHS England and NHS Improvement supports the pooling of the 4 entrectinib studies in order to maximise the patients included in the analyses of clinical and cost effectiveness.

Cost effectiveness

Parametric extrapolation

17. Given the immaturity of the entrectinib data, NHS England and NHS Improvement recognises the need for parametric extrapolation of the data on progression-free and overall survivals. NHS England and NHS Improvement considers that the Weibull extrapolation for both progression free and overall survivals is just as clinically plausible as the company-choice of the exponential. NHS England and NHS Improvement notes that use of the Weibull would significantly increase the ICER.

Generalisability as regards costs

18. NHS England and NHS Improvement again notes that the cost effectiveness analysis for the comparator population is based on which cancers were treated in the entrectinib pooled analysis. Given the tumour agnostic marketing authorisation that is expected for entrectinib, it is highly likely that the case mix of the NHS England and NHS Improvement treated population will significantly differ from the biased case mix of the pooled entrectinib analysis e.g. the real world NHS England and NHS Improvement population will not be constituted by a 13% proportion made up by patients with the mammary analogue secretory carcinoma variant of the salivary gland. It is likely that a real-world mix of patient case mix would increase the ICER by both reducing the incremental survival and by the case mix adjustment increasing the costs of ascertaining NTRK gene fusion e.g. from a very small cost for testing such salivary gland tumours (NTRK gene fusion present in 90-100%) versus lung cancer (NTRK gene fusion present in 1% or so).

Utilities

19. NHS England and NHS Improvement notes that the mean utility values gained from the entrectinib pooled analysis for the progression free and post progression survival states of the economic model were [redacted] and [redacted]. It is counterintuitive for the progressed state utility to exceed that of the progression free utility. The corresponding weighted mean utility values for the comparator were 0.73 and 0.59. If these two entrectinib and comparator populations are comparable then they both must start with the same utility value: by keeping this differential for the progression free state, the company's estimates of QALY gain for entrectinib are biased. The company however then ditches the counter-intuitive figure of 0.84 for the post progression free state for entrectinib in favour of the 0.59 figure borrowed from the weighted comparator analysis. It is NHS England and NHS

Improvement's view that Roche must be consistent and start with equal utility values in the progression free state.

Costs of chemotherapy

20. Roche has used incorrect costs for systemic therapy (mainly chemotherapy) in the comparator population. It has used BNF costs which are not the costs borne by the NHS. The correct costs are those set out in the eMIT tool in the Commercial Medicines Unit of NHS England and NHS Improvement. The cost differences are stark: a 100mg dose of paclitaxel is costed at £200 by the company whereas the real cost is £9, a 200mg dose of doxorubicin is costed as £391 whereas the real cost is £16 etc. NHS England and NHS Improvement acknowledges that the cost of the comparator is not a key driver of cost effectiveness in the economic model but is concerned that other costs used in the model may also be as unrealistic as these chemotherapy costs.

21. In terms of drug administration costs, Roche has omitted any chemotherapy tariff costs for any oral treatment: this is important for entrectinib which has a substantial mean treatment duration. The SB11Z oral chemotherapy tariff (£120 per visit) should have been used and this incremental cost applies almost completely to the entrectinib arm.

Further chemotherapy after entrectinib and comparator treatment

22. Roche assumes that comparator patients do not receive any further active therapy and that 35% of entrectinib patients receive further active treatment. NHS England and NHS Improvement considers that both of these assumptions are reasonable provided that the correct treatments (last line of chemotherapy or best supportive care) have been chosen for the individual cancers in the comparator arm.

NTRK gene fusion testing

23. Proof of a NTRK gene fusion requires either whole genome sequencing (WGS) or next generation sequencing (NGS), the latter providing the technology for multigene panels (which provide testing for anything between 5 and 500 genes). There are some screening TRK immunohistochemistry tests which greatly reduce the need for NGS but these also have a significant false negative rate.

24. As part of the establishment of the NHS Genomic Medicine Service (including the Genomic Laboratory Hubs), NHS England and NHS Improvement are making fundamental changes to how cancer genomic testing is provided, commissioned and funded. A national service has been created and is regionally organised by 7 Genomic Laboratory Hubs. The hubs are responsible for processing samples for WGS, performing NGS testing and interpreting all NGS and WGS results before returning the results to the requesting clinician. The WGS is done by Genomics England which receives samples from and returns WGS results to the hubs. 2019-20 is a critical set up year for the Genomic Laboratory Hubs for both their establishment and for the diversion of previous genomic funding from the many hospitals who have done a variety of gene testing until now. The NHS England and NHS Improvement Genomics Medicine Service is the first national service to be set up in the world: its ambition is matched by the revolution occurring in the organisation and funding of the 7 Genomic Laboratory Hubs and in the types of NGS now becoming available.

25. During 2019, the NHS will start to offer whole genome sequencing for patients with paediatric cancer and for those with all types of sarcoma. The current timeline for the operation of WGS is end of the summer of 2019, however full implementation will take time (NHS England and NHS Improvement's working assumption is that it will be the autumn of 2020 before all WGS pathways are fully operational across the country). Full implementation requires significant changes to the diagnostic pathway including the establishment of pathways of

care such that fresh frozen tissue can be processed by the Genomic Laboratory Hubs in a timely fashion so that DNA of the appropriate quality is obtained before then being sent to Genomics England for testing. Funding from NHS England and NHS Improvement is in place for the provision of WGS for paediatric cancer and sarcoma although it is recognised that NGS may be necessary for NTRK fusion testing in the short term until WGS is fully operational.

26. For some rare tumours, such as mammary analogue secretory carcinoma of the salivary gland and the secretory variant of breast cancer, the National Genomic Test Directory for 2019 already sets the expectation that NTRK testing should be performed. Although NHS England and NHS Improvement does not have robust data about existing testing activity, it is aware that cancer genomic testing for such a test is not currently performed systematically across the country. Funding is therefore in place for the NTRK gene fusion testing for these 2 rare cancers.

27. In all other adult solid cancers, NTRK gene fusion testing is not currently required by the National Genomic Test Directory and is not systematically performed. However, by the end of the 2019/20 financial year, the Genomic Laboratory Hubs plan to introduce gene panels for solid tumour testing, which will include the capability to identify NTRK gene fusions. This could be for example with a 50-60 gene panel (cost ~£250) or a 500 gene panel (cost ~£400). To facilitate testing for NTRK gene fusion in solid tumours, NHS England and NHS Improvement will need to include NTRK gene fusion testing in the National Genomic Test Directory and determine the funding required. Some of the Genomic Laboratory Hubs are currently more advanced in their ability to deliver NGS multigene panel testing and hence there is likely to be some initial sharing of NGS testing until all 7 of the hubs are fully operational.

28. As is clear from the preceding paragraphs and apart from the rare cancers in which NTRK gene fusions are more commonly expressed, large numbers of patients have to be screened to find the NTRK gene fusion. For a tumour agnostic drug which has a high chance of benefitting patients who harbour the NTRK gene fusion, the logical potentially eligible population is in all patients with solid cancers which are incurable (i.e. the patients who have locally advanced or metastatic disease). Some cancers already have some genetic testing embedded in the treatment pathway (e.g. melanoma and lung, colon, thyroid, breast and ovarian cancers). For patients and clinicians to be able to best use the information of NGS panel testing, such testing has to be done prior to the initiation of all systemic therapy for the locally advanced/metastatic disease. The cost of NGS panel testing is therefore very great as NHS England and NHS Improvement estimates that it would need to test approximately 100,000 patients in all. About 3,000 will be eligible for WGS and 30,000 already receive some genomic testing as part of existing standard of care (and this is assumed to cover the cost of NGS panel testing at least in melanoma and lung and colorectal cancers). Thus 67,000 patients represent additional and new activity. The estimated assay cost of this new activity would be £16.8m if the 50-60 gene panel is used, £26.8m if the 500 gene panel is used and £21.8m if an average cost of £325 per multigene test is used. If the £325 figure is used and since 33% of the total testing cohort is assumed to already receive testing, this means that the average incremental diagnostic cost per patient tested is £218. If the incidence of NTRK gene fusion is 1 in 200, then the total cost per positive NTRK gene fusion patient is £43,500. If the incidence of NTRK gene fusion is 1 in 100, then the total cost per positive NTRK gene fusion patient is £21,800. If WGS initially does not deliver information within a timetable required for clinical decision making and all 3000 patients have to have NGS, then the average incremental diagnostic cost per patient tested would initially be £227

with a cost per positive NTRK fusion of £45,000 if its incidence is 1:200 and £21,500 if 1:100.

29. In addition to the costs of WGS and gene panel testing, there are capital costs to consider: laboratory equipment, bioinformatics and the increased need for expert interpretation of results to aid clinical decision-making. NHS England and NHS Improvement is currently working through these issues as the 7 Genomic Laboratory Hubs are starting from different baseline positions.

30. In summary, it is anticipated that WGS will be fully operational by Q2 2020/21 and panel testing will be available by Q1 2020/21. Uptake of molecular testing across the 7 genomic hubs will increase during 2020/21 as genomic pathways are embedded and links are made with the clinical teams. Given the complexity of implementation, it may take a further 12 months for molecular testing to become fully embedded in practice.

Costing of NTRK gene fusion testing for this appraisal

31. The established approach in NICE technology appraisals of cancer drugs which require genomic testing has been to ensure that the full cost of the testing has been included in the cost effectiveness analysis. In the appraisal of entrectinib, there are 4 important differences to previous appraisals of targeted cancer drugs which have required genomic testing. Firstly, entrectinib is a tumour agnostic drug and hence all cancers have to be tested as there is currently no evidence to indicate that certain cancers never have NTRK gene fusions. The consequence of this is that the number of patients to be tested is very great. Secondly, the incidence of NTRK gene fusions in solid tumours is very low. Thirdly, the need for NTRK fusion testing is coming at a critical set up time for a new and national genomic medicine service in England. Fourthly, this new service must embed at set-up the technologies which will it will need to provide the huge

benefits of such a national service i.e. a service for WGS and NGS panel tests has to be built.

32. NHS England and NHS Improvement recognises that the national availability of WGS and NGS multi-gene panel testing for patients with incurable solid tumours will bring many future treatment opportunities: for NTRK fusion inhibitors, for other tumour agnostic cancer drugs (several of which are likely in the next few years), for the many expected future targeted drugs which will require genomic testing in patients with specific tumours and for greater entry into clinical trials. It should be noted that the current plans for NHS investment in these genomic services is primarily for improving geographical equity of access and pump priming the new genomics infrastructure. NHS England and NHS Improvement therefore considers that it is appropriate that at least part of the cost for multi-gene panel testing be covered by each company that benefits from this new service provision (in line with the standard approach employed in technology appraisals). As a consequence, NHS England and NHS Improvement would wish NICE to explore scenario analyses in its appraisal of the cost effectiveness of entrectinib in which various percentages of the costs of multi-gene panel testing are borne by entrectinib: 100%, 50%, 33%, 25% and 0%.

33. To reach a proportionate and reasonable position on how much of this cost should be borne by the NHS vs an individual company with a tumour agnostic product, NHS England and NHS Improvement will wish to see these scenario analyses and will decide on the appropriate level of contribution by October 2019, i.e. in advance of the final point of submission before the NICE committee considers entrectinib in its November 2019 meeting.

Roche costing of detecting NTRK gene fusions

34. The company in its submission understands the great changes currently occurring in terms of the setting up of a NHS England and NHS Improvement Genomics Medicine Service. It proposes a 2 stage approach to the majority of patients in whom WGS is not currently funded: a screening immunohistocytochemical assay for pan-TRK (the Ventana assay test) and then NGS on the 10% of patients whose Ventana test is positive. NHS England and NHS Improvement notes that Ventana is a Roche company and that there is a significant false negative rate to its TRK test. NHS England and NHS Improvement is critical of this screening approach for three reasons. Firstly, histological services in England are currently stretched in terms of delivering capacity and this would add a very substantial workload to pathology departments. By the time the whole process of set up and training had occurred and TRK immunocytohistochemical testing was up and running, it would be time to dismantle the service on account of NGS being available. Secondly, the Roche approach requires NGS to be in place in any case. Thirdly, modern oncology medicine will be founded on genomic testing and it would be a retrograde step now to pour effort into in effect 20th century technology when it is 21st century genomics that NHS England and NHS Improvement wishes to promote and deliver to the benefit of patients.

35. Roche has produced an analysis of what it would cost by disease to find 1 patient with a NTRK gene fusion and then weighted this according to the type of cancer and the proportion of patients with this cancer in the pooled entrectinib clinical studies. When the number of patients is so small (n=54), the weighted average could easily change with a different case mix and hence this weighted average of cost of testing carries very significant uncertainty (this issue has been described in greater detail in preceding paragraphs).

End of life cost effectiveness threshold

36. NHS England and NHS Improvement agrees that entrectinib would satisfy NICE's End of Life threshold criteria given that using a weighted average of survival for the comparator is a reasonable approach and NHS England and NHS Improvement believes that survival may be overestimated as Roche has assumed use of entrectinib at earlier points in the treatment pathway in some diseases (see above for detailed discussion of this).

Cancer Drugs Fund

NHS England and NHS Improvement supports Roche's aim for entrectinib to enter the Cancer Drugs Fund. NHS England and NHS Improvement regards entrectinib as a highly promising drug which needs clinical data of much greater maturity and testing in a real world setting across many cancers and in much greater numbers. NHS England and NHS Improvement is concerned that in Roche's own economic analysis, Roche has [redacted]: with the full QALY weighting of the End of Life threshold (i.e. at £50,000 per QALY), Roche's base case deterministic ICER is £54,600 at the discounted entrectinib price. [redacted] These are of course the issues on which the Appraisal Committee will be deliberating and making its own conclusions.

Implementing a positive NICE recommendation

NICE recognises that in the event of a positive recommendation, more prescriptive clinical commissioning criteria for treatments commissioned via

Specialised Services will be implemented by NHS England and NHS Improvement to ensure appropriate use within the NHS.

NHS England and NHS Improvement is responsible for ensuring that the final clinical commissioning criteria are aligned with final guidance (section 1 – recommendation and section 3 – committee discussion).

Draft commissioning criteria

37. If entrecinib for treating NTRK gene fusion locally advanced/metastatic solid tumours is recommended for use within its marketing authorisation, NHS England and NHS Improvement proposes to use the following commissioning criteria:

- The patient's cancer must have the presence of an NTRK gene fusion as determined by WGS or following a NHS multigene panel test
- The patient must have locally advanced or metastatic disease
- The patient must have progressed following treatment with all NICE-recommended systemic therapies or established standard therapies in clinical practice or have a documented ineligibility for such treatments
- The patient must have an ECOG performance score of 0-2
- If the patient has metastases in the central nervous system, then these must be asymptomatic if untreated or treated and controlled
- Entrectinib is to be used as monotherapy
- The prescription of entrectinib and care of the patient on entrectinib to be by a consultant oncologist specifically trained and accredited in the use of systemic anticancer therapy
- The patient is to be treated until progressive disease or unacceptable toxicity or the patient choice to discontinue treatment, whichever is the sooner.

If this technology is recommended for routine commissioning in a subpopulation or with certain specifications (for example, a treatment continuation rule), the final commissioning criteria will reflect these conditions.

38. If entrectinib for treating NTRK gene fusion positive locally advanced or metastatic cancer is recommended for use in the Cancer Drugs Fund, the final commissioning criteria will reflect the patient eligibility criteria in the managed access agreement.

Issues for discussion

39. These have all been outlined above.

Issues for decision

40. These relate to the above and principally relate to:

- Incidence of NTRK gene fusion cancers
- Interpretation of the wording of the marketing authorisation
- Generalisability of the trial population
- Treatment pathway and comparators
- Parametric modelling of progression free and overall survivals
- Utilities in the progression free and post progression health states
- Costs of chemotherapy
- NTRK gene fusion testing, implementation and costs
- CDF entry

Equality

41. NHS England and NHS Improvement recognises that the 7 Genomic Laboratory Hubs are at different stages of being able to implement NGS multigene panel testing and this variation will be resolved over the next 1-2 years. NHS also recognises that WGS will take time to embed within clinical treatment pathways, particularly in respect of the need for the collection and processing of fresh tissue.

Author

Professor Peter Clark, NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund

July 2019

CONFIDENTIAL UNTIL PUBLISHED
Evidence Review Group's Report

Entrectinib for treating NTRK fusion-positive solid tumours

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD
Authors	Robert Hodgson, Research Fellow, CRD
	Alexis Llewellyn, Research Fellow, CRD
	Peter Murphy, Research Fellow, CRD
	Lindsay Claxton, Research Fellow, CRD
	Lucy Beresford, NIHR Training Fellow, CRD
	Matthew Walton, Research Fellow, CRD
	David Glynn, Research Fellow, CHE
	Kath Wright, Information Service Manager
	Sofia Dias, Professor of Health Technology Assessment
Correspondence to	Sofia Dias, Professor of Health Technology Assessment, CRD, University of York, York YO10 5DD
Date completed	Date completed (29/07/2019)

Source of funding

This report was commissioned by the NIHR HTA Programme as project number NIHR128763

Declared competing interests of the authors

None

Acknowledgements

We thank Dr Robert Metcalf, Honorary Consultant in Medical Oncology at The Christie NHS Foundation Trust, Professor Christian Ottensmeier, Chair in Experimental Cancer Medicine at University Hospital Southampton, Dr Helene Schlecht, Principal Clinical Scientist at Manchester University NHS Foundation Trust, and Marta Pereira, Clinical Scientist at Manchester University NHS Foundation Trust for their expert advice throughout this project. We also thank Professor Steve Palmer of the Centre for Health Economics, for his advice on the project.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Hodgson R, Llewellyn A, Murphy P, Claxton L, Beresford L, Walton M, Glynn D, Wright K, Dias S. Entrectinib for treating *NTRK* fusion-positive solid tumours: A Single Technology Appraisal. CRD and CHE, University of York, Technology Assessment Group, 2019.

Contributions of authors

Robert Hodgson, Lindsay Claxton, Peter Murphy, David Glynn and Lucy Beresford wrote the cost effectiveness sections and conducted the ERG economic analyses. Alexis Llewellyn, Matthew Walton and Lucy Beresford wrote the clinical effectiveness sections of the report. Kath Wright wrote the sections on the search strategies. Robert Hodgson provided advice, commented on drafts of the report and took overall responsibility for the cost effectiveness sections. Sofia Dias provided advice, commented on drafts of the report and took overall responsibility for the report.

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in [redacted], all academic-in-confidence (AIC) data are highlighted in [redacted]

Copyright statement

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

ADR	Adverse events reactions
AE	Adverse events
ALK	Anaplastic lymphoma kinase
BHM	Bayesian hierarchical model
BSC	Best supportive care
CCOD	Clinical data cut-off date
CDF	Cancer Drugs Fund
CEA	Cost-effectiveness analysis
CI	Confidence interval
CNS	Central nervous system
CR	Complete response
CRC	Colorectal cancer
CrI	Credible interval
CS	Company submission
CSR	Clinical study report
DNA	Deoxyribonucleic Acid
DoR	Duration of response
DSA	Deterministic sensitivity analysis
EEA	Efficacy evaluable analysis dataset
EMA	European Medicines Agency
eMIT	Electronic market information tool
ERG	Evidence review group
ESMO	European Society for Medical Oncology
FISH	Fluorescence in situ hybridisation
HRQoL	Health related quality of Life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
IHC	Immunohistochemistry
IPD	Individual participant data
ITT	Intention to treat
IV	Intravenous
KM	Kaplan Meier
LYG	Life years
MAIC	Matched adjusted indirect comparison
MASC	Mammary-analogue secretory cancer
NGS	Next generation sequencing

NICE	National Institute for Health and Care Excellence
NNS	Number needed to screen
NSCLC	Non-small cell lung cancer
NTRK	Neurotrophic tyrosine receptor kinase
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PFS	Progression free survival
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSM	Partitioned survival model
QALY	Quality adjusted life-year
RCT	Randomised Controlled Trial
RNA	Ribonucleic acid
ROS1	Proto-oncogene tyrosine-protein kinase ROS
RT-PCR	Real-time polymerase chain reaction
SLD	Sum of longest tumour diameter
SLR	Systematic literature review
SOC	Standard of care
Trk	Tropomyosin receptor kinase
TTO	Time trade off
WGS	Whole-genome sequencing
WTP	Willingness-to-pay

1 Summary

1.1 Critique of the decision problem in the company's submission

Entrectinib is a potent inhibitor of tropomyosin receptor kinases A, B, and C, encoded by the neurotrophic tyrosine receptor kinase genes *NTRK1, NTRK2,* and *NTRK3*, anaplastic lymphoma kinase (*ALK*) and *ROS* proto-oncogene 1 receptor tyrosine kinase (*ROS1*). The recommended dose for entrectinib is 600 mg orally once daily for adults, and 300 mg/m^2 orally, once daily for paediatric patients who have the ability to swallow whole capsules. Entrectinib is currently awaiting European marketing authorisation.

The NICE scope reflects the anticipated licence, which presents entrectinib as a treatment option for [redacted]

The ERG found that the intervention and outcomes presented in the company submission (CS) evidence match the NICE scope. The comparators selected by the manufacturer were all therapeutic options offered in established management without entrectinib, as defined in the NICE scope. The ERG is concerned that the population presented in the evidence submitted does not match the NICE final scope. Only a small subset of tumour types known to harbour *NTRK1*/2/3 fusions were represented in the CS and only one *NTRK2* patient was included. A significant proportion ([redacted]) of trial patients received entrectinib as first line systemic therapy, including for several tumour types where the company placed entrectinib in subsequent lines of therapy. The high proportion of patients receiving entrectinib in earlier lines of therapy across tumour types may mean that survival benefits are overestimated.

1.2 Summary of clinical effectiveness evidence submitted by the company

The efficacy evidence in the CS was supported by four uncontrolled basket trials that included a total of 66 efficacy evaluable patients with metastatic or locally advanced *NTRK* fusion-positive solid tumours, including seven paediatric patients. Most of the efficacy evidence came from an *NTRK* positive subgroup of an uncontrolled phase 2 basket trial. Clinical efficacy for ten tumour types across 54 patients were included in the company's submission: sarcoma, non-small cell lung cancer (NSCLC), Mammary analogue secretory carcinoma (MASC), breast, thyroid, colorectal cancer (CRC), neuroendocrine tumours, pancreatic cancer, gynaecological cancers and cholangiocarcinoma. Following an ERG request, response data for [redacted] further patients across [redacted] were provided. Each tumour type was represented by between one and 13 patients in the whole *NTRK* population.

At the latest clinical data cut-off date (CCOD) provided [redacted], the objective response rate (ORR) was [redacted]; complete response was reported in [redacted], and partial response in [redacted]. Median duration of response was [redacted] in responders [redacted] and the Kaplan-Meier (KM) estimated median progression free survival (PFS) was [redacted]. At CCOD [redacted], [redacted] had died and the Kaplan-Meier estimated median overall survival (OS) was[redacted]. Following a request from the ERG, the company provided responder analyses as well as individual patient-level

response data for 66 *NTRK* positive patients by tumour type and line of therapy, but not for PFS and OS. The company's Kaplan-Meier curves from responder analyses showed that the OS benefit observed in responders ceased approximately at [redacted], at which point the two survival curves cross.

Health-related quality of life outcomes were reported. The safety population included 355 patients across four trials, of which 68 had an *NTRK* fusion. [redacted]. AEs leading to discontinuation of entrectinib were reported in [redacted] of the safety population.

In the absence of a control group in the trial evidence, the company adopted a pragmatic approach to identify PFS and OS comparator data for established management without entrectinib, by searching NICE pathways to identify NICE approved comparators for each of the tumour types represented in the CS efficacy evidence. Median PFS and OS from each tumour type were averaged and then pooled to calculate mean overall PFS and OS across all tumour types, weighted by the prevalence of each tumour type within the trial population.

The ERG found that the population included in the comparator trials is unlikely to match the entrectinib efficacy population, notably due to the unknown prevalence of *NTRK* fusions in most of the comparator evidence, and mismatches in the lines of therapy previously received with the treatment pathway in practice. In the base case analysis no attempts were made to adjust for differences in population characteristics between the entrectinib and comparator trial populations; comparisons were naïve and did not account for any potentially important prognostic factors. The ERG found that the methods used to identify, select and combine comparator data are inappropriate, and that the comparator data used to inform the company model is highly unreliable.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

Overall, the trial evidence showed a clinically meaningful overall response rate across tumour types. However, there is considerable uncertainty regarding the extent to which the response observed translates into clinically meaningful survival benefits. The ERG identified a number of important issues, particularly due to the significant immaturity of the PFS and OS data. Despite substantial censoring and the small number of patients at risk in the tails of the Kaplan-Meier curve, the crossing of OS curves between entrectinib responders and non-responders is of some concern.

The ERG were concerned that the large number of tumour types not represented in the trial, the previously discussed issues concerning trial power, and the naïve comparisons with somewhat arbitrary comparator data meant that the evidence submitted in the CS may not have allowed the company to meaningfully address the decision problem.

The ERG explored heterogeneity in response rates between the 13 tumour types included in the EEA dataset using a Bayesian hierarchical model, which assumes the response probabilities are similar

across tumour types, rather than identical (the company's preferred assumption). The ERG's analyses found that overall response rates obtained were similar to those observed when equal response probabilities are assumed, although there was considerable uncertainty in the level of heterogeneity of response rates across tumour types. Based upon this analysis, the response probability for an unrepresented tumour type could range from [redacted] Therefore, the possibility that some tumour types could have response rates that differ significantly from the pooled estimate of [redacted] cannot be excluded.

1.4 Summary of cost effectiveness submitted evidence by the company

The company's economic submission included a systematic review of published evidence on the cost-effectiveness, health-related quality of life, and resource use associated with entrectinib in the treatment of patients with *NTRK* fusion–positive solid tumours. No studies were, however, found to meet the review inclusion criteria and as such, no published evidence was identified on the cost-effectiveness, health-related quality of life, and resource use associated with entrectinib.

The CS presented a *de novo* cohort cost-effectiveness model to estimate the cost-effectiveness of entrectinib compared with established practice in a population of adult and paediatric patients with *NTRK* fusion-positive solid tumours. Established practice consisted of a composite comparator represented through a weighted average of comparators from the tumour types represented in the integrated analysis for entrectinib. Cost-effectiveness was assessed over a lifetime time horizon of 30 years with a 3.5% discount rate applied to both costs and quality adjusted life years (QALYs). No other discount rates were explored in the CS.

The model structure is based on a partitioned survival model (PSM) or "area under the curve" analysis comprising of three mutually exclusive health states: (i) PFS (progression free), (ii) progressive disease (PD; progression), and (iii) death. Within the PFS and PD health states, the model distinguished between patients who are receiving treatment and those who are not. The model predicted the total costs and QALYs separately for the entrectinib arm and the pooled comparator arm. The distribution of patients in each health state was determined by using estimates of PFS and OS.

For entrectinib, these distributions were based on KM data from the *NTRK* efficacy evaluable analysis set. In the comparator arm, estimates of mean OS and PFS for each tumour type were modelled to estimate time in each health state. These estimates of time in state were then used to estimate total costs and QALYs for each tumour type. Total costs and QALYs for the comparator arm were then estimated as weighted averages using the distribution of tumours in the integrated analysis of entrectinib.

The OS and PFS extrapolations for entrectinib were based on the integrated analysis which pooled data from three trials: ALKA, STARTRK-1, and STARTRK-2. The integrated analysis set included 54 patients across 13 different tumour types, but excluded 6 patients with primary CNS and a paediatric patient. The data-cut off used in the economic model was the 31st of May 2018, later updated to the [redacted] cut off at points for clarification. To extrapolate the observed OS and PFS data, the company fitted a number of standard parametric models. The models selected for the company's base-case analysis were extrapolated exponential OS and PFS survival functions.

Comparator OS and PFS data for each tumour type was generated from multiple NICE Technology Appraisals (TAs), which were then weighted by the distribution of tumour types in the integrated efficacy analysis. The OS and PFS data were extrapolated assuming an exponential survival function. As the company extracted only median OS and PFS values and not KM data, no other survival functions were considered.

The estimates used in the company's base-case analysis for health-related quality of life of patients in the PFS and progressive disease health states for entrectinib were based on EQ-5D-3L data collected in the STARTRK-2 study. Due to the small sample size and associated uncertainty, the post-progression utility from the integrated efficacy analysis was not used in the economic analysis. The company therefore assumed that utilities in the PD health state was equal to that of established management. The utilities used for established management were taken from the relevant NICE TAs identified in the clinical effectiveness section. The utilities for each tumour type were weighted according to the distribution of tumour types in the integrated efficacy analysis. In contrast with the approach taken for the comparator efficacy, where a range of estimates for each tumour type were pooled, the utility values extracted for each tumour type were obtained from a single selected TA.

Resource use and costs included: drug acquisition and administration costs, monitoring costs, costs related to health states and adverse events, the cost of subsequent treatments and screening costs. Patient access scheme (PAS) discounts are available for entrectinib, nintedanib, nab-paclitaxel, trifluridine/tipiracil, everolimus, eribulin and trabectedin. For the purpose of simplicity, the company grouped interventions into three classes: oral, simple intravenous (IV) and complex IV and used these costs to estimate drug administration costs as well as the progression-free health state costs based on the interventions comprising established management. For estimation of the screening costs, the company used a hierarchical approach to testing assuming immunohistochemistry (IHC) followed by next generation sequencing (NGS) for the majority of tumour types.

The company found entrectinib to be more costly (cost difference of [redacted]) and more effective ([redacted] QALYs gain) compared with established management. The deterministic base case incremental cost-effectiveness ratio (ICER) was £52,609 per QALY and the mean probabilistic ICER was £52,052 per

QALY. These results do not include PAS discounts available for nintedanib, nab-paclitaxel, trifluridine/tipiracil, everolimus, eribulin and trabectedin. The majority of the QALYs gained were generated as a result of additional life years. The company reported that the most influential parameters in the one-way sensitivity analysis included the comparator OS estimates and the screening costs.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG highlights that there are significant number of issues that contributed to uncertainty in the cost-effectiveness results presented by the company.

The focus of the company's submission was on a single answer to indicate the cost-effectiveness of entrectinib in the population covered by the marketing authorisation. The general view of the ERG is that optimised decisions are preferable and while the ERG acknowledges the challenges presented by the current decision problem, the company could have gone further in justifying the use of a single ICER. In particular, the ERG considers that the company could have explored further the variability in the treatment effect across tumour types, as well as further considering how variability in testing costs impact on the tumour-type specific ICER. The ERG notes the possibility for heterogeneity in the treatment effect across tumour types, as well as across other clinical characteristics such as age (paediatric vs adults), fusion type and position in the treatment pathway, which were not accounted for.

The ERG has several concerns about the representativeness of the modelled population, which was based on the integrated efficacy analysis. These include concerns about the distribution of tumour types modelled, which appear to over represent some tumour types, while under-representing others. Further, the modelled population includes only the 13 tumour types represented in the trials, while there is evidence to suggest that *NTRK* fusions occur in at least another 11 tumour types representing a minimum of 20% of the eligible population. The omission of these patients has a number of implications for the model and potentially impacts upon a number of the inputs used to model established management including, comparator effectiveness, comparator treatment cost, testing costs, and health state utilities. The ERG is also concerned that the analysed integrated efficacy data set excluded available evidence on patients with primary CNS tumour as well as a number of paediatric patients,

[redacted]

There are also significant uncertainties regarding whether the appropriate comparators have been modelled. The anticipated marketing authorisation for entrectinib allows entrectinib to be used and multiple points in the treatment pathway, meaning there is significant uncertainty regarding the

patient group in which entrectinib may be used in practice. It is therefore unclear whether the modelled comparators represent current NHS practice. Further, because the model only considers 13 tumour types and not all tumour types in which *NTRK* fusions may occur, there are a number of relevant comparators not covered by the model. The model therefore implicitly assumes that the modelled population is representative of the eligible population which appears to be unlikely given available evidence on the distribution of tumour types with *NTRK* fusions.

The ERG highlights that the observed data for entrectinib was immature with median OS not yet met. As such, there is significant uncertainty regarding the longer-term survival benefits of entrectinib. The company base-case fits an exponential function to the available KM data, selected from a range of standard parametric functions on the basis that the exponential function has the best statistical fit to the observed data. The ERG considers the exponential function to represent a potentially plausible extrapolation of OS, but is concerned that it implies that post-progression survival is significantly longer than pre-progression survival. The ERG questions the clinical plausibility of this given that entrectinib therapy is discontinued on progression and that only [redacted] of patients received any subsequent therapy. The ERG's preference is therefore for the Weibull function, which produces a more reasonable balance between pre- and post-progression survival while also having good statistical fit to the observed data.

Because the available effectiveness evidence for entrectinib was from single arm studies, it was necessary to generate an appropriate comparator dataset. The company does this by using previous NICE TAs as a source of effectiveness data, which are then weighted by the distribution of tumour types in the integrated efficacy analysis. While the ERG considers the broad approach adopted by the company to be reasonable, there are significant challenges associated with implementing this approach successfully, as well as further issues resulting from the company's execution of this approach.

The ERG's principal concerns regarding the company's approach to generating a comparator is that it relies on an unadjusted naïve comparison between the weighted comparator and the integrated efficacy analysis with significant scope for confounding bias. The ERG in particular notes that a significant proportion of the patients in the integrated efficacy population (37.0%) received entrectinib as a first-line systemic therapy, while the comparator dataset draws predominantly from patients in later lines of therapy. Further, the use of NICE TAs as a source of effectiveness evidence means that comparator effectiveness data is being drawn from a population who are primarily *NTRK* fusion negative. This is problematic because there is evidence to suggest that *NTRK* fusions are prognostic, with variable impact upon prognosis depending upon tumour type.

Because of these significant concerns about confounding bias and the challenges of generating a truly comparable comparator data set, the ERG considers that the company should have also considered other approaches to generating a comparator data set to further explore the uncertainties associated with generating a comparator data set. For example, the company could have utilised two alternative methods outlined in Hatswell *et al.*[1], which would have provided alternative estimates of comparator effectiveness and could have been used to validate the company's base-case.

The ERG also has substantive concerns regarding the companies approach to modelling *NTRK* fusion testing. The ERG in particular is concerned that the company appears to have included extensive testing costs in the comparator arm of the model. The ERG considers that the focus of modelled testing costs should be on the incremental testing costs associated with identifying *NTRK* fusion positive patients.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

ORR rates were clinically meaningful and objective response was observed across all tumour types and lines of therapy included. Clinical efficacy evidence included 13 tumour types in mostly metastatic patients, a paediatric population, and several cancers expected to harbour a larger proportion of patients who may be eligible for entrectinib according to the anticipated marketing population.

1.7 Weaknesses and areas of uncertainty

The main weaknesses and areas of uncertainty identified by the ERG include:

Uncertainty surrounding the homogeneity of the treatment effect

The ERG considers the company's assumption that all tumour types will have identical response rates when treated with entrectinib to be very strong and subject to considerable uncertainty. Analyses presented by the ERG suggest that there is heterogeneity in response and that response rates in tumour types not represented in the trial data could vary considerably from what has been presented.

Uncertainty surrounding the relevant patient population

Significant uncertainties exist regarding the position of entrectinib in the patient pathway. The anticipated marketing authorisation for entrectinib allows patients to be treated when there is [redacted] is ambiguous and is likely to be influenced by subjective assessments of the response rates and adverse event burden associated with existing options.

The choice of comparator regimens

Because of the significant uncertainties surrounding the position of entrectinib in the treatment pathway, it is not clear whether the comparators considered reflect current established management in the treated population.

The uncertainty surrounding the extrapolation of OS for entrectinib

The ERG notes that significant uncertainties remain regarding the extrapolated OS estimates for entrectinib. While the ERG considers that the company's approach based on an exponential function provides reasonable estimates of long-term survival there are concerns about what this implies regarding the split between pre- and post-progression survival.

Uncertainty surrounding the costs of identifying patients the NTRK fusions

Current testing for the majority of tumour types does not routinely include testing for *NTRK* and the rarity of the *NTRK* fusions means that the number needed to screen (NNS) to identify a single *NTRK* fusion positive patient is often high. Testing costs therefore represent a substantial proportion of the incremental costs associated with implementing entrectinib.

A number of plausible testing strategies exist that could be implemented, should entrectinib be approved for use in the NHS, with a range of advantages in terms of the costs and diagnostic performance. There are also significant uncertainties around who will receive testing and when testing will be implemented across tumour types, as knowledge on the tumour types which harbour *NTRK* fusions is current incomplete.

Uncertainty surrounding broader infrastructure and training requirements

The provision of entrectinib on the NHS is likely to substantially increase the number of patients requiring molecular testing. The ERG considers that important uncertainties remain concerning whether the additional resource/cost implications for the NHS have been fully quantified. The ERG notes that particular consideration should be given to whether there are additional infrastructure or training requirements for the NHS which have not been captured.

1.8 Summary of exploratory and sensitivity analyses undertaken by the ERG

The key uncertainties addressed by the ERG scenario analyses relate to:

- The testing costs associated with the implementation of entrectinib;
- The population modelled and the distribution of eligible patients across tumour types;

- Unit costs associated with the chemotherapy regimens that constitute established management;
- Drug wastage associated with entrectinib.

Further to the above, the company presented additional analysis as part of the points for clarification response which incorporated the latest [redacted] data cut; incorporated available effectiveness evidence available for patients with primary CNS tumours as well as several paediatric patients; and made alternative assumptions about the duration of subsequent therapies received by entrectinib patients.

The results of these scenario analyses including the ERG's base-case are summarised in Table 1. Due to time constraints, deterministic ICERs are presented throughout.

The ERG alternative base-case analysis incorporated a number of alternative assumptions, a number of which were also explored by the company in scenario analyses. The changes made by the ERG include:

- Inclusion of children and primary CNS tumours in the population;
- Weibull distribution for extrapolation of entrectinib OS and PFS;
- Inclusion of marginal testing costs only;
- Confirmatory RNA-based NGS test after whole genome sequencing (WGS) test, and removal of NGS testing costs for lung cancer patients;
- Testing costs estimated using the number needed to screen based on the whole *NTRK* population;
- WGS test to identify *NTRK* tumours in paediatric patients,
- Second-line therapy following discontinuation of entrectinib, limited to 6 month duration;
- electronic market information tool (eMIT) costs for therapies in the established management arm;
- Inclusion of drug wastage of entrectinib.

Under the ERG's alternative set of assumptions, the ICER for entrectinib versus established care is £77,109 per QALY.

Table 1 Summary of ERG exploratory analyses

	Incremental costs	**Incremental QALYs**	**ICER (£/QALY)**
Base case	[redacted]	[redacted]	£52,609
Scenario 1: Alternative distribution of tumour types	[redacted]	[redacted]	£69,747
Scenario 2: Remove testing costs in established management arm	[redacted]	[redacted]	£63,329

Scenario 3: Remove lung cancer cost of testing	[redacted]	[redacted]	£59,465
Scenario 4: Confirmatory RNA-based NGS in WGS patients	[redacted]	[redacted]	£64,608
Scenario 5: Prevalence of *NTRK* fusions (tumour types represented in the trial)	[redacted]	[redacted]	£56,914
Scenario 6: Prevalence of *NTRK* fusions (based on the whole *NTRK* population)	[redacted]	[redacted]	£65,981
Scenario 7: Cumulative impact of 2, 3, 5, 7	[redacted]	[redacted]	£64,115
Scenario 8: No testing costs	[redacted]	[redacted]	£36,914
Scenario 9: WGS for identifying *NTRK* tumours in paediatric patients *	[redacted]	[redacted]	£48,860
Scenario 10: eMIT costs for therapies in the established management arm	[redacted]	[redacted]	£52,081
Scenario 11: With drug wastage	[redacted]	[redacted]	£55,357
ERG alternative base-case analysis **	[redacted]	[redacted]	**£77,109**

* These results should be compared to the analysis including primary CNS and paediatric patients, see Table 50.
** These results have been updated by the ERG following the factual accuracy check to include the change made in Scenario 9

The ERG also presented a further scenario analysis using the ERG's base assumptions in which an alternative model structure was used where PFS and OS were determined according the ORR. This method used the survival of non-responder patients to estimate survival predictions in the established management arm. The entrectinib arm was based on a weighted average of responder and non-responder survival predictions, which allowed for the exploration of cost-effectiveness in different tumour types by varying the response rate used to estimate the weighted average. The ICER for the pooled group was £95,705 per QALY. When varied by tumour type, the ICERs ranged from £57,451 per QALY for sarcoma patients, to £128,663 for thyroid tumours.

In further exploratory analysis using the response-based model, the ERG also presents an example of how a response-based model can be used estimate the value of heterogeneity and the population net health effect, so as to potentially permit optimised decisions that would limit the provision of entrectinib to those patients in which it is most cost-effective. Using the tumour type CRC as an example, an 'optimised' recommendation which excludes CRC might result in an additional 12.99 QALYs per year to the health system.

2 Background

2.1 Critique of company's description of underlying health problem

The present appraisal concerns the treatment of locally advanced or metastatic solid tumours exhibiting gene fusions involving neurotrophic tyrosine receptor kinase (*NTRK*) genes 1, 2, and 3 in any solid tumour. This is the first time a technology has been appraised for a histology-independent indication, with treatment determined by the presence of a specific type of genomic alteration, rather than the location of the tumour.

The CS describes that advances in techniques used to identify particular gene fusions have enabled the development of therapies directed specifically at the molecular targets responsible for the growth of cancer cells, and that *NTRK* gene fusions are 'clinically actionable' drivers of solid tumour formation and development across a wide variety of sites. The underlying health condition considered in this appraisal is therefore defined with respect to the presence of *NTRK* fusions and not tumour type. As such, in contrast with other NICE appraisals of cancer therapies (where the indication considered is a single tumour type), this appraisal considers any solid tumour exhibiting the *NTRK1*, *2* or *3* gene fusions.

The ERG considers the company's description of the underlying health problem to be appropriate and relevant to the decision problem under consideration. The company describes the role of the tropomysin receptor kinases (Trks) in the development and function of neurons in the central and peripheral nervous system. These receptor proteins can be expressed in a variety of tissue types and are involved in the regulation of function, proliferation, and survival of cells. *NTRK* gene fusions occur when the 3' region of *NTRK* gene is joined with the 5' sequence of a fusion partner gene by a chromosomal rearrangement event. This results in the over-production of a chimeric Trk protein which is permanently 'switched on', meaning cell survival and proliferation are decoupled from normal regulatory processes, which may lead to oncogenesis. The ATP-binding sites of the TrkA/B/C proteins share high structural similarity,[2] which entrectinib exploits to inhibit the activity of chimeric receptors to stop or reverse the growth of *NTRK* fusion-positive tumours.

The company suggests that the prognosis of patients with *NTRK*-fusion positive tumours is worse than those without this genomic alteration, and provides an example of a study in colorectal cancer patients in which shorter median overall survival (OS) is observed for patients with *NTRK*, *ALK*, or *ROS1* gene rearrangements.[3] However, this is a small study and does not report survival data by gene arrangement type, and therefore in the ERG's view cannot be considered conclusive. Furthermore, the ERG considers it more likely that the relative prognosis of patients with *NTRK* fusions will vary between cancer types, and that outlook could also plausibly vary by which of the three *NTRK* genes is

involved. This is supported by evidence from other cancer types. For example, in a study of patients with papillary thyroid cancers, prognosis was similarly found to be worse in patients with *NTRK* fusions when compared to those without,[3-5] while the presence of *NTRK* fusions in a mesoblastic nephroma patient population was associated with more favourable outcomes in another study.[6] From the evidence available, it also is unclear whether *NTRK* fusions are in themselves prognostic, or whether it is their association with other specific prognostic factors such as age and ECOG status that drives the observed differences in prognosis.

2.1.1 Prevalence of *NTRK* gene fusions

The CS estimates that *NTRK* gene fusions are present in 0.7% of all cancers, based on a weighting of literature prevalence estimates with figures observed in the entrectinib clinical trial; however, the ERG notes this is significantly higher than other figures reported in the literature sources referenced by the company. Excluding the estimate derived from the entrectinib trial, the prevalence of *NTRK* gene fusions is reported to be between 0.25% - 0.31% in the adult population[7-9] and 0.34% – 0.49% in the paediatric/adolescent population.[7, 8, 10, 11] The Foundation Medicine Inc. dataset cited by the company found [redacted]% of ~116,000 samples harboured an *NTRK* gene fusion. As the largest epidemiological study available, the ERG considers this figure the most reliable estimate of *NTRK* gene fusion prevalence.

This lower figure impacts upon the number of patients who would be eligible to receive entrectinib in clinical practice. The CS estimates that the number of patients eligible to receive entrectinib, i.e. those with *NTRK* fusion positive advanced or metastatic cancer, would be 648 per year in England. This figure is based on a number of assumptions: any cancer can harbour an *NTRK* gene fusion; fusions occur in 0.7% of all cancers; 34% of all cancers are advanced or metastatic; and that 90% of patients are fit enough for treatment. The ERG suggests this figure may be an overestimate, and provides an alternative estimate based on a different set of assumptions.

The company's population size estimate includes patients with any type of cancer, rather than just those with solid tumours as described in the anticipated marketing authorisation of entrectinib. Using the company's assumption of eligibility by cancer stage, but limiting the eligible population to solid tumours with a prevalence of *NTRK* of [redacted]%, the number of patients eligible for entrectinib in England is reduced to [redacted] individuals annually.

The ERG's estimate of patients eligible for entrectinib uses a bottom-up approach, where a total population size was calculated by using tumour-specific rates of *NTRK* gene fusions and disease incidence. The tumour types included in these calculations are presented in Table 2.

Table 2 Tumour types included in ERG population size calculations

Tumour Type	
MASC	Cervix
NSCLC (Adenocarcinoma & squamous cell carcinoma)	Soft tissue sarcoma
Breast cancer	Head and neck squamous cell carcinoma
Secretory breast carcinoma	Salivary gland (non MASC)
Papillary thyroid tumour	Sinonasal adenocarcinoma
Thyroid tumour	Gastro-oesophageal junction
Colon/colorectal	Prostate cancer
Melanoma	Renal cell carcinoma
Neuroendocrine	Low-grade glioma
Gastrointestinal stromal tumour	High grade glioma (inc. glioblastoma multiforme)
Cholangiocarcinoma	Paediatric high grade glioma
Pancreatic	Congenital mesoblastic nephroma
Appendix	Paediatric melanoma
Uterine	Infantile fibrosarcoma
Ovarian	Paediatric low grade glioma

The ERG considered it appropriate to limit the population to patients at the relevant stage of the treatment pathway for each tumour type (i.e. in line with the proposed positioning of entrectinib), thus yielding a more representative estimate of the population eligible for entrectinib in practice. Using these assumptions, the ERG estimate that 196 patients would become eligible for entrectinib every year in England. Clinical advisers to the ERG suggested that it is possible that *NTRK* fusions may present in any tumour type, so the ERG's estimate of the eligible population is likely to be conservative, as it does not account for cancers in which an *NTRK* fusion has not yet been identified.

Further details on the calculation of the size of the eligible population are presented in Appendix A.

2.2 Critique of company's overview of current service provision

2.2.1 Treatment pathways

The company states that there is currently no established treatment pathway for patients with *NTRK* fusion-positive tumours, with treatment guided by tumour type-specific care guidelines. The position at which *NTRK* fusion-positive cancer patients would be offered entrectinib is likely to vary by the availability of other effective treatments in each tumour. This is reflected in the anticipated marketing authorisation, which covers entrectinib as a treatment option for [redacted] The company's interpretation appears to position entrectinib as an alternative to standard chemotherapy when one or more other

options have been exhausted, or as a first-line option where there are no acceptable alternatives. However, what constitutes an '[redacted]' is ambiguous and may be affected by the availability of entrectinib itself. Clinical advice to the ERG suggested that 'acceptability' would be a subjective assessment of the response rates and adverse event burden associated with existing options, but the threshold at which a decision to offer entrectinib would be made is likely to vary between indications. Current availability of testing for targeted therapies in each indication is also likely to influence the positioning of entrectinib if *NTRK* fusion testing is added to existing screening processes; in those indications with early testing for other genetic oncodrivers it is likely that entrectinib will be used in place of other chemotherapy options. However, in their clarification response the company stated that they anticipate entrectinib to be used in later lines of treatment in the majority of cases, at the point where therapeutic options are very limited or exhausted altogether. The company also provided an outline of where they expect entrectinib to be offered within existing treatment algorithms for patients included in the integrated efficacy analysis, reproduced in Table 3.

Table 3 Proposed positioning of entrectinib for the treatment of *NTRK* fusion-positive, locally advanced or metastatic solid tumours (Reproduced from CS Table 6, Page 30)

Position of entrectinib in line of systemic therapy	
First-line*	**Second-line and beyond†**
MASC	NSCLC
Soft-tissue sarcoma	Breast
Pancreatic cancer	Thyroid cancer
Cholangiocarcinoma	Colorectal cancer
Gynaecological cancers	Neuroendocrine tumours

*Patients ineligible for curative surgery or radiotherapy with no immunotherapy or targeted therapy options
†Some patients may receive first-line entrectinib treatment if not eligible for targeted or immunotherapies

The ERG do not consider the company's definition of current treatment pathways and the anticipated positioning of entrectinib to sufficiently address the decision problem. Firstly, the as-yet undetermined timing of testing within each tumour type will inevitably define the eligible population. Secondly, the groupings of tumour types as presented by the company are too broad to accurately represent the diversity of cancer types and different treatment options available within each. For example, neuroendocrine and gynaecological cancers comprise numerous specific indications with differing prognoses and treatment options recommended by NICE. Furthermore, it is likely that entrectinib will be offered at different points in the respective treatment pathway of the tumour types covered by these umbrella terms.

2.2.2 *NTRK* fusion diagnostic pathways

2.2.2.1 Testing for *NTRK* gene fusions

A number of testing strategies are available for the identification of *NTRK* gene rearrangements across different tumour types. These include fluorescent in-situ hybridisation (FISH), immunohistochemistry (IHC), ribonucleic-acid (RNA)-based next generation sequencing (NGS) or reverse transcriptase polymerase chain reaction (RT-PCR).[12]

FISH testing is commonly used to detect chromosomal abnormalities such as *ALK* and *ROS1* gene rearrangements.[13, 14] FISH is used to identify a single specific gene fusion, so if a particular fusion is common in a tumour type, it can be efficient to use FISH. The NHS currently offers FISH for the detection of the highly prevalent *ETV6-NTRK3* gene fusion in MASC patients.[15] However, where the type of *NTRK* fusion is unknown, then three individual tests would need to be conducted in order to detect the presence of one of the three *NTRK* rearrangements.

Immunohistochemistry detects overexpression of Trk proteins, a subset of which may be the result of NTRK gene fusions. Unlike FISH, all three *NTRK* fusions can be tested in one IHC test by using a pan-*TRK* fusion panel. IHC is quick and inexpensive, and is currently used for a variety of gene rearrangements across tumour types in the NHS.[12]

Next generation sequencing methods that use rapid sequencing of RNA and DNA can be used to detect *NTRK* fusions. DNA-based NGS can be used to detect multiple chromosomal rearrangements from a single sample,[16] and is currently used as a diagnostic and prognostic method in oncology for a range of tumour types.[15] However, there are concerns that DNA-based NGS panels will not identify all *NTRK* fusions; for fusions where there is a large intron size, DNA-based NGS may be limited and may provide inaccurate results.[17] In research, DNA-based NGS panels to detect *NTRK* fusions have to be confirmed with RNA-sequencing or IHC.[17] RNA-based NGS can detect *NTRK* fusions independent of *NTRK* fusion type,[18] and is often seen as the 'gold standard' of testing for gene fusions if RNA quality is high.[17] NGS is substantially more resource intensive than FISH and IHC, with longer turnaround times and higher quality sample requirements.

More recently, hybrid DNA/RNA NGS panels have been developed, allowing DNA and RNA to be extracted and run simultaneously in one test.[12] The Oncomine Focus Fusion Assay, for example, screens for 161 cancer-associated gene rearrangements. Like FISH, knowledge of fusion partners is necessary in order to identify gene rearrangements. As these assays are in constant development, the addition of newly-identified mutations is relatively straightforward.[12]

2.2.2.2 Testing strategy options

There is currently no established strategy for detecting for *NTRK* fusions across tumour types in the NHS. For the tumours where WGS is currently unavailable, and with the exception of MASC, where patients receive the *ETV6-NTRK3* FISH test, *NTRK* fusions are not routinely tested for in solid tumours.

The company propose a two-tiered testing approach. First, IHC testing is used to detect the presence of an *NTRK* gene fusion. For individuals with a suspected *NTRK* fusion, confirmatory NGS will be used to identify the particular gene rearrangement. The CS does not include MASC patients and patients eligible for whole genome sequencing (paediatric tumours and sarcoma) in this testing strategy. However, the ERG was advised that current NHS WGS may not accurately detect *NTRK* fusions or other structural abnormalities, therefore confirmatory RNA-based NGS would still be required in these patients.

The company suggest the use the Roche Ventana pan-TRK assay for IHC testing, which detects Trk proteins A, B, and C, identifying any *NTRK* gene rearrangement in one test. The CS states that Ventana assay eliminates 89% of *NTRK* fusion-negative samples. However, the sensitivity of other pan-TRK IHC assays have been estimated to be as low as 55% for *NTRK3* fusions.[7] Thus, up to half of the individuals with an *NTRK3* rearrangement could incorrectly test negative. IHC also has limited predictive value in neural and smooth muscle tumours, where false positives occur due to the natural expression of Trk in these tissues.[12] The Oncomine Focus Fusion assay, a hybrid DNA/RNA assay recommended by the company, has a high specificity and sensitivity (both 100%), from the small sample available.[19]

There have been a variety of alternative testing algorithms proposed for the identification of *NTRK* fusions,[17] most of which suggest that the testing approach should vary depending on the prevalence of *NTRK* fusions, and the provision of genomic testing currently available.[12, 16]

The extent and purpose of current testing provision varies across tumour types, with some genomic and histological testing for specific genetic abnormalities already in place for specific cancer types.[20] Table 4 provides details of genomic and molecular testing currently available for tumour types with known *NTRK* fusions. There is some form of genetic or molecular testing available in the majority of tumour types with a known *NTRK* fusion. With the exception of gastrointestinal stromal tumours, NGS is not routinely provided to every patient with a particular tumour type. Eligibility for testing often depends on the histology and the sub-type of the tumour. For example, pre-surgery IHC is routinely offered for all individuals with invasive breast cancer at the time of diagnosis, and only women under 50 years old with triple negative breast cancer are eligible for *BRCA1* and *BRCA2* NGS testing.[21] The majority of NGS testing available on the NHS is currently DNA-based. There are

concerns that DNA-based NGS panels will not identify all *NTRK* fusions[17]. For the of cancer types where DNA-based NGS is currently offered,additional RNA fusion-based or RNA/DNA hybrid NGS will be required to confirm *NTRK* rearrangements.

Table 4 Current molecular and genetic testing for tumour types with *NTRK* fusions on the NHS

Tumour Type	**Frequency of *NTRK* fusion**	**Current Molecular Testing**
MASC	100.00%	FISH (*ETV6-NTRK3*)
NSCLC (Adenocarcinoma & squamous cell carcinoma)	[redacted]	IHC (22C3) Multi-target NGS panel (*EGFR*) [22]
Breast cancer	[redacted]	IHC (HER2)[21] Multi-target NGS panel: (Oncotype DX)[22]
Thyroid tumour	[redacted]	IHC for Papillary Thyroid Tumour[23] Multi-target NGS Panel (*BRAF*, *KRAS*, *NRAS*, *HRAS*)[23]
Colon/colorectal	[redacted]	IHC for Lynch Syndrome (hereditary CRC)[24] Multi-target NGS Panel (*BRAF*, *KRAS*, *NRAS*) [22]
Melanoma	[redacted]	Multi-target NGS Panel (*BRAF*, *NRAS*, *KIT*) [22]
Neuroendocrine	[redacted]	No Routine Testing Available
Gastrointestinal stromal tumour	[redacted]	IHC (*CD117*, *C134*, *DOG1*) [25] Multi-target NGS Panel (*KIT*, *PDGFRA*)[22]
Cholangiocarcinoma	[redacted]	No Routine Testing Available
Pancreatic	[redacted]	No Routine Testing Available
Appendix	[redacted]	No Routine Testing Available
Uterine	[redacted]	IHC (*EMA*, Ber-EP4, *PAX8*, *CK7*)[26] (REF) FISH (*EPC1-PHF1*)
Ovarian	[redacted]	IHC[27] Multi-target NGS panel (*BRAC1*, *BRAC2*)[22] Multi-target NGS panel (*SMARCA4*)[22]
Cervix	[redacted]	IHC[28]
Soft tissue sarcoma	[redacted]	Whole Genome Sequencing
Head and neck carcinoma	0.24%	IHC (HPV) [29] Multi-target NGS Panel – (*CDKN2A*, *EGFR*, *TP53*)[22]
Prostate cancer	[redacted]	IHC (PSA)[30]
Renal cell carcinoma	[redacted]	FISH/RT-PCR (*TFE3*)

High grade glioma (inc. glioblastoma multiforme)	[redacted]	IHC[31] Multi-target NGS (*IDH1*, *IDH2*, *ATRX*, *TERT*, *H3F3A*)[22] Multi-target NGS (*BRAF*), MGMT promotor hypermethylation[22]
Paediatric high grade glioma	5.30%	Whole Genome Sequencing
Congenital mesoblastic nephroma	60.70%	Whole Genome Sequencing
Paediatric melanoma	11.11%	Whole Genome Sequencing
Infantile fibrosarcoma	90.90%	Whole Genome Sequencing
*The frequency NTRK fusions in appendix tumours in the FMI data set was reported to be 0%, however it has been reported to be higher than 0% in the literature[2] MASC, mammary analogue secretory carcinoma; NSCLC non-small cell lung cancer		

According to the expert advice received by the ERG, the only RNA-based NGS fusion panel available NHS is for a specific subgroup of NSCLC patients, targeting a range of genes including *EGFRALK*, and *ROS1*. Whilst this panel does not currently target *NTRK1-3* rearrangements, genomic advisers informed the ERG that the costs of adding additional gene targets to an RNA-based NGS panel are nominal.

The European Society for Medical Oncology (ESMO) propose that the standard testing pathway should differ depending on the frequency of *NTRK* fusions in each tumour type, and whether sequencing is currently provided by the NHS. It is recommended that FISH, RT-PCR, or targeted NGS assays are used first line in tumour types known to have a high prevalence of *NTRK* fusions and where other NGS is already available, and IHC where it is not.[17] In those tumour types thought to have lower frequencies of *NTRK* fusions and where current genomic testing is available, ESMO recommends the use of front-line NGS, followed by confirmatory IHC. In the tumour types where there is thought to be a lower frequency of *NTRK* fusions and where there is no genomic testing available, it is suggested that IHC is used for initial screening; *NTRK* gene rearrangements are then confirmed using NGS. A similar approach, suggested by Penault-Llorca *et al.*[16] is presented in Table 5.

Table 5 Alternative screening pathways according to prevalence

Prevalence of *NTRK*	**Testing strategy**
High prevalence of *NTRK* gene fusions	FISH or IHC
5-25% prevalence of *NTRK* gene fusions	NGS panel
< 5% prevalence of *NTRK* gene fusions	NGS panel
< 5% prevalence of *NTRK* / gene fusions not common	IHC then confirmatory NGS

2.2.2.3 Feasibility of *NTRK* fusion Screening

As discussed in Section 2.1.1, there are 28 tumour types observed to harbour *NTRK* gene fusions, but the ERG's clinical advisers suggested that *NTRK* fusions could potentially occur in any tumour type. Therefore, the feasibility of screening for *NTRK* fusions in all tumour types should be considered.

Using the company's proposed diagnostic testing strategy (IHC followed by confirmatory NGS), and their top-down estimates of the annual population eligible for entrectinib (CS, Budget Impact Model), [redacted] individuals would require testing using IHC every year. If the assumptions of the diagnostic accuracy reported of IHC (89% *NTRK* fusion-negatives identified) are used, [redacted] confirmatory NGS tests would be required per year. This is likely to be an overestimate, as the population defined by the company includes patients with haematological cancers. If these patients are excluded, the estimated number of individuals that would require NGS is reduced to [redacted], with [redacted] individuals requiring confirmatory NGS per year. The company do not consider the positioning of entrectinib, or when testing would be offered, so the size of this population is still likely to be higher than what would be expected in practice. For further details on the company's assumptions and calculations, see Appendix B.

The ERG used a conservative, bottom-up approach to calculate the number requiring testing, based on the tumour types in which there is a known *NTRK* fusion (see Table 2). Using the company's proposed diagnostic algorithm (IHC followed by confirmatory NGS), the number of additional IHC tests required to identify patients meeting the anticipated marketing authorisation of entrectinib would total approximately 51,958 a year in England. Based on the diagnostic accuracy figures supplied by the company, this would mean 5,806 patients would require confirmatory NGS tests annually. For further details on the ERG's assumptions and calculations, see Appendix B.

In order to provide testing for *NTRK* gene fusions, sufficient capacity in genomic testing services is required. With the increasing number of targeted medicines available, the number of individuals requiring genetic testing is increasing. Cancer Research UK estimated that in 2014, 16,000 patients with colorectal cancer or NSCLC did not receive molecular testing, with 3,500 of those individuals expected to be eligible for some form of targeted medicine.[32] To ensure that individuals are able to access the appropriate testing, and consequently, correct targeted medicine, substantial investment in the NHS genomics services are needed to increase capacity and to ensure that staff have appropriate skills and training for specific genetic analysis. There is also an additional need for education and training to ensure that clinicians are aware of where targeted medicines could fit within each cancer type's treatment pathway. Clinical advisers to the ERG report that the provision of testing for patients

can be dependent on their clinician's knowledge of genomic medicine and available targeted therapies.

In addition to the requirement for a larger workforce, investment in laboratory infrastructure is needed to ensure sufficient equipment is available to deal with increasing demands.[33] In laboratories where RNA-based NGS or RNA/DNA hybrid-based NGS is not available, substantial investment would be required to provide infrastructure to enable NGS testing.

While the company acknowledge that screening for *NTRK* is likely to impact significantly upon the total cost of identifying and treating patients with entrectinib, the scale of practical and infrastructural considerations associated with the introduction of such a vast number of tests to NHS pathology services is not addressed. As new tests for molecular markers are introduced, increasing numbers of patients being referred for increasingly complex diagnostic investigations continues to outstrip the ability of the service to increase testing capacity. Cancer Research UK predicts a "severe crisis" in pathology capacity in the next 5-10 years,[34] and the ERG's clinical advisers agreed that existing infrastructure could not accommodate the proposed increases in IHC testing without significant NHS investment. Capacity constraints have been identified as a key barrier to the introduction of precision medicines onto the NHS, but investment in increasing capacity is rarely considered in cost-effectiveness evidence.[33] Therefore economic evaluations that fail to integrate these considerations may not provide meaningful evidence on how to implement precision medicines in a cost-effective way.

3 Critique of company's definition of decision problem

3.1 Population

The clinical efficacy evidence submitted by the company included Trk inhibitor-naïve patients with *NTRK* fusion-positive solid tumours that is limited to 10 tumour types included in the entrectinib clinical trials. This includes sarcoma, NSCLC, MASC, breast cancer, thyroid cancer, CRC, neuroendocrine tumour, pancreatic cancer, gynaecological and cholangiocarcinoma. In descriptions of the integrated efficacy analysis, the company do not differentiate between patients with different breast cancers, thyroid cancers or gynaecological cancers. Throughout this report the ERG therefore refers to 13 tumour types to reflect these subtypes. Following request for clarification, the company provided further ORR data for three additional tumour types, including three in a paediatric population (primary CNS, infantile fibrosarcoma and skin cancer) and one in adults (primary CNS).

Table 6 presents an overview of the 13 tumour types for which the CS presented efficacy data, for a total of 66 patients. The number of patients representing each of the tumours included in the entrectinib efficacy evidence is small, ranging from one (cholangiocarcinoma, paediatric skin cancer) to 13 (soft tissue sarcoma). The most frequently represented solid tumour types in the trial evidence were sarcomas (19.7%), NSCLC (15.2%), salivary gland tumours (MASC) (10.6%), and breast cancer (9.1%), which together accounted for over half of patients (54.7%). However, there is a mismatch between the distribution of tumour types in the efficacy population and the estimated yearly prevalence in England calculated by the ERG. For instance, the efficacy evidence included four patients with secretory breast carcinoma, over 10 times the estimated prevalence of the eligible population (0.3/year). Other over-represented populations include sarcoma (13 patients included, yearly prevalence 4) and MASC (7 patients, vs. 2). Conversely, the tumour types included in the efficacy evaluable population with the highest estimated eligible population in England were represented by relatively fewer patients: three papillary thyroid tumour cancers (26/year), three pancreatic cancer (15/year) and four CRC patients (14/year).

The ERG estimate that, of the 13 tumour types included in the trial evidence (including CNS primary and the included paediatric population), approximately 159 patients per year will be eligible for entrectinib. This represents 81.0% of the ERG's estimated annual Trk-inhibitor eligible population of 196 patients, which includes CNS primary and paediatric patients (see Appendix A). This indicates that the trial evidence includes a number of tumour types likely to harbour a larger number of patients who would be eligible for entrectinib according to the anticipated marketing authorisation.

The CS noted that in clinical practice, *NTRK* gene fusions may be present in additional tumour types and histologies. Clinical advisers to the ERG noted that theoretically *NTRK* fusions may be present in over 400 solid tumour types. The Foundations Medicine Inc. dataset only identified *NTRK* fusion in

41 tumour types, out of circa 116,000 samples. Therefore a significant number of tumour types and populations known to harbour *NTRK* fusions are not represented in the entrectinib efficacy evidence.

Table 6 Tumour types included in the entrectinib efficacy evidence

Tumour type/population (high level)	**Tumour Type (low level)**	**N included in efficacy evidence**	**ERG estimated prevalence per year (eligible population)[#]**
Salivary gland (MASC)	[redacted]	[redacted]	2
Lung	[redacted]	[redacted]	10
Breast	[redacted]	[redacted]	4
	[redacted]	[redacted]	0.3
Thyroid	[redacted]	[redacted]	26
	[redacted]	[redacted]	NE
	[redacted]	[redacted]	NE
Colon/colorectal	[redacted]	[redacted]	14
Neuroendocrine	[redacted]	[redacted]	4
Cholangiocarcinoma	[redacted]	[redacted]	0.3
Pancreatic	[redacted]	[redacted]	15
Sarcoma/soft tissue sarcoma	[redacted]	[redacted]	NE
	[redacted]	[redacted]	NE
	[redacted]	[redacted]	NE
	[redacted]	[redacted]	NE
	[redacted]	[redacted]	3
	[redacted]	[redacted]	NE
	[redacted]	[redacted]	4
Gynaecological	[redacted]	[redacted]	1
	[redacted]	[redacted]	3
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]		66	97

[#] according to the positioning of entrectinib presented in CS table 6; values were rounded to nearest integer unless <1. [+] Estimated prevalence of thyroid tumour (NOS) was 5.6; [redacted]

Table 7 summarises the characteristics of *NTRK* fusion positive patients included in the efficacy evaluable population (EEA). This includes the combined population of 54 adult patients across 10

tumour types (NSCLC, MASC, sarcoma, breast cancer, thyroid cancer, CRC, neuroendocrine tumour, pancreatic cancer, gynaecological and cholangiocarcinoma) with at least 6 months follow-up enrolled into entrectinib studies up to November 30th 2017, and excludes adults with CNS primary tumours and paediatric patients. Given the limited evidence on the *NTRK* population the extent to which trial population characteristics reflect those of the broader population under the NICE scope is difficult to assess. Clinical advisers to the ERG confirmed the EEA population characteristics were broadly representative of the population defined in the anticipated marketing authorisation beyond standard of care

Table 7 Characteristics of *NTRK* trial population (from CS Table 9)

Characteristic	**Description**	***NTRK* efficacy cohort (n=54)**
Age (years)	Median (range)	57.5 (21-83)
	<65	34
	≥65	20
Gender	Female	59.3%
Race	White	79.6%
	Asian	13.0%
	Not reported	7.4%
Mean BSA, m^2 (SD)		1.85 (0.26)
Mean BMI, kg/m^2 (SD)		25.68 (5.30)
Median time since diagnosis, months (range)		21.4 (2.1–433.1)
Disease stage at initial diagnosis, n (%)	0, I or II (A/B)	15 (28.3)[a]
	III (A/B/C) or IV	33 (62.3)[a]
	Unknown	5 (9.4)[a]
Performance status	ECOG 0	42.6%
	ECOG 1	46.3%
	ECOG ≥2	11.1%
Smoking status	Never-smoker	56.6%
Metastatic disease	Any site	96.3%
	Baseline CNS metastases	20.4%
No. of lines of therapy since metastatic disease[b], n (%)	0	37.0%
	1	20.4%
	2	25.9%
	3	7.4%
	≥4	9.3%
Previous therapy[c]	Any systemic therapy	88.9%

	Surgery	79.6%
	Radiotherapy	66.7%

CNS, central nervous system; ECOG, Eastern Cooperative Oncology Group
[a] Percentages calculated based on denominator of 53 patients as one patient in the ALKA study for whom the initial diagnosis field on the Case Report Form was blank was excluded.
[b] Patients may have received other therapies in the adjuvant or neo-adjuvant setting that are not included as a line of therapy from the time of metastatic disease diagnosis.
[c] Includes chemotherapy, immunotherapy, targeted therapy or hormonal therapy.

The EEA population includes only one *NTRK2* fusion positive patient. The company did not provide data on *NTRK2* prevalence in the adult primary CNS and paediatric population. Epidemiological evidence of *NTRK* suggests that the prevalence of subtypes 1/2/3 vary across tumour types,[2] although the ERG agrees with the company that estimates are uncertain given the rarity of *NTRK* fusions and variation in testing methods. Clinical advisers to the ERG noted that there is theoretically no reason to suggest that only one type of *NTRK* fusion should be present between patients within any given tumour type. In response to a clarification request from the ERG, the company stated that the low *NTRK2* prevalence in the trial population is reflective of that observed in the wider *NTRK* population ([redacted]) according to the Foundation Medicine Inc. dataset. This is much higher than the 1.9% prevalence reported in the EEA population, therefore the ERG believe that the *NTRK2* population is significantly underrepresented in the efficacy evaluable population.

The company's proposed positioning of entrectinib is as first line therapy for five tumour types (MASC, soft-tissue sarcoma, pancreatic cancer, cholangiocarcinoma and gynaecological cancers) and second line or beyond for five tumour types (NSCLC, breast, thyroid cancer, CRC, neuroendocrine tumours) although the CS noted that some of these patients may receive entrectinib as first line systemic therapy if not eligible for targeted treatments or immunotherapies. In response to clarification, the company positioned entrectinib as second line for the following tumour types: CNS primary (adults), CNS primary (paediatric), sarcoma (paediatric); and as second line or beyond for paediatric skin cancer (paediatric). The company stated they anticipated use of entrectinib in [redacted]

Following a request from the ERG, the company reported individual participant data (IPD) including the number of lines of prior systemic therapy since diagnosis for the efficacy evaluable population, as well as five *NTRK* patients with primary CNS and seven paediatric patients who were excluded from the efficacy evaluable population, from the [redacted] clinical cut-off data. Table 8 presents the distribution of this population by line of therapy. This shows that [redacted] of patients received entrectinib as first line systematic therapy, [redacted] as second line, and [redacted] as third line or beyond. The company provided no breakdown of line of therapy by tumour type received between 3rd line and subsequent lines. Table 8 shows that entrectinib was administered as either first line or as a subsequent line of therapy in all tumour types except cholangiocarcinoma, gynaecological cancers and paediatric skin cancer. The

absence of an alternative "acceptable therapy" was not an eligibility criterion in STARTK-2, which formed the large majority of the total clinical efficacy population. For this reason, some trial participants may not match the population as defined in the NICE scope.

Table 8 indicates there is a mismatch between the proposed positioning of entrectinib and the trial population evidence submitted. For all five cancer types proposed as 1st line therapy, only 35% of the efficacy evaluable population received entrectinib as 1st line therapy. For the remaining cancer types that were positioned as 2nd line or beyond, 40% of patients received entrectinib as first line therapy; however, it is not clear what proportion of these patients received entrectinib as first line because they were not eligible for targeted treatment or immunotherapies. There was no overlap between the proposed positioning of entrectinib and the trial population submitted for two cancer types: although positioned as first line therapy, patients with cholangiocarcinoma and gynaecological cancers received entrectinib as 3rd line therapy and/or beyond. Overall, this mismatch limits the extent to which the trial evidence supports the company's proposed positioning of entrectinib. Although a higher response rate and better survival outcomes may be expected from patients receiving entrectinib as first line therapy, there is insufficient survival outcomes evidence to determine whether this mismatch may have favoured entrectinib. This matter is further discussed in section 4.2.6.1.

Table 8 Distribution of *NTRK* participants by line of systemic therapy and tumour type (efficacy evaluable population +5 CNS primary adults and 7 paediatric patients)

Line of therapy	**1st line**	**2nd line**	**3rd line & beyond**	**Total**
*Company proposed positioning: 1st line**				
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	1
Gynaecological (endometroid, ovarian)	[redacted]	[redacted]	[redacted]	2
Pancreatic	[redacted]	[redacted]	[redacted]	3
Salivary glands (MASC)	[redacted]	[redacted]	[redacted]	7
Sarcoma	[redacted]	[redacted]	[redacted]	13
Total (of tumour types proposed as 1st line)	[redacted]	[redacted]	[redacted]	**26**
Company proposed positioning: ≥2nd line*				
Breast	[redacted]	[redacted]	[redacted]	6
CRC	[redacted]	[redacted]	[redacted]	4
Neuroendocrine	[redacted]	[redacted]	[redacted]	3
NSCLC	[redacted]	[redacted]	[redacted]	10
Thyroid	[redacted]	[redacted]	[redacted]	5
CNS primary (adults)	[redacted]	[redacted]	[redacted]	5
CNS primary (paediatric)	[redacted]	[redacted]	[redacted]	4
Sarcoma (paediatric)	[redacted]	[redacted]	[redacted]	2

Skin cancer (paediatric)	[redacted]	[redacted]	[redacted]	1
Total (of tumour types proposed as ≥2nd line)	[redacted]	[redacted]	[redacted]	**40**
Total (all tumour types)	[redacted]	[redacted]	[redacted]	**66 (100%)**

*Patients ineligible for curative surgery or radiotherapy with no immunotherapy or targeted therapy options

Some patients may receive first-line entrectinib treatment if not eligible for targeted or immunotherapies

In response to clarification questions, the company provided data on the subsequent therapies received by the trial participants. [redacted] patients in the efficacy evaluable population received a subsequent chemotherapy after progression. A number of these were [redacted]. This suggests that for these patients there are 'acceptable' alternative standard therapies available besides chemotherapy, hormone therapy or best supportive care. No information was provided on the total number of patients receiving subsequent targeted therapies, or on which patients received which subsequent targeted therapy. This further limits the extent to which the integrated efficacy analysis matches the proposed population as defined in the NICE scope.

In summary, the trial population includes only a subset of the total population that is potentially eligible for entrectinib. Although the trial evidence includes a number of tumour types expected to harbour a larger number of patients who may be eligible for entrectinib according to the anticipated marketing population, the large majority of tumour types potentially harbouring *NTRK* fusions are not represented in the evidence submitted. There is also a mismatch between the trial population and the proposed positioning of entrectinib. For an unknown number of patients, there appeared to have been 'acceptable' alternative standard therapies available besides chemotherapy, hormone therapy or best supportive care. Due to concerns about the large number of missing tumour types, the under-representation of *NTRK2* patients, the small sample size of the *NTRK* efficacy trial population, and concerns about the positioning of entrectinib in the trial evidence, the ERG is concerned that the population presented in the evidence submitted does not match the NICE final scope. In particular, the high proportion of patients receiving entrectinib in earlier lines of therapy across tumour types may lead to overestimating its survival benefits.

3.2 Intervention

The intervention is entrectinib at the recommended dose of 600 mg orally once daily for adults, and 300 mg/m^2 orally, once daily for paediatric patients who have the ability to swallow whole capsules. This is in line with the NICE scope.

3.3 Comparators

As the CS evidence only includes trials with no control arms, the company adopted a pragmatic approach to identify comparator data for established management without entrectinib. The company conducted a search of NICE pathways to identify comparators approved by NICE for each of the tumour types represented in the CS efficacy evidence population, using tumour type search terms. Where searches resulted in multiple possible pathways on the NICE Pathways website, the company made a decision on the pathway most relevant to the decision problem, for example the pathway referring to management of advanced/metastatic patients. Although the ERG understand that the company chose not to conduct separate systematic reviews to identify comparator data due to the large number of comparators and tumour types, the risk that relevant evidence may be been omitted cannot be excluded.

Therapies including chemotherapy, hormone therapy and best supportive care that had received a positive NICE recommendation were included. Excluded therapies were: surgery with curative intent, radiotherapy (non-palliative), immunotherapy, targeted agents and biological therapy. The clinical advisers to the ERG confirmed that these criteria are likely to be generally applicable for the majority of clinical scenarios.

Choices of lines of therapy by tumour histology were made by the company, and comparators were selected following current NICE recommendations. Median PFS and OS data were extracted from the clinical effectiveness data presented within the Committee slides, or where not available, from the company's submission. Trial participant characteristics, estimates of precision, or the committee-preferred parametric models used to extrapolate OS and PFS data were not extracted. Where chemotherapies were recommended by NICE but clinical evidence was not specifically available, the comparator data was not included. The company did not clarify which comparators and tumour types this criterion was applied to. The ERG believe this approach to be inappropriate and that a targeted systematic search for relevant evidence, and where possible, extracting data directly from survival curves for a better estimate of median PFS/OS and its variance, would have been preferable.

Where multiple PFS and OS values were available for a given comparator, the company extracted median values from primary analyses of individual trials informing the NICE TA analyses. Subgroup values were not used. The company stated that technology appraisals were informed primarily by one randomised controlled trial, and there was only one median value provided for each outcome that was relevant to the decision problem or the scope of the technology appraisal for the given comparator. However, the ERG found that in some instances, two different estimates where used for a single agent within the same line of therapy. For instance, for best supportive care for 2L+ thyroid cancer, figures adjusting and not adjusting for cross-over were both extracted.

In response to a clarification request, the company stated that comparator efficacy data was drawn from multiple technology appraisals for individual comparators in the same line of therapy where available (for example, docetaxel in NSCLC). They noted that this decision was taken to increase the robustness of the comparator data, by taking a mean of multiple values, and to ensure that an outlying or extreme value was not inadvertently used. The ERG found that in a number of instances (such as trifluridine-tipiracil for CRC, or eribulin for breast cancer) where more than one trial informed the TA, or where subgroups where combined (e.g. everolimus and best supportive care for neuroendocrine tumours in different sites) using inputs from robust meta-analytical techniques (for instance, as reported in company submissions or conducted by ERGs) would have been preferable to naively pooling unweighted means of medians, which is statistically inappropriate.

Where no chemotherapies were recommended by NICE, no additional targeted systematic reviews were conducted. Instead, the company used one of two approaches to identify relevant comparator data. In the case of MASC, surrogate trial data for best supportive care was used to derive OS data. However, it is not clear how this trial data was identified. Other tumour types for which only one patient was included in the efficacy population (cholangiocarcinoma, endometroid and ovarian cancers) were grouped into a single "other" category. PFS and OS estimates were derived for this category by calculating an average of PFS and OS median estimates from comparator data selected for the other tumour types (colorectal cancer, NSCLC, breast cancer, soft tissue sarcoma, pancreatic cancer, neuroendocrine tumours and thyroid cancer). The same method was used to derive PFS comparator data for MASC, reportedly due to lack of evidence. The ERG finds this method inappropriate, as prognosis for patients with a given tumour type such as cholangiocarcinoma, gynaecological cancers or MASC may differ significantly from patients with other unrelated tumour types.

The final choice of comparators was validated by clinical advice for seven of the ten cancer types included in the entrectinib efficacy evaluable population. These include: colorectal cancer, NSCLC, breast cancer, soft tissue sarcoma, pancreatic cancer, neuroendocrine tumours and thyroid cancer. The company did not state whether PFS and OS values were validated by clinical advice for salivary gland cancer (including MASC), cholangiocarcinoma, and gynaecological cancers. It is not clear whether comparator data provided in response to clarification was clinically validated.

Extracted comparator data is presented in table 30 of the CS appendix, and in the company model. The ERG identified some discrepancies between the two sources, which the company addressed in response to clarification. Table 9 summarises PFS and OS values for comparator data selected by the company that informed the economic model. These also include ORR values extracted by the company. This shows that all tumour types except one (MASC) were assigned at least two comparators across multiple lines of therapy. For each cancer type, individual median PFS and OS

estimates from each comparator treatment were pooled to calculate an unweighted mean PFS and OS. This method implies there is an even distribution of patients receiving each of the therapeutic options within a given line of therapy, which does not reflect clinical practice. For instance, due to its toxicity, irinotecan is less likely to be administered than FOLFIRI as second line therapy for CRC, therefore these therapies should not be given the same weight when pooling estimates across comparators. Similarly, this approach assumes there is an equal distribution of patients receiving different lines of therapy within a tumour type, which is not reflective of clinical practice. For instance, doxorubicin and trabectedin, respectively ≥1L and ≥2L treatments for soft tissue sarcoma, were given equal weight to generate an average value of all comparator efficacy data for this tumour type. This approach also did not take into account the distribution of the underlying data including heterogeneity in prognosis factors across different comparators within each tumour type, or estimates of variance and precision in survival estimates, and is therefore invalid.

Median and mean survival estimates were then applied at an individual patient level to calculate an overall mean PFS and OS across all tumour types, weighted by the number of individual patients with each tumour type in the integrated efficacy population. These estimates were used to inform naïve, unadjusted comparisons with entrectinib efficacy data. These comparisons do not account for any potentially important patient characteristics, such as age, performance status, *NTRK* fusion status or the prevalence of CNS metastases.

Following request for clarification, the company provided comparator data for the following tumour types: primary CNS (adults and paediatric), infantile fibrosarcoma and malignant melanoma (Table 10).

Table 9 Selected comparator data (from CS, entrectinib Roche model, Inputs for SoC *NTRK*+)

Tumour type	Therapy	Entrectinib proposed line of therapy (from CS table 6)	Comparator line of systemic therapy		ORR	Median PFS (months)	Median OS (months)	Reference	Average Estimation	
									PFS	OS
			CS appendix table 30	ERG extracted from trial data						
Non Small-cell Lung Cancer	Docetaxel	≥2L	≥2L	≥2L	NR	3.3	8.7	Average of values from NICE TAs 520, 428, 483, 484, 403, 347, 124		
	Docetaxel + nintedanib		≥2L	2L	4.7	4.2	12.6	NICE TA347	3.75	10.65
Colorectal Carcinoma	FOLFIRI	≥2L	2L	2L	11.1	4.7	12.1	NICE TA307		
	Irinotecan		2L	2L	34.8	4.4	14.3	NICE Guideline CG121 - Kim et al 2009		
	Trifluridine-tipiracil		≥3L	≥3L	0.9	2	9	NICE TA405		
	Trifluridine/Tipiracil		≥3L	≥3L	1.6	2	7.2	NICE TA405		
	Best supportive care		≥3L	≥3L	0.0	1	6.6	NICE TA405		
	Best supportive care		≥3L	≥3L	0.0	1.7	5.2	NICE TA405	2.63	9.07

Breast cancer incl. secretory breast	Capecitabine	≥2L	2L	1 to 3L	11.5	4.1	14.5	NICE TA515		
	Eribulin		≥3L	1 to 6L	12.2	3.6	13.2	NICE TA423		
	Vinorelbine		≥3L	NR	4.7	2.2	10.5	NICE TA423		
	Gemcitabine + paclitaxel		≥3L	NR	4.7	2.2	10.5	NICE TA423	3.03	12.18
Salivary Gland Cancer (incl. MASC)	Best supportive care (Platinum+Gemcitabine data used as surrogate)	1L	≥1L	1-3L	NR	4.3	13.8	Surrogate data for BSC - Laurie et al. 2010	4.35	13.80
Soft Tissue Sarcoma	Doxorubicin	1L	≥1L	≥1L	7.5	4.1	14.7	NICE TA465		
	Trabectedin		≥2L	≥2L	5.1	3.7	13.9	NICE TA185	3.90	14.30
Pancreatic	Gemcitabine + nab-paclitaxel	1L	≥1L	1L	23	5.5	8.7	NICE TA476		
	Gemcitabine		≥1L	1L	7	3.7	6.6	NICE TA476		
	FOLFIRINOX		≥1L	1L	31.6	6.4	11.1	NICE Guideline NG85 - Conroy et al 2011	5.20	8.80

Thyroid (papillary), unsuitable/refractory to radioactive iodine	Best supportive care	≥2L	≥2L	1 to 2	1.5	3.7	19.1 (after cross-over adjustment)	NICE TA535 (Cross-over adjusted value from Guo et al 2015)		
	Best supportive care		≥2L	1 to 2	0.5	5.4	42.8	NICE TA535	4.55	30.95
Neuroendocrine tumours	Everolimus (pancreatic)	≥2L	≥2L	≥1L	4.8	11	44.02	NICE TAs 449 and 539		
	Everolimus (GI & lung)		≥2L	≥1L	2	11	37.16	NICE TAs 449 and 539		
	Best supportive care		≥2L	≥1L	2	4.6	37.68	NICE TAs 449 and 539		
	Best supportive care		≥2L	≥1L	1	5.5	39.56	NICE TAs 449 and 539	8.025	39.605
Others*	[redacted]					4.6#	17.2^			

*Cholangiocarcinoma, uterine and ovarian ; #unweighted average of all PFS estimates except MASC. ^unweighted average of all OS estimates.
ORR: objective response rate; OS: overall survival; PFS: progression free survival; GI: gastrointestinal; TA: technology assessment

Table 10 Selected comparator data (from company clarification, entrectinib Roche model, Inputs for SoC NTRK+)

	Treatment	Entrectinib proposed line of therapy (from CS table 6)	Line of systemic therapy		ORR	PFS	OS	Source	Average Estimation	
			CS appendix table 30	ERG extracted from individual trials					PFS	OS
High grade glioma (after surgery/radiotherapy)	Temozolomide	NR	2L	2L	5.4	2.89	7.34	TA23		
	Procarbazine, CCNU (lomustine) and vincristine	NR	2L	1L	NR	3.6	6.7	Brada M et al, 2010 35		
	Single agent CCNU (lomustine)	NR	2L	2L	14.4	3.0	9.8	Batchelor T et al, 2013 36	3.16	7.95
Infantile Fibrosarcoma (after surgery/chemotherapy)	Best supportive care	NR	2L	NA	NA	4.1 (average of known)	15.8 (average of known)	NA	4.11	15.85
Malignant melanoma	Dacarbazine	NR	2L+	1L	12.1	1.5	6.4	NICE guideline NG14 (Middleton MR et al, 2000) 37	1.5	6.4

Limitations in reporting meant that in some cases calculation of comparator data could not be replicated. For instance, reporting was insufficient to replicate calculation of average PFS and OS values for the large number of NSCLC comparators and TAs used.

The company noted that, given that *NTRK* fusion status and high prevalence of CNS metastases (20.4%) in the entrectinib trial population are not accounted for, the comparator OS and PFS values may be overestimated. The ERG found insufficient evidence to support this statement. There is insufficient evidence to suggest that *NTRK* fusions are associated with a worse prognosis for most of the tumour types presented in the trial efficacy population (see section 2.1). The prevalence of CNS metastases in the comparator data is also uncertain. Relevant participant characteristics from trials informing PFS and OS inputs were not extracted. Therefore, the ERG checked key participant characteristics of all trials informing the comparator data reported in the publications from the trials. Most comparator trial publications did not report whether CNS metastases were excluded, except for the following comparators: FOLFIRI and irinotecan for CRC, and best supportive care (BSC) for MASC. Other trials reported inclusion restrictions, and only included patients with treated/stable CNS metastases (capecitabine & eribulin for breast cancer, doxorubicin and trabectedin for sarcoma), similarly to STARTRK-2 (see Section 4.2.2). Most comparator trials did not report baseline prevalence of CNS metastases, except for NSCLC trials (ranging from 6% to 14%). Further details on comparator trial extracted by the ERG on population characteristics, end of life and survival distributions used in TAs, are reported in Appendix C.

The ERG found a mismatch between the lines of therapy used in the comparator data and those reported in the efficacy evaluable population for some tumour types. As reported in Table 8, just over a third of the company's trial participants received entrectinib as first line systemic therapy, and entrectinib was administered in treatment naïve patients in 10 of the 13 tumour types (all except cholangiocarcinoma, gynaecological, and paediatric melanoma) represented in the trial evidence. The company identified comparator data including treatment naïve patients in all of those 10 tumour types, with the exception of NSCLC and CRC.

The ERG also identified a mismatch between the company's proposed positioning of entrectinib (as reported in CS table 6) and the line of therapy in which the comparators were used in the trials identified by the company. Soft tissue sarcoma and MASC comparator trials included patients in second line therapy and beyond, although the company placed entrectinib as first line for these tumour types. Conversely, comparator trials included first line patients where entrectinib was positioned as 2nd line or beyond for the following tumour types: breast cancer, thyroid and neuroendocrine tumours.

In response to clarification, the company stated that their proposed lines of therapy by tumour type are provided in the appendix comparator table, which the ERG presents in Table 9. The ERG found a discrepancy between this source and the proposed positioning presented in CS table 6 for soft tissue sarcoma, the appendix table includes a second line comparator (trabectedin) while CS table 5 positions entrectinib as first line for soft tissue sarcoma.

Adverse events of comparator therapies were not extracted, therefore the safety of entrectinib could not be compared to other relevant therapies. Utilities were extracted from TA documentation and are further discussed in Section 5.2.7.2.

The ERG consider the methods used to identify, select and combine comparator data to have a number of important limitations. Overall, the populations included in the comparator trials do not match the entrectinib efficacy population, notably due to likely limited prevalence of *NTRK* fusion in comparator evidence, and mismatch in lines of therapy within the treatment pathway. Comparisons were naïve and do not account for any potentially important prognostic factors, such as age, performance status, *NTRK* fusion status, prevalence of CNS metastases, or specific tumour mutations within each tumour type. In the base case analysis ,no attempts were made to adjust for differences in population characteristics between the entrectinib and comparator trial populations. Overall, the ERG conclude that the comparator data used to inform the company model is highly unreliable. Due to a high risk of confounding bias, comparisons with entrectinib are unlikely to be reliable. Alternative methods for addressing the uncertainty associated with the comparator evidence are presented in Section 5.2.6.1.

3.4 Outcomes

The outcomes presented in the CS include overall survival, progression free survival, overall response rate, duration of response, adverse effects of treatment and health-related quality of life. These match the outcomes specified in the NICE scope.

3.5 Other relevant factors

The CS provided analyses of ORR for the subgroups specified in the NICE scope (previous therapy and tumour type) alongside other subgroups the company considered relevant. Following a request from the ERG, the company provided IPD level data for ORR outcomes by tumour type and line of therapy, but not for PFS and OS.

The company proposed a data collection plan via the cancer drug fund (CDF). A Patient Access Scheme discount of [redacted] off the entrectinib list price has been agreed with NHS England. The CS did not identify any equality issues.

4 Clinical Effectiveness

This section contains a critique of the methods of the review of clinical effectiveness data, followed by a description and critique of the trials included in the review, including a summary of their quality and results, and the results of evidence syntheses performed by the company.

4.1 Critique of the company review methods

4.1.1 Searches

The company performed a systematic search for randomised controlled trial (RCT), non-randomised, and observational studies that investigated the efficacy and safety of entrectinib for the treatment of patients with *NTRK*-positive solid tumours.

The literature searches for each of the systematic literature reviews (SLRs) – clinical effectiveness, economic evaluation, health state utility value – were carried out "in parallel". Consequently one search was conducted of each resource/database that included each of these aspects. For the cost/resource use systematic literature review, separate searches were conducted. Overall, the ERG considers that the searches carried out were well conducted and reported and appropriate sources were used so the likelihood of relevant studies not being identified is low.

The databases used for the effectiveness review are reported as being MEDLINE (segments used were 1946 to present, Daily and In Process & Epub Ahead of Print), Embase, EconLit and EBM Reviews. The latter resource includes a range of other databases: Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Health Technology Assessment database, NHS Economic Evaluation Database, ACP Journal Club, Cochrane Clinical Answers and Cochrane Methodology Register. The search strategies used in each of the databases are fully reproduced in section D.1.1.3 of the CS and the date that they were conducted is given. The numbers of records retrieved matches the number given in the PRISMA diagram (CS page 25).

Additional searches of conference websites (ASCO, ECCO, ESMO,AACR) were conducted to identify potentially relevant posters and abstracts and the reference lists of identified studies were reviewed. Searches of the trials registers ClinicalTrials.gov and the WHO ICTRP were also conducted to find ongoing studies. HTA websites were scanned by the company to identify previous regulatory submissions, including NICE, SMC, AWMSG, PBAC, CADTH including pCODR.

Other searches were conducted using the Cost Effectiveness Analysis Registry, RePEc (EconPapers within Research Papers in Economics), www.euroqol.org, www.inahata.org, www.hta.org.ac.uk, ScHARRHUD utility database, University of York Centre for Reviews and Dissemination, and Google Scholar. No details are reported about the terms in the searches of these additional resources.

The strategy used in the Embase, MEDLINE and EBM Reviews databases consists of three sections i.e. 1) *NTRK* fusion 2) entrectinib OR comparators and 3) quality of life. Sets were combined to retrieve studies about: a) *NTRK* fusion and entrectinib or comparators or b) *NTRK* fusion and quality of life and c) comparators and quality of life. The strategy for the systematic review of resource use consists of terms for the condition (*NTRK* Fusion) combined with search terms for resource use/health care costs.

The overall structure of the strategy is appropriate and there are no errors in how the sets are combined. Neither are there any typographical errors within the search terms used. A validated search filter to identify HSUV was incorporated into the search strategy

The search strategy used for Embase, MEDLINE and EBM Reviews consists mainly of free text terms rather than a combination of thesaurus and free text terms. This broad approach can be successful when seeking to identify studies that are available in Embase as conference abstracts as they have less detailed indexing applied to the database record. The Embase search strategy did not include the EMTREE heading "protein tyrosine kinase inhibitor" although after testing it was clear that this did not change the overall numbers retrieved.

The search of EconLIT appropriately was much broader than that conducted in Embase and MEDLINE, consisting solely of terms for *NTRK*, TRK fusion combined with terms for the intervention and comparators. No information is given about how the website searches were conducted

4.1.2 Inclusion criteria

The company provided full details of the inclusion criteria used in the systematic literature review in Table 1 of CS Appendix D. Studies were single-screened for inclusion and independently checked by a second reviewer, with any discrepancies resolved through consensus. Studies were eligible for inclusion if they recruited patients with *NTRK*-positive solid tumours to prospective RCT (stage 2-4), non-randomised, or retrospective/prospective observational cohort studies. Studies evaluating the efficacy, safety, and/or HRQoL associated with entrectinib and a number of comparators (e.g. belizatinib, cabozantinib, larotrectinib, repotrectinib) were eligible for inclusion. While these interventions were included in the company's original SLR, as per the NICE scope only those studies assessing entrectinib were included in the main CS. Efficacy outcomes included overall survival, progression-free survival, time-to-progression, duration of response, time-to-response, and objective response rate. Safety outcomes of interest were any treatment-related adverse event and tolerability issues, i.e. dose reductions and interruptions, treatment discontinuation. Details of HRQoL and patient reported outcome measures administered as part of clinical trials were also included.

There were no restrictions by location and date; the primary focus of the review was on English language publications, or non-English language publications with an abstract in English.

4.1.3 Critique of data extraction

Data extraction was performed and reported adequately. Appendix D of the CS stated that data were extracted from the included studies by one reviewer, with all extracted data checked against the source document by a second reviewer.

The company's main submission presents detailed information about the included studies (ALKA, STARTRK-1, STARTRK-2, and STARTRK-NG), with a summary of the methods, participant characteristics, and results presented in Table 7 and Table 8 of the CS (pages 31-38). The company provided further patient characteristic and efficacy data in their clarification responses at the ERG's request.

4.1.4 Quality assessment

The company did not present a formal quality assessment of any of the studies included in the SLR, reasoning that as they were not primary full publications there was insufficient available evidence to adequately assess the quality of the study.

The ERG did not consider the company's lack of quality assessment appropriate, particularly given the availability of evidence to the company about their own trials. Therefore the ERG conducted its own quality assessment on STARTRK-2, the primary source of clinical data used in the company's analysis, based on the Downs and Black checklist[38] (see Section 4.2.2), which assessed quality of reporting, external validity, internal validity, confounding, and study power.

4.1.5 Evidence synthesis

The four included studies enrolled a total of 357 patients; however, as the majority of patients recruited to these trials were not *NTRK* fusion-positive, the company pooled across the trials the patients who met the following criteria:

- Had at least 6 months follow-up
- Had *NTRK* gene fusion positive tumours
- Received at least 1 dose of entrectinib
- Had not been previously treated with a Trk inhibitor

This population is referred to by the company as the *NTRK* Efficacy Population (n=62), from which further *post hoc* exclusions were made; six patients with primary CNS tumours, one paediatric patient, and one patient with non-measurable disease. The resulting patient group is the '*NTRK* Efficacy Evaluable Analysis Set' (EEA) (n=54), which forms the basis of the company's efficacy analyses and

is the population upon which the economic model is focused. Following the ERG's request, the company provided updated analyses which included the 5 primary CNS and 7 paediatric patients for whom outcomes were available.

4.2 Critique of trials of the technology of interest, their analysis and interpretation

Four non-randomised single-arm Phase I/II basket studies were included in the company's 'integrated efficacy analysis set', which was a *post hoc* pooling of participants designed to maximise the number of patients included in the analysis. These studies investigated the efficacy and safety of entrectinib in adult patients (ALKA, STARTRK-1/-2), and paediatric/adolescent patients (STARTRK-NG) with tumours positive for *NTRK1/2/3, ROS1,* or *ALK* molecular alterations. The STARTRK-2 trial contributed 51 of the 54 patients included in the company's efficacy analyses, therefore this study, and the pooled integrated efficacy dataset forms the focus of the following section.

4.2.1 Design and analysis of basket trials

A common approach to evidence generation in histology independent cancer therapeutics is the basket trial. In the context of the present appraisal, a basket trial designed to address the decision problem would evaluate a single drug targeting a single mutation (i.e. *NTRK* gene fusion) in multiple disease cohorts defined by histology or tumour type. Typically, two-stage studies are designed to recruit a certain number of patients to each 'basket', and if a pre-specified proportion of patients in a particular basket respond, then recruitment is expanded within this disease area. If too few responses are observed within a basket then recruitment is stopped due to low promise of efficacy.

While the company cites FDA, EMA, and EUnetHTA opinion stating that basket trials are acceptable for HTA of tumour agnostic therapies, the basket study (STARTRK-2) comprising the majority of the company's evidence submission was not designed as a basket trial in the sense intended by these bodies. The baskets in this study were based on molecular targets (*ALK*, *ROS1*, *NTRK*) rather than tumour type for each molecular target. Therefore assumptions underpinning the analysis of a basket trial may not hold for the analysis of *post hoc* subgroups within the *NTRK* fusion positive basket.

Heterogeneity of response across baskets is an important issue in the design and analysis of conventional basket trials, and in this case extra care must be taken to accommodate the potentially large variation and imprecision in response rate estimates introduced by very small sample sizes. There are a number of possible analytical approaches; one method is to analyse each basket separately as though it were an independent study. However, this approach does not allow for the possibility that some populations may respond in a homogeneous way, which is plausible given the common molecular target. The approach taken in the company's analysis was to assume equal efficacy across all baskets and to generate a pooled response estimate, but in doing so reject the potential for heterogeneity of response across baskets. A third approach assumes similar efficacy across baskets,

with the different histologies not determining a particular ordering of effectiveness *a priori*, i.e. the baskets are exchangeable and a Bayesian hierarchical model (BHM) can be used.[39] This type of design acknowledges the heterogeneity of response across baskets by assuming that response rates are exchangeable, rather than equal across baskets. This allows borrowing of information on the probabilities of response across baskets and increases precision of estimates, whilst reducing the chances of obtaining extreme estimates in specific baskets with few patients. Alternative forms of BHM have also been proposed, which allows borrowing of information across similar baskets while avoiding optimistic borrowing from extreme baskets.[40]

However, this hierarchical approach may increase uncertainty unnecessarily. Therefore, when there is a strong rationale for expecting a uniform level of response it may be preferable to use a simple pooling of information across subgroups as in the CS.[41] However, the company did not state any reasons to expect homogeneity of response across tumour types *a priori*, and indeed previous basket trials have shown heterogeneity in effectiveness of chemotherapeutic agents across tumour types. A recent trial of vemurafenib in 122 patients with *BRAF* V600–mutated cancers across multiple tumour types (including CRC, NSCLC, Erdheim–Chester disease and Langerhans'-cell histiocytosis, primary brain tumours, cholangiocarcinoma, anaplastic thyroid cancer) found evidence of response in some tumour types including NSCLC and Erdheim–Chester disease and Langerhans'-cell histiocytosis, but not in colorectal cancer.[42] A trial of imatinib, a tyrosine kinase inhibitor, that included 196 patients across 40 different subtypes, found evidence of activity of imatinib in only five malignancies.[43] Another basket trial of imatinib in 10 histologic subtypes of advanced sarcoma concluded that although rare dramatic responses were seen, imatinib was not an active agent in these subtypes, although it had previously shown effectiveness in another subtype of soft tissue sarcoma, gastrointestinal stromal tumour.[44]

Thus, it does not seem reasonable to make an assumption of homogeneity across tumour types, given the variability of the entrectinib trial results in the absence of a plausible clinical argument. Furthermore, as the included trials were not designed or sufficiently powered to test the assumption of heterogeneity of response across subgroups, the ERG consider it inappropriate and overly optimistic to assume equal response independent of tumour histology.

The ERG explores the effect of heterogeneity of response across tumour types using BHM in section 4.3.1.

4.2.2 STARTRK-2

STARTRK-2 is an ongoing multicentre, single arm, open-label, phase II basket study of entrectinib in patients aged ≥18 years with solid tumours harbouring *NTRK1/2/3, ROS1*, or *ALK* molecular alterations. As discussed in Section 4.2.1, the baskets in this study were based upon the three types of

genetic alteration, and not on tumour type. A total of 206 patients were enrolled and treated with entrectinib, with 63 enrolled to the *NTRK* basket. Patients were recruited across 84 sites in 15 countries, including 3 centres in the UK. Fifty-one patients met the criteria for inclusion (i.e. >6 months follow up and measurable disease at baseline) and thus formed the main part of the *NTRK* efficacy evaluable analysis set. The study design of STARTRK-2 is summarised below in Table 11, and eligibility criteria are summarised in Table 12. STARTRK-2 is a non-randomised, uncontrolled, open-label trial. Only a small subgroup (*NTRK*-fusion positive patients) of the trial informed the submission. Therefore, the evidence from this trial is considered at high risk of bias, and is not appropriately designed to assess the relative efficacy and safety of entrectinib against current established management. The trial eligibility criteria did not specify that inclusion into the trial was dependent on the lack of alternative effective and suitable standard therapy. As discussed in section 3.1, this may limit the extent to which the trial population matches the anticipated licence.

Table 11 Study design of STARTRK-2

Study details	
Location	84 sites in Australia, Belgium, France, Germany, Hong Kong, Italy, Japan, South Korea, The Netherlands, Poland, Singapore, Spain, Taiwan, United Kingdom, USA
Design	Non-randomised, one-arm, open-label
Duration of core study	4 years
Method of randomisation	None
Method of blinding	None
Intervention(s)	Entrectinib (RXDX-101)
Comparator(s)	None
Primary outcome	Objective Response Rate (ORR)
Data cut-off	35 months (Nov 2015 – [redacted])
Secondary outcomes	Duration of response (DOR), best overall response (BOR) time to response (TTR), progression free survival (PFS), safety outcomes.

Table 12 Eligibility criteria for STARTRK-2 (adapted from CS Table 8, Pages 34-37)

Inclusion criteria	**Exclusion criteria**

Age ≥18 Histologically- or cytologically-confirmed diagnosis of locally advanced or metastatic solid tumour that harbours an NTRK1/2/3, ROS1, or ALK gene rearrangement that is predicted to translate into a fusion protein with a functional TrkA/B/C, Ros1, or Alk kinase domain, respectively Measurable disease as assessed locally using RECIST v1.1 ECOG performance status ≤2 and minimum life expectancy of ≥4 weeks Prior anticancer therapy is allowed but 2 weeks must have elapsed following prior chemotherapy, and 4 weeks since completion of antibody-directed therapy Patients with CNS involvement, which is either asymptomatic or previously-treated and controlled.	Concomitant secondary oncodrivers (e.g., epidermal growth factor receptor, KRAS) Prior treatment with an approved or investigational TRK, ROS1, or ALK inhibitor in patients with tumours testing positive for the respective gene rearrangements Active gastrointestinal disease or other malabsorption syndromes

The ERG used the Downs and Black checklist to quality assess STARTRK-2 using the Interim Clinical Study Report provided by the company. This checklist scores the quality of reporting, external validity, internal validity, internal validity-confounding, and power of non-randomised trials. Results of the ERG's quality assessment using the Downs and Black checklist are presented in Appendix D. Overall, the ERG considers this trial to be at high risk of bias given that it is uncontrolled and only a fairly small subgroup of patients from this trial are included in the analysis.

4.2.3 ALKA

ALKA is an ongoing multicentre, single arm, open-label, phase I ascending dose and dose escalation study of entrectinib in patients aged ≥18 years with advanced/metastatic solid tumours with *NTRK1/2/3, ROS1*, or *ALK* molecular alterations. The primary objective of this study was to determine first cycle dose-limiting toxicity and the maximum tolerated dose of entrectinib. While only one patient from this study was included in the efficacy evaluable analysis dataset, a total of 57 patients were evaluable for safety outcomes.

4.2.4 STARTRK-1

STARTRK-1 is an ongoing multicentre, single arm, open-label, phase I ascending dose and dose escalation study of entrectinib in patients aged ≥18 years with solid tumours harbouring *NTRK1/2/3, ROS1*, or *ALK* molecular alterations. This study contributed two patients to the efficacy evaluable analysis dataset, and 76 patients were evaluable for safety outcomes.

4.2.5 STARTRK-NG

STARTRK-NG is an ongoing multicentre, single arm, open-label, phase I/Ib dose escalation and expansion study of entrectinib in patients aged 2-22 years with solid tumours harbouring *NTRK1/2/3, ROS1*, or *ALK* molecular alterations. At the time of the data cut used in the CS, only one patient from the STARTRK-NG trial had at least 6 months of follow-up, thus, no patients were included from this trial in the EEA dataset as originally presented by the company. Sixteen patients from this trial were analysed in the pooled safety population.

In response to a request by the ERG, the company provided a scenario analysis including 7 paediatric patients from the STARTRK-NG who had reached 6 months of follow-up by the latest data cut (█). This group included four patients with primary CNS tumours, two with sarcoma and one with malignant melanoma.

4.2.6 *NTRK* Efficacy Evaluable Analysis Set

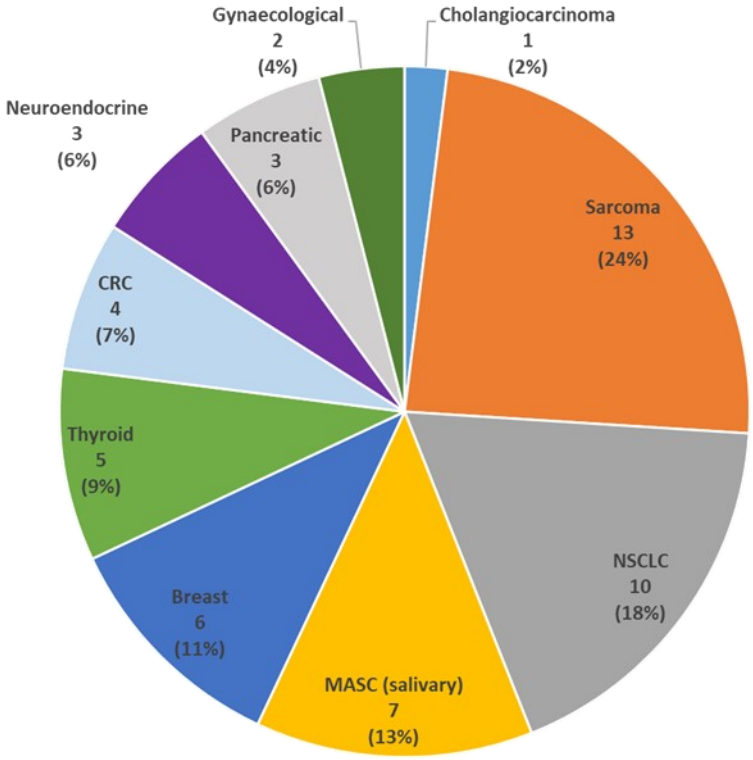
The primary source of efficacy data used in the company's efficacy analyses was based on the *NTRK* Efficacy Evaluable Analysis Set, which included patients derived from the four entrectinib trials who met the criteria described in Section 4.1.5, and excludes paediatric patients and those with primary CNS tumours. The company provided further analyses including 5 patients with primary CNS tumours and 7 paediatric patients in response to a request by the ERG (see section 4.2.6.1).

The demographics and baseline characteristics of the EEA population are summarised in Section 3.1 (Table 7).

As discussed in Section 3.1, there was a high degree of heterogeneity in the numbers of previous therapies received by patients within tumour types, making it likely that patients within subgroups were at different stages of the treatment pathway with varying disease history and prognosis. The ERG considered it unfeasible to conduct formal analyses on such data due to very small numbers of patients with such a high degree of heterogeneity in characteristics and response. The effect of heterogeneity across different tumour types is further explored in Section 4.3.

Figure 1 illustrates the distribution of cancer types included in the EEA dataset. The most represented solid tumour types were 'sarcomas' (24.1%), NSCLC (18.5%), salivary gland tumours (13.0%), and breast cancer (11.1%). The majority of patients had gene fusions involving *NTRK1* (40.7%), and *NTRK3* (57.4%), while only one patient was included with an *NTRK2* gene fusion. The company suggested in their clarification response that this was simply due to a lower absolute prevalence of *NTRK2* gene fusions, which comprise only █ of *NTRK* fusions. However, the ERG were concerned that the distribution of tumour types and gene fusion types in this patient population did not closely match or represent that expected in the NHS population, an issue discussed further in Section 3.1.

Figure 1 Tumour types included in the EEA dataset (n=54) (CS Fig. 7, Page 46)



CRC, colorectal cancer; MASC, mammary analogue secretory carcinoma; NSCLC, non-small cell lung cancer

4.2.6.1 Summary of clinical efficacy results

This section presents a critical summary of efficacy results presented by the company for the *NTRK* fusion trial population, including the EEA dataset, CNS primary and paediatric population.

Overall, the trial evidence showed a clinically meaningful objective response rate (ORR) ([redacted] in EEA dataset, CNS primary and paediatric population) including in patients with CNS metastases at baseline. However, there is considerable uncertainty regarding the extent to which the high response rates observed translate into clinically meaningful survival benefits. The ERG identified a number of important issues, particularly due to the significant immaturity of the PFS and OS data.

Despite substantial censoring and the small number of patients at risk in the tails of the Kaplan-Meier curve, the crossing of OS curves between entrectinib responders and non-responders is of some concern. Due to limited data there is considerable uncertainty about the precision of response and survival benefit estimates and heterogeneity by tumour type and line of therapy. Due to the lack of control group in the entrectinib trial evidence, the relative clinical benefits of entrectinib compared with relevant alternative cancer therapies are highly uncertain.

Response rate

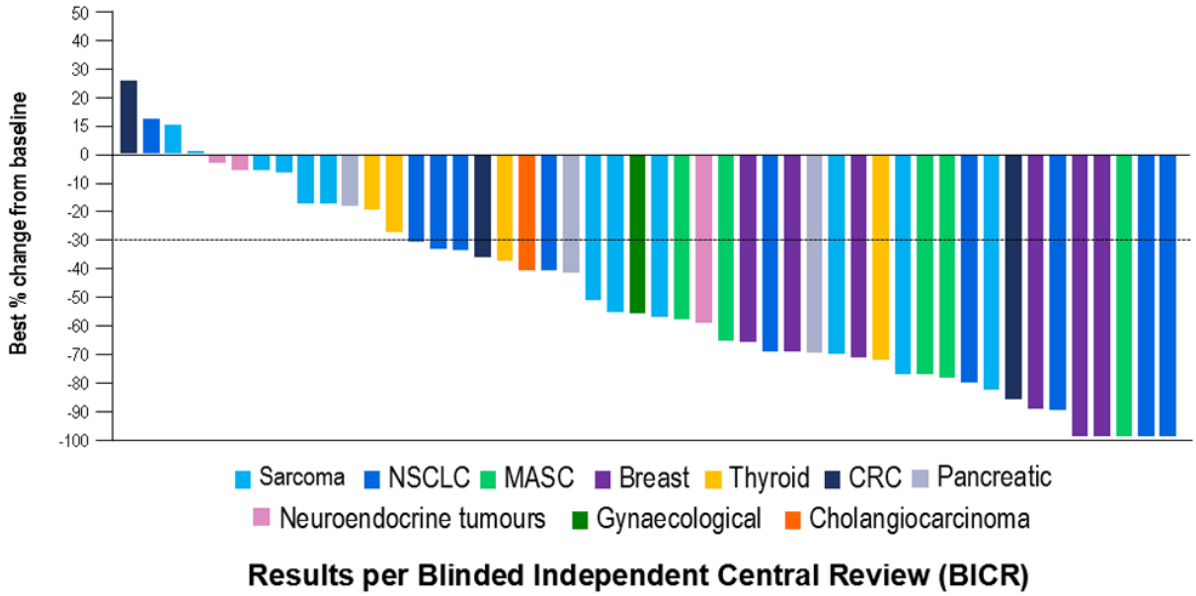
Objective response was assessed according to RECIST 1.1, assessed independently by blinded independent central review (BICR, primary analyses) and by the investigator. Response was defined as partial response (PR) or complete response (CR) confirmed by repeat imaging at least 28 days following first documentation of response. Table 13 presents the overall response rate (ORR) and best overall response for the efficacy evaluable population (data cut-off [redacted] BICR). Objective response was achieved in a high proportion of patients (57.4%, 95% CI: 43.2% to 70.8%). [redacted] achieved CR and [redacted] had a PR. Disease progression was found in four patients (7.4%). Investigator-assessed response rate estimates were consistent with the BICR (53.7%, 95% CI: 39.6% to 67.4%).

Table 13 Objective response rate and best overall response (efficacy evaluable population, data cut-off [redacted], from CS table 13)

	N (% of 54)
Responders	[redacted]
95% CI for response rates	[redacted]
Non-responders	[redacted]
Complete response (CR)	[redacted]
Partial response (PR)	[redacted]
Stable disease (SD)	[redacted]
Progressive disease (PD)	[redacted]
Non-CR/PD	[redacted]
Missing or unevaluable	[redacted]

Figure 2 presents individual patient responses measured as best percentage change from baseline in sum of longest tumour diameter (SLD) for the efficacy evaluable population. The 30% line of best percentage change corresponds to the RECIST 1.1 definition of partial response. Six patients from the efficacy evaluable population had missing SLD % change and were excluded from this plot. The company stated that response was observed across tumour types. Although the ERG agrees with this statement, no clear trend in response by tumour type can be inferred from visual inspection of this plot due to the small sample size and large number of subgroups.

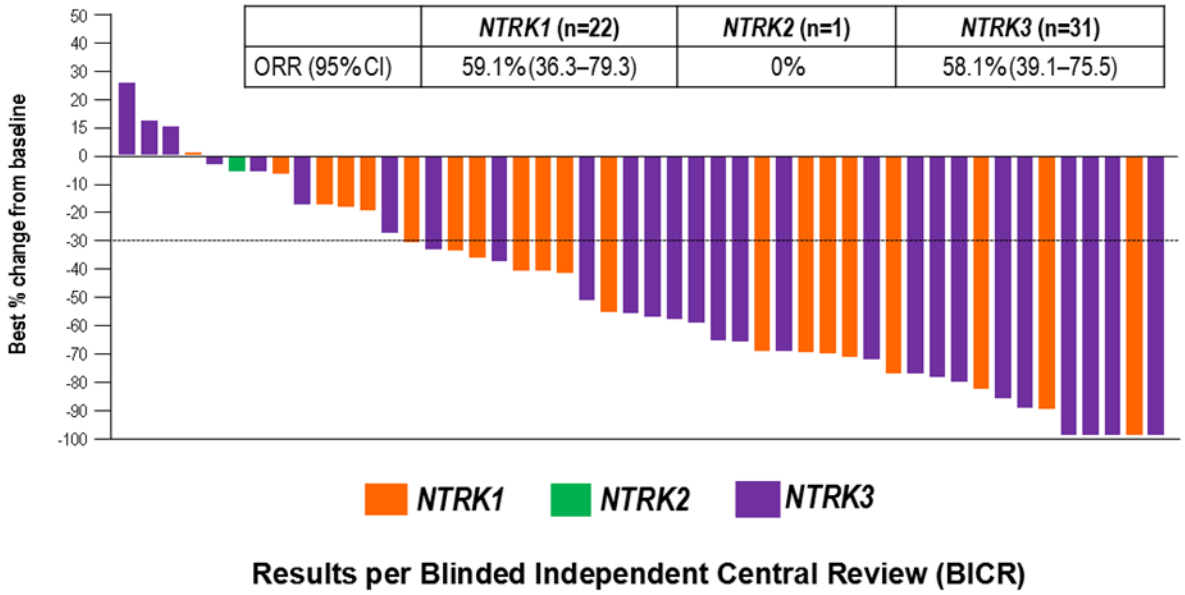
Figure 2 Entrectinib activity in *NTRK* fusion-positive solid tumours: individual patient responses by tumour type, BICR assessment - data cut-off 31 May 2018 (efficacy evaluable analysis, N=48*, from CS Figure 8)



*6 had missing SLD % change

Figure 3 presents individual patients' response for the efficacy evaluable population. This shows similar ORR for patients with *NTRK1* fusion (59.1%; 95% CI 36.3-79.3) and *NTRK3* (58.1%, 95% CI 39.1-75.5). The only patient with *NTRK2* fusion did not respond. The company stated that responses were independent of the *NTRK* fusion gene. The ERG believes this interpretation to be highly uncertain due to the lack of evidence for *NTRK2* fusions and small size of the *NTRK1* and *3* subgroups.

Figure 3 Entrectinib activity in *NTRK* fusion-positive solid tumours: individual patient responses by *NTRK* gene, BICR assessment - data cut-off 31 May 2018 (efficacy evaluable analysis N=48*, from CS Figure 9)



*6 had missing SLD % change

Figure 2 and Figure 3 exclude CNS primary and paediatric patients. In response to an ERG request, the company provided individual patient response data by line of therapy and tumour type for 66 patients, including the EEA population, as well as five adult primary CNS tumours patients, and seven paediatric patients (four primary CNS tumours, two sarcoma and one malignant melanoma) from the [redacted] clinical cut-off date. For primary CNS tumours, response was measured according to different criteria (Response Assessment in Neuro-Oncology Criteria, RANO) than other included tumours (RECIST v1.1), therefore the ERG agree with the company these results should be interpreted with some caution.

Detailed results assessed by BICR and investigator are presented in tables 1 and 2 and figures 1 and 2, in the company's clarification response (entrectinib clarification response 04072019K). Table 14 summarises response rates by line of therapy for 66 patients for EEA population, adult primary CNS tumours patients, and paediatric patients. Response data were extracted from the Company's clarification response (entrectinib clarification response 04072019K). BIRC-assessed response data by tumour type based on the more recent [redacted] clinical cut-off date (Company's clarification response Table 1) were used.

In the adult primary CNS tumour population, investigator-assessed response data are available for the five patients. However, the BIRC data only include one primary CNS tumour patient, as BIRC data from the four STARTRK-2 adult primary CNS patients are not available. In the paediatric population,

only investigator-assessed response data are available. Investigator-assessed response rates are provided in the Company's clarification response Table 2 and were used to impute response status where BIRC-assessed response is missing. However, patients were not listed in the same order in Tables 1 and 2 of the Company's clarification response, and no patient identification numbers were provided. Therefore, imputation was done by carefully matching the missing patients by tumour type and line of therapy, checking that a value for the correct patient was imputed. The ERG is confident that imputation was adequate. Table 14 shows high response rates across first, second and third line therapy and beyond. ORR and CR rate were higher in patients receiving entrectinib as 1st line therapy [redacted] than as 2nd line [redacted] and third or subsequent line [redacted], although these findings are based on small subgroups.

Table 14 Response rates by and line of therapy (data cut-off [redacted]

	[redacted]	[redacted]	[redacted]	[redacted]
Response	[redacted]	[redacted]	[redacted]	[redacted]
CR	[redacted]	[redacted]	[redacted]	[redacted]
PR	[redacted]	[redacted]	[redacted]	[redacted]
No response	[redacted]	[redacted]	[redacted]	[redacted]

[#] if BICR was missing, data imputed from investigator assessment where possible. * Missing patients treated as non-responders

ORRs were [redacted] and [redacted] in patients with and without baseline CNS metastases, respectively (cut off [redacted]). As above, these results may not be reliable due to the small number of patients in each subgroup as reflected in the wide confidence intervals (11 patients had CNS metastases at baseline) and the exclusion of CNS primary patients from this subgroup analysis.

Duration of response

Table 15 shows that responses in the efficacy evaluable population were durable with a median DOR of [redacted] among the 32 responders, although this was subject to significant censoring, as [redacted] of the 32 responders had an event at the [redacted] cut-off.

Table 15 Duration of response, BICR assessment (efficacy evaluable analysis), data cut-off [redacted]

Pts included in analysis (Responders)	[redacted]
Pts with event (%)	[redacted]
Progressive Disease	[redacted]
Death	[redacted]
Median	[redacted]
95% CI for Median	[redacted]
25% and 75%-ile	[redacted]

Range	[redacted]

* Subject to censoring. NE: not estimable

KM data were also reported for earlier cut-off of 31st May 2018. A swimmer plot for the 31 responses in the *NTRK* efficacy evaluable analysis set is shown in CS Figure 11, although once again this was subject to significant censoring (15/31 censored).

Progression free survival

Table 16 and Figure 4 present Kaplan-Meier analyses results for PFS based on the BICR assessment in the EEA population (data cut-off [redacted]). The estimated median PFS was [redacted]. The ERG note that these results are subject to significant censoring, as only [redacted] patients had an event. In addition, these results only apply to the EEA population, and do not account for heterogeneity across tumour types. In response to a clarification request, the company stated they were not able to provide PFS data stratified by line of therapy or tumour type.

Table 16 Progression-free survival BICR assessment (efficacy evaluable analysis) – data cut-off [redacted]

	BICR-assessed PFS (n = 54)
Patients with event (%)	[redacted]
Progressive Disease	[redacted]
Death	[redacted]
Median PFS (95% CI)	[redacted]
25% and 75%-ile	[redacted]
Range	[redacted]

* Subject to censoring. NE: not estimable

Figure 4 Kaplan-Meier curve for BICR-assessed PFS (efficacy evaluable analysis) - data cut-[redacted] *Figure redacted*

The ERG requested further survival data stratified by response status from the company. PFS KM data was provided for the EEA with and without CNS primary and paediatric populations from the [redacted] cut-off. Table 16 and Figure 5 present KM results for the combined EAA, adult primary CNS and paediatric populations. This shows that [redacted] of these patients were responders, and [redacted] were non-responders. As expected, median PFS was higher in responders ([redacted] months) compared with non-responders ([redacted] months). The reliability of these results may be limited, notably due to the immaturity of the PFS data. From visual inspection of KM curves, approximately [redacted] of the trial population [redacted] were censored. In addition, these analyses are limited by the lack of adjustment for potential confounding factors between responders and non-responders, including differing baseline risk and use of subsequent therapy. However, in the

absence of direct comparator data, these provide a proxy for the potential magnitude of PFS benefits observed in entrectinib responders. Similar results were reported for the EAA population only.

Table 17 Progression-free survival BICR assessment (efficacy evaluable analysis + CNS primary adults and paediatric population) – data cut-off [redacted]

	[redacted]	[redacted]
Progressive Disease	[redacted]	[redacted]
Median PFS	[redacted]	[redacted]
95% CI	[redacted]	[redacted]

Figure 5: PFS Kaplan-Meier curves for responders vs non-responders - integrated analysis population plus primary CNS and paediatric patients data cut-off [redacted]

Figure redacted

Overall survival

Table 18 presents OS results for the EEA population with and without the CNS adult and paediatric population. At the cut-off date of [redacted], [redacted] had died. The KM estimated median OS for the total efficacy population was [redacted]. Although this is potentially clinically significant, these estimates are highly uncertain due to significant data immaturity. The extent to which OS is driven by the efficacy of subsequent therapies is also unclear. As discussed previously, [redacted] of the trial population received entrectinib as first line therapy, and [redacted] received subsequent cancer treatments. There is insufficient evidence to explore whether survival outcomes may have been greater in the first line population compared to patients further down the treatment pathway. The extent to which OS may vary by tumour type is also uncertain. In response to a clarification request, the company stated they were not able to provide OS data stratified by line of therapy or tumour type, but provided further OS data stratified by response status.

Table 18 Overall survival, BICR assessment (efficacy evaluable analysis set with and without CNS primary adults and paediatric population), [redacted] (from CS Table 22)

	Total n=54	Total n=66
Pts with event (%)	[redacted]	[redacted]
Median	[redacted]	[redacted]
95% CI for Median	[redacted]	[redacted]
25% and 75%-ile	[redacted]	[redacted]
Range	[redacted]	[redacted]

* subject to censoring

OS KM data was provided for the EEA with and without CNS primary and paediatric populations from the [redacted] cut-off. Table 19 and Figure 6 present KM results for the combined EAA, adults primary CNS and paediatric populations. Table 19 shows that at the [redacted] cut-off, the numbers of deaths recorded in responders and non-responders was [redacted] and [redacted] respectively. Median OS was [redacted] for responders, but was not reached in non-responders (survival was 50.1% at the last time point). The KM data and the survival curves presented in Figure 6 indicate that the OS benefit observed in responders ceases approximately at [redacted], at the point where the two survival curves cross. The ERG found this potentially concerning as it suggests there may be no long-term OS benefit for those who respond to entrectinib compared with those who do not. However, these OS data are immature. From visual inspection of KM curves, approximately [redacted] of the analysed population, including [redacted] of responders and [redacted] non-responders were censored; the survival curve for non-responders reaches a plateau at [redacted] at which point the estimated survival probability was still [redacted] The crossing of survival curves and substantial immaturity of the data mean that the longer-term OS benefit of entrectinib in this population is highly uncertain. KM data and survival curves were also reported for responder analyses excluding the adult CNS primary and paediatric population, and are presented in Figure 7 below. These data include [redacted] responders and [redacted] non-responders. These responder analyses [redacted] Again, this raises concerns about the true longer-term OS benefits of entrectinib in treatment responders, and emphasises the need for more mature survival data.

As discussed above, the lack of adjustment for potential baseline imbalances between responders and non-responders (as noted by the company) and other confounding factors including subsequent therapies means that these results may not be reliable. As reported in section 3.1, the company clarified that [redacted] of the EEA population received a wide range of subsequent cancer therapies, although these data are not broken down by response status. In addition, the responder analyses do not take into account the heterogeneity in survival by age, tumour type, line of therapy or presence of CNS metastases. The ERG agree with the company that uncertainties associated with these analyses are further compounded by the small number of patients, limited follow-up, exclusion of non-responders who died prior to outcome assessment, and the use of different definitions of response in CNS primary patients. However, a responder analysis approach avoids some of the significant limitations of the company's naïve comparison with external comparator data as discussed in section 3.3. In particular, all patients included in these analyses had an *NTRK* fusion, and the distribution of *NTRK1* and *NTRK3* fusions were similar between responders and non-responders in the EEA responder analysis (*NTRK* fusion subtypes were not reported for the adult primary CNS and paediatric populations).

The company noted that it cannot be assumed that entrectinib had no activity in patients classed as non-responders, and that in these patients tumour progression may have been temporarily halted or slowed down by treatment, thereby improving survival outcomes of the non-responder group. The ERG agrees that SLD reductions between 10% and 30% were observed for several patients across lines of therapy and tumour types (see response to clarification response 04072019K figure 1). Although it is theoretically possible that entrectinib may have improved survival results in patients not classed as responders according to the RECIST definition, the company did not provide evidence to support this, and the ERG believe this is unlikely to fully explain the positive survival outcomes observed in non-responders and the crossing of curves in the KM responder analyses. Therefore, the ERG believes there to be significant uncertainty about the longer-term survival benefits of entrectinib, regardless of response status or depth of response.

Table 19 OS BICR assessment (efficacy evaluable analysis + CNS primary adults and paediatric population) – data cut-off [redacted]

	[redacted]	[redacted]
Death	[redacted]	[redacted]
Median OS	[redacted]	[redacted]
95% CI	[redacted]	[redacted]
25% and 75%-ile	[redacted]	[redacted]

Figure 6 OS Kaplan-Meier curves for responders vs non-responders - integrated analysis population plus primary CNS and paediatric patients data cut-off [redacted]

Figure redacted

Figure 7 OS Kaplan-Meier curves for responders vs non-responders - EEA population (without primary CNS and paediatric patients) data cut-off [redacted] (from company follow-up clarification response 11072019KM, figure 2)

[redacted]

Figure redacted

To further explore uncertainty and heterogeneity in the survival estimates presented by the company, the ERG requested from the company individual patient data on PFS and OS by line of therapy and tumour type for the integrated analysis population, adult primary CNS tumours and paediatric populations. The company replied that this was not possible due to legal and governance reasons, although they noted they may be able to conduct further prospective analyses as required.

Health-related quality of life

HRQoL was evaluated for 51 of the 54 *NTRK* fusion patients included in the EEA population enrolled in the STARTRK-2 trial, and was assessed prior to the first dose of each cycle and at the end of treatment. The following questionnaires were administered: European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (QLQ-C30) and EuroQol-5 Dimension (EQ-5D). Nine NSCLC patients also completed the lung cancer module (QLQ-LC13) and three patients with metastatic colorectal cancer (mCRC) completed the colorectal cancer module (QLQ-CR29). Completion rate for QLQ questionnaire was reported to be ≥80% at most study visits. Results were reported in the interim STARTRK-2 CSR (31st May 2018 cut-off) in tables and narratively.

The QLQ-C30 assesses five functional domains (physical, role, cognitive, emotional, social) and an overall global health status score. Baseline functioning scores were moderate-to-high for QLQ-C30 for global health status (69.79), physical functioning (74.17), role functioning (67.01), and cognitive functioning (84.72) on a score ranging from 0 to 100 with 100 reflecting better functioning. While receiving entrectinib, mean GHS scores were generally maintained or improved (-4.17 to 9.72). Physical functioning and role functioning scales results were moderate to high, with a trend towards clinical improvement. Cognitive functioning showed a negative trend (worst mean change score of -11.11 at Cycle 20 Day 1). Further results for QLQ-LC13 and QLQ-CR29 were reported in the STARTRK-2 interim CSR. Results for emotional and social functioning were not reported. Results of the EQ-5D questionnaire are discussed in section 5.2.7.

The ERG generally agree with the company's interpretation of the quality of-life questionnaire results, which suggest that overall general health functioning is not signigicantly affected duing entrectinib treatment. However, a reduction in cognitive functioning was observed, and there was no evidence on emotional and social functionning specifically. As above the reliability of the reported results is limited due to the small sample size.

Adverse effects of treatment

The CS provided adverse events data for 355 patients from three ongoing adult studies: ALKA (n = 57), STARTRK1 (n = 76), STARTRK2 (n = 206)) and one paediatric trial STARTRK-NG (n = 16). Patient safety data from these four trials have been pooled and analysed as the 'integrated safety population', with a data cut-off of 31st May 2018. Patients included in the integrated safety population were followed up for at least 6 months. The results for patients with < 6 months at the 31st May 2018 cut-off are reported separately. The integrated safety population includes adult and paediatric patients with *NTRK*, *ROS1* and *ALK* as well as paediatric patients with other/no known gene fusions. Table 20 presents the demographic characteristics of the integrated safety analysis.

Table 20: Baseline characteristics of the integrated safety population (adapted from CS Table 36)

	Adult Patients (N = 339)	Paediatric Patients (N = 16)	*NTRK* Adult Patients (N = 68)	Integrated safety population (N=355)
Sex, n (%)				
Male	151 (44.5)	10 (62.5)	31 (45.6)	161 (45.4)
Female	188 (55.5)	6 (37.5)	37 (54.4)	194 (54.6)
Median age, years (range)	55.0 (15, 86)	9.5 (4, 20)	57.5 (21, 83)	55.0 (4–86)
Age group, years, n (%)				
<65	249 (73.5)	16 (100.0)	43 (63.2%)	265 (74.6)
≥65	90 (26.5)	0	25 (36.8%)	90 (25.4)
Race, n (%)				
Asian	82 (24.3)	3 (18.8%)	9 (13.2)	82 (23.2)
White	222 (65.7)	13 (81.3%)	52 (76.5)	235 (66.4)
Black of African American	13 (3.8)	0	1 (1.5)	16 (4.5)
Other	5 (3.6)	0	0	5 (1.4)
Not reported	6 (4.4)	0	6 (8.8)	16 (4.5)
Mean BSA, m2 (SD)	1.79 (0.26)	1.07 (1.07)	1.83 (0.28)	1.76 (0.30)
Mean BMI, kg/m2 (SD)	24.79 (5.17)	17.33 (4.45)	25.12 (5.63)	24.45 (5.36)
ECOG PS, n (%)				
0	140 (41.3%)	0	26 (38.2)	140 (41.3)
1	170 (50.1%)	0	33 (48.5)	170 (50.1)
2	25 (7.4%)	0	7 (10.3)	25 (7.4)
3	3 (0.9%)	0	2 (2.9)	3 (0.9)
4	1 (0.3%)	0	0	1 (0.3)
Metastatic disease at baseline, n (%)				
Any site	NR	12 (75)	NR	311 (87.6)
CNS lesions	NR	0	NR	138 (38.8)

A summary of adverse events reported in the integrated safety population are reported in Table 21.

[redacted]

Grade 3/4 AEs were reported for [redacted] of patients in the overall safety population. Treatment-related grade 3/4 AEs were found in [redacted] of the patients. Treatment-related serious adverse events were reported in [redacted] of the overall safety population. Deaths associated with adverse events were seen in [redacted] AEs leading to discontinuation of entrectinib were reported in [redacted] of the integrated safety population.

Table 21: All-causality and treatment Related Adverse Events (Integrated Safety Population, data cut-off [redacted], from CS Table 38, Page 80)

Adverse Event, No. Patients (%)	Adult Patients[a] (n=339)		Paediatric patients (n=16[c])		*NTRK* Fusion patients (n = 68)		All Patients (N=355)	
	All causality	Treatment-related[b]	All causality	Treatment-related[b]	All causaslity	Treatment-related[b]	All causality	Treatment-related[b]
Number of Patients:								
with AE	[redacted]							
with SAE	[redacted]							
with Grade ≥3 AE	[redacted]							
Adverse events associated with:								
Discontinuation	[redacted]							
Dose reduction	[redacted]							
Drug interruption	[redacted]							
Death	[redacted]							

AE, adverse event; SAE, serious AE

[a] Includes 68 NTRK, 134 ROS1 NSCLC and 137 other adult patients

[b] Treatment Related Adverse Events refer to adverse events that was considered by the investigator to be related to entrectinib treatment.

[c] Paediatric patients include 1 NTRK patient

[d] The ERG noted a small discrepancy in reporting of the proportion of all-cause Grade 3/4 adverse events (61.1% in the CS Table 38 and CS confidential docs 6.2.4.7 Table 9. 60.3% reported in CS page 81; CS confidential docs 6.2.4.7 and CS Appendix F. Table 19).

Table 22 presents adverse drug reactions (ADR) by organ class in the integrated safety population as reported in the Summary of Product Characteristics provided in the CS. [redacted]

Table 22: Adverse drug reactions, integrated safety population. (from CS Appendix C, Summary of Product Characteristics, Version 1, 10/2018. Table 5).

System Organ Class Adverse Reaction	All Grades (%)	Grade 3 – 4 (%)	Frequency Category (All Grades)
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
ALT: Alanine aminotransferase. AST: Aspartate aminotransferase. Very common: ≥1/10, common: ≥1/100 to <1/10			

The ERG agrees with the manufacturer that the rates of AEs were broadly similar between adult patients across *NTRK*, *ROS1* NSCLC and other *ROS/ALK* patients.

The most frequently reported all-causality adverse events in the *NTRK* adult population were similar to those seen in the total integrated safety population.

[redacted]

The Summary of Clinical Safety reports the frequency of patients who experienced weight gain (Table 23); [redacted]

Table 23: Patients experiencing treatment-related weight gain (Integrated Safety Population, data cut-off 31st May 2018, from Summary of Clinical Safety (Roche Confidential Docs, 6.2.7.4) Page 61)

Adverse Event	No. Patients (%) [redacted]
≥ 5% weight increase	[redacted]
10 to ≥ 20% weight increase	[redacted]
≥ 20% weight increase	[redacted]

Overall, the ERG found that adverse events were generally well reported in the CS. The company did not extract safety data for comparator included in the NICE scope (see section 3.3). Due to the lack of comparator data, the relative safety of entrectinib compared with established management is highly uncertain. Rarer adverse events may not have been identified due the relatively small size of the safety population, particularly in paediatric patients.

4.3 Additional work on clinical effectiveness undertaken by the ERG

4.3.1 Exploring heterogeneity in response rates across tumour types

The ERG considers the company's assumption that such a variety of tumour types will have identical response rates when treated with entrectinib to be very strong and, as yet, untested. Therefore. an analysis of the potential heterogeneity in response rates across tumour types represented in the EEA, adult CNS primary and paediatric populations, and the additional uncertainty around this potential variability, was conducted.

Regardless of how the STARTRK trials were originally designed and analysed (Section 4.2), we can consider each of the tumour types as a "basket" or group and analyse the response data using a Bayesian Hierarchical Modelling (BHM) framework[39] to explore the potential heterogeneity in effects across tumours. Although originally developed as an adaptive trial design with stopping rules for unpromising treatments, we can ignore the adaptive phase and use the method to estimate posterior probabilities of response for each tumour type, as well as a pooled posterior probability of response across all tumour types, accounting for the potential lack of uniformity of effect across tumours. An additional advantage of this type of model is the ability to predict the response probability that would be expected in a "new" tumour type (i.e. a tumour that is not represented in the trial data), which will give a measure of the uncertainty in the response rates in tumour types in the target population but for which no data are available (see Appendix A).

4.3.1.1 Methods

For the response outcome, data available for each of the tumour types in the integrated analysis population, plus primary CNS tumours and paediatric tumours, are the number of responders, x_j, out of the total number of patients, n_j for tumour type j, which are assumed to follow a binomial likelihood

$$x_j \sim \text{Binomial}(n_j, p_j)$$

where p_j is the probability of response for tumour type j, with $j = 1,...,G$, and G is the total number of tumour types. We model the log-odds of response in tumour type j, θ_j, on the log-odds scale: $\text{logit}(\theta_j) = p_j$. The BHM assumes that for each of the G tumour types, the log-odds of response, θ_j, are exchangeable and follow a Normal distribution

$$\theta_j \sim \text{Normal}\left(\mu, \sigma^2\right)$$

where σ is the standard deviation quantifying the between-tumour heterogeneity and μ is the pooled mean effect across all tumours. Prior distributions must be selected for μ and σ and are likely to have some influence on the posterior estimates,[39, 45] particularly when a small number of groups and patients per group are included.

A relatively conservative normal prior distribution for μ is used, centred around a probability of response of 0.3 (a log-odds of -0.8473) which is often considered as a promising response rate, with a variance of 10 across all tumour types. Sensitivity of results to a more favourable prior distribution where the prior probability of response across all tumour types is centred around a mean of 0.5 (i.e. a log-odds of 0) with the same variance.

The prior for the between-tumour heterogeneity standard deviation is specified as Uniform(0,5) which was found to be robust in a simulation study.[45] An Inverse Gamma(2, 20) prior distribution for the between-tumour variance had previously been proposed[39], which means the between-tumour precision has prior mean 0.10 and variance 0.005. Inverse-gamma prior distributions were found to lead to posterior distributions which are highly sensitive to the chosen parameters and are therefore not recommended in most cases.[45] For completeness we present the results obtained using this prior distribution for the base-case dataset in Appendix E.

We also calculate the probabilities that the response rate for each tumour type is at least 30% or at least 10%.

Because the tumour types included in the integrated analysis population, plus primary CNS tumours and paediatric tumours are not reflective of the full licensed indication (i.e. some tumour types are missing, see section 3.1), the predictive distribution for the response rate in a new tumour type is calculated to reflect the full degree of uncertainty both due to the sample size and the observed heterogeneity in effects across the observed tumours. The resulting distribution is the probability of response in a "new", i.e. unrepresented tumour type.

The model was adapted from Thall et al[39] and estimated using Markov chain Monte Carlo in OpenBUGS,[46] implemented in R[47] (version 3.6.0) using R2OpenBUGS[48] (version 3.2.3.2). Code and implementation details are presented in Appendix E.

Model fit was assessed by plotting individual tumour contributions to the residual deviance (in a well-fitting model these are expected to be close to 1) and by comparing the total residual deviance to the number of tumour types, G.

4.3.1.2 Description of included data

Response data were extracted from the Company's clarification response (entrectinib clarification response 04072019K). BIRC-assessed response data by tumour type based on the more recent [redacted] clinical cut-off date (Company's clarification response Table 1) were used. Where BICR data was missing, data were imputed where possible using investigator-assessed response data. Further details are reported in section 4.2.6.1. The number of patients and responses by tumour type are given in Table 24.

Table 24 Number of responders by tumour type ([redacted] clinical cut-off date, with imputed response status where BIRC-assessed response is missing)

Tumour ID	Tumour type	Number of patients (n)	Number of responders (x)
1	Sarcoma	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]
9	Breast	[redacted]	[redacted]
10	Thyroid	[redacted]	[redacted]
11	CNS Primary	[redacted]	[redacted]
12	Paediatric CNS Primary	[redacted]	[redacted]
13	Paediatric (non-CNS)	[redacted]	[redacted]
	Total	[redacted]	[redacted]

The company advised that caution should be exercised in the interpretation of response for CNS tumours as it is measured according to different criteria than for systemic solid tumours (Section 4.2.6.1). Whilst the ERG agrees with this advice, it is still valid to assess the heterogeneity in response across all included tumours regardless of how response is defined, as the overall response rate is an important clinical result.

4.3.1.3 Results

Results for the base-case analysis, which includes all adult and paediatric tumours (Table 24), are presented in this section. The prior distributions used for the base-case analysis are

$$\mu \sim \text{Normal}(-0.8473, 10)$$
$$\sigma \sim \text{Uniform}(0, 5) \quad (1)$$

The BHM estimates moderate between-group heterogeneity (posterior median [redacted], on the log-odds scale) although there is considerable uncertainty 95% credible interval (CrI) ([redacted]) (Figure 8). This suggests that there could be considerable variability across tumour types, although the possibility of very little variability is also not ruled out.

Figure 8 [redacted]

[redacted]*Figure redacted*

The estimated mean response rate across all tumour types is [redacted] with 95%CrI [redacted]. This is similar to the response rate that would be obtained if the tumour types were all assumed to have identical response probabilities [redacted] 95%CrI [redacted], which is consistent with the company's submission. The 95% CrI for the response probability predicted for an unrepresented tumour type is wide (Table 25, Figure 9), meaning that this probability could be as low as [redacted], or as high as [redacted].

Table 25 Probabilities of response according to the BHM.

	Overall posterior probability of response		
	mean	median	95% CrI
Posterior probability of response	[redacted]	[redacted]	[redacted]
Predictive probability of response	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3;
Uniform prior distribution for the between-tumour standard deviation

Figure 9 [redacted]

[redacted] *Figure redacted*

The estimated probabilities of response for each tumour type are shown in Table 26. The effect of allowing borrowing of information across the tumour types is to shrink the observed response probabilities towards the pooled mean response probability in Table 25. Tumour types with few patients borrow more information than tumour types with more patients.

Table 26 Probabilities of response for all tumour types

	Tumour type	Observed response (%)	Estimated mean response based on BHM (%)	Prob of response rate at least 30%	Prob of response rate at least 10%
1	Sarcoma	[redacted]	[redacted]	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]	[redacted]	[redacted]
9	Breast	[redacted]	[redacted]	[redacted]	[redacted]
10	Thyroid	[redacted]	[redacted]	[redacted]	[redacted]
11	CNS Primary	[redacted]	[redacted]	[redacted]	[redacted]
12	Paediatric CNS Primary	[redacted]	[redacted]	[redacted]	[redacted]
13	Paediatric (non-CNS)	[redacted]	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3; Uniform prior distribution for the between-tumour standard deviation

Figure 10 shows the posterior distributions of the probabilities of response for each of the 13 tumour types. Whilst all distributions overlap, the distributions of response for [redacted] and their 95% CrI (Appendix E) suggest that response rates [redacted] are plausible. These tumour types also have the lowest probabilities of having a response rate greater than 30% (Table 26).

Figure 10 Posterior distribution for the probabilities of response in each tumours type, including primary CNS and paediatric.

[redacted] *Figure redacted*

Three sensitivity analyses were carried out and presented in Appendix E.

1. To assess sensitivity of results to the inverse-gamma prior distribution for the between-tumour heterogeneity variance, as suggested by Thall *et al* [39]
2. To assess sensitivity of results to the use of a more favourable prior for the log-odds of response;
3. To assess sensitivity to excluding primary CNS and paediatric patients from the data.

4.3.2 Exploring heterogeneity in time-to-event outcomes across tumour types

Heterogeneity in time to event outcomes (PFS, OS) can be explored using the BHM in a similar way.[39] The model assumes a common parametric distribution for each tumour type, but with a different location parameter. Information on this parameter can be borrowed across the different tumours, according to an estimated heterogeneity parameter. The results from this type of model would be different distributions of PFS or OS for each tumour type which could be incorporated in the economic model in order to further explore how heterogeneity in outcomes by tumour type influences the expected ICERs.

Although the BHM can borrow information across tumour types, and is designed to allow inferences with few events per tumour type, it is unclear whether this type of model would provide useful results in this appraisal, given the immaturity of the survival data and the small number of patients in most tumour types. PFS and OS data were not available to the ERG by tumour type so the feasibility of this type of analysis could not be assessed. Nevertheless, as more data become available, this could be a useful way to determine the extent of heterogeneity in PFS and OS across the different tumour types, and would allow predictive distributions of PFS and OS to be used to inform the survival of patients with unrepresented tumour types.

4.4 Conclusions on clinical effectiveness

The CS efficacy evidence was supported by four uncontrolled basket trials that included a total of 66 patients with metastatic or locally advanced *NTRK* fusion positive solid tumours, including seven paediatric patients. Thirteen tumour types were included: sarcoma, NSCLC, MASC, breast, thyroid, CRC, neuroendocrine tumours, pancreatic cancer, gynaecological cancers, cholangiocarcinoma, CNS primary, infantile fibrosarcoma and paediatric melanoma. Each tumour type was represented by between one and 13 patients.

The ERG found that the intervention and outcomes presented in the CS evidence match the NICE scope. However, due to concerns about the large proportion of unrepresented tumour types, the under-representation of *NTRK2* patients, and the small sample size of the *NTRK* efficacy trial population, the ERG is concerned that the population presented in the evidence submitted is not representative of the population defined in the NICE final scope. A significant proportion (redacted) of trials patients received entrectinib as first line systemic therapy, and for some there appeared to have been 'acceptable' alternative standard therapies available.

The company adopted a pragmatic approach to identify PFS and OS comparator data for established management without entrectinib, by searching NICE pathways to identify NICE approved comparators for each of the tumour types represented in the CS efficacy evidence. Median PFS and OS from each tumour types were averaged and then pooled to calculate mean overall PFS and OS

across all tumour types, weighted by the prevalence of each tumour type within the trial population. The ERG found that the population included in the comparator trials is unlikely to match the entrectinib efficacy population, notably due to the unknown prevalence of *NTRK* fusion in most of the comparator evidence, and mismatch in lines of therapy within the treatment pathway. In the base case analysis, no attempts were made to adjust for differences in population characteristics between the entrectinib and comparator trial populations; comparisons were naïve and do not account for any potentially important prognostic factors. The ERG conclude that the methods used to identify, select and combine comparator data are inappropriate, and that the comparator data used to inform the company model is highly unreliable.

Overall, the trial evidence showed a clinically relevant overall response rate across tumour types. However, there is considerable uncertainty regarding the extent to which the response observed translate into clinically meaningful survival benefits. The ERG found a number of important issues, particularly due to the significant immaturity of the PFS and OS data. Despite substantial censoring and the small number of patients at risk in the tails of the Kaplan-Meier curve, the crossing of OS curves between entrectinib responders and non-responders is of some concern.

The ERG explored heterogeneity in response rates across tumour types using a Bayesian hierarchical model, which assumes the response probabilities are similar (i.e. exchangeable) across tumour types, rather than identical (the company's preferred assumption). The ERG's analyses found that response rates obtained were similar to those observed when equal response probabilities are assumed, although there was considerable uncertainty in the level of heterogeneity of response rates across tumour types. Therefore, the possibility that some tumour types could have response rates that differ significantly from the pooled [redacted] response rate cannot be excluded. Due to limited data there is considerable uncertainty about the precision of response and survival benefit estimates and heterogeneity by tumour type and line of therapy. The lack of control group in the entrectinib trial evidence means that the relative effectiveness and safety of entrectinib compared with relevant alternative cancer therapies are highly uncertain. Due to lack of appropriate data and the uncertainty in response rates, the efficacy of entrectinib in tumour types not represented in the company's trials is unknown.

5 Cost Effectiveness

This section focuses on the economic evidence submitted by the company and the additional information provided in response to the points for clarification. The submission was subject to a critical review on the basis of the company's report and by direct examination of the electronic model. The critical appraisal was conducted with the aid of a checklist to assess the quality of the economic evaluation and a narrative review to highlight key assumptions and uncertainties.

5.1 ERG comment on company's review of cost-effectiveness evidence

The CS describes the search strategies used to identify relevant cost-effectiveness studies for the treatment of patients with *NTRK* fusion–positive solid tumours. Full details of the search strategy used are provided in Appendix D of the CS.

5.1.1 Searches

The ERG considers the searches undertaken by the company to be appropriate. For details of the searches undertaken by the company, see Section 4.1.1.

5.1.2 Inclusion/exclusion criteria used for study selection

The inclusion/exclusion criteria are summarised in Table 27 (Appendix G) of the CS and follow the usual PICOS framework. In brief, the review included any economic analyses and systematic reviews of pharmacological treatments for patients with *NTRK* fusion–positive solid tumours. Articles were assessed by a single reviewer against each eligibility criteria and independently checked by a second reviewer. Any discrepancies between reviewers regarding the inclusion of studies were resolved by discussion.

The ERG considers that the inclusion/exclusion criteria applied were appropriate and likely to identify any relevant studies.

5.1.3 Studies included and excluded in the cost effectiveness review

A total of 2,645 studies were identified in the searches following de-duplication. Of these, 80 full text articles were screened for inclusion in the review. No studies were, however, found to meet the review inclusion criteria and as such no published evidence was identified on the cost-effectiveness of entrectinib. Supplemental searches conducted by the ERG also did not identify any studies on the cost-effectiveness of pharmacological treatments for patients with *NTRK* fusion–positive solid tumours.

5.1.4 Conclusions of the cost effectiveness review

In the absence of any previously published cost effectiveness studies in patients with *NTRK* fusion–positive solid tumours, the *de novo* analysis in the CS represents the most relevant evidence for the stated decision problem.

5.2 ERG's summary and critique of company's submitted economic evaluation

The company presented a *de novo* economic analysis comparing entrectinib with established management in 13 tumour types. The model estimates a single composite ICER considering cost-effectiveness across all 13 tumour types and does not attempt to estimate individual ICERs for each individual tumour types. A summary of the company's economic evaluation is presented in Table 27, with justifications for key aspects and signposts to the relevant sections of the CS. The ERG has considered the methods applied in the company's economic evaluation in the context of a detailed checklist, reported in Appendix F.

Table 27 Overview of the company's economic evaluation

	Approach	Source / Justification	Location in CS
Model	Cost-effectiveness (cost-utility) analysis: Entrectinib: A partitioned survival analysis (PartSA) approach. Established management: Based on a PartSA approach, but rather than deriving time in state from parametrically extrapolated curves fitted to KM data, derives mean time from extracted median PFS and OS estimates for comparator therapies.	PartSA approach allows modelling relevant outcomes (PFS, TTOT and OS). The model structure is stated to be in line with the NICE decision support unit guidance [49] Modelling of comparator data attempts to simulate the PFS and OS benefit of the comparator therapies.	Section B.3.2.2; p91-92
States and events	The PartSA model contains 3 states: progression free (PF), progressed disease (PD) and death.	The PartSA model health states partitions OS into states of interest pre and post progression, and on and off treatment.	Section B.3.2.2; p91
Comparators	Entrectinib was compared to established practice which was a pooled comparator consisting of chemotherapy regimens and BSC.	Established practice was modelled as a "simulated "chemotherapy comparator generated by averaging clinical outcomes derived from previous NICE appraisals. These were weighted according to the proportion of patients in the integrated analysis with each tumour type.	Section B.3.2.2; p94
Natural History	Based on partitioned survival model. Transitions for patients receiving entrectinib were based on the integrated analysis of the ALKA, STARTRK-1 and STARTRK-2 single arm trials. Transition for patients receiving established management were based on median PFS and OS data extracted from relevant NICE technology appraisals (TAs) and extrapolated assuming an exponential function.	PFS and OS estimates were modelled independently, with the proportion of progressed patients at each cycle, calculated as the difference between the OS and PFS curves. Comparator outcomes were based on mean PFS and OS.	Section B.3.2.2; p91
Treatment effectiveness	Clinical outcomes included PFS and OS. Entrectinib PFS and OS were extrapolated from pooled analysis of relevant patients from the ALKA, STARTRK-1 and STARTRK-2 trials using single standard parametric models.	In the base-case analysis, an uncontrolled and unadjusted comparison was established between the pooled patient level data from the three entrectinib single arm trials and pooled data from previous NICE appraisals used to estimate comparator outcomes.	Section B.3.2.2; p94 Section B.3.3.1; p94-97 Section B.3.3.2; p97-100 Section B.3.3.3; p100-103

	Approach	Source / Justification	Location in CS
	Comparator OS and PFS was extrapolated based on median PFS and OS extracted from relevant NICE appraisals assuming an exponential parametric survival function		
HRQoL	Pre-progression health state utilities for patients receiving entrectinib were estimated from EQ-5D-3L collected in the STARTRK-2 trial. Health state utilities in the comparator arm and post-progression health states for both treatment arms were based on a weighted average of utility values used in previous NICE TAs Utility decrements for adverse events were applied in scenario analysis only.	The health state utilities (PF and PD) were assumed to differ across treatment arms in the pre-progression health sate and assumed to be equal in the post progression health state. Scenario analysis was also undertaking assuming no utility difference between entrectinib and comparator arms of the model. EORTC and EQ-5D-3L data were collected at baseline; day 1 of each subsequent treatment cycle, and after treatment discontinuation. Questionnaires were also completed in the period after treatment discontinuation. Observed EQ-5D-3L responses were classified into three categories according to the patient treatment or progression status: 1) Base-line assessment (assessment prior to treatment start date) ([redacted]), 2) patients in PFS (after treatment star data but prior to disease progression ([redacted]), and 3) patients post-PFS (after IRC assessed progression) ([redacted]). Utility values were derived from the collected EQ-5D-3L values and assigned to the entrectinib pre-progression health states. Post progression data was not used as the reported mean was considered implausible. Pre- progression utilities for established practice and post progression values for both treatment arms were based on a weighted average of utility values reported in previous NICE TA's. These were weighted according to the proportion of patients in the integrated analysis with each individual tumour type. The specific source of individual utility values used in the model other than the source TA were not reported in CS. Utility decrements associated with adverse events relating to entrectinib and chemotherapy treatment were not included in the model as it was assumed that these were already captured in the trial-based utility values used. Scenario analysis was also undertaken incorporating additional disutilities associated with adverse events specifically associated with entrectinib treatment (weight gain).	Section B.3.4.1. p109-110 Section B.3.4.3. p111-112 Section B.3.4.4. p111-113

	Approach	Source / Justification	Location in CS
		All AE disutilities were applied as a one-off decrement applied to the first cycle of the model.	
Adverse events	Adverse events were included for grade 3/4 events occurring in ≥5% of subjects. Event rates were assumed to be identical for intervention and comparator arms with the exception of increase in weight which was assumed to only occur in patients receiving entrectinib.	Event rates were drawn from the integrated analysis of the ALKA, STARTRK-1 and STARTRK-2 trials. In the base-case analysis, the AE rates were applied in the model to estimate associated costs only. Scenario analysis was also presented in which rates were used to estimate treatment related disutilities.	Section B.3.5.3 p123
Resource use and costs	Cost categories were: • Treatment and administration costs • Subsequent therapy • Health state resource use and costs • Testing of *NTRK* status • AE costs	Drug and administration unit costs were sourced from BNF, and NHS reference costs. Resource use was informed by UK hospital chemotherapy protocols. A Patient Access Scheme (PAS) discount of [redacted] off the entrectinib list price has been agreed with NHS England. PAS are also available for a number of the comparator therapies, but are not included in the company's base-case analysis. To estimate health state resource use comparator therapies were classified into three categories (oral chemotherapy, single agent chemotherapy and combination therapy). Resource use for each category was drawn from recent TA's (TA515[50], TA520[51], TA476[52]). Unit costs were sourced from the Personal Social Services Research Unit (PSSRU) and NHS reference costs. Testing costs were based on current testing algorithms already used in practice and a proposed testing algorithm for *NTRK* fusions based on IHC followed by confirmatory NGS. Costs of individual tests were based on values used in TA406, costs cited in a Scottish science advisory council report and values elicited from five national genomic laboratory hubs. The costs of adverse events grade 3-4 with incidence ≥ 5% were included in the base-case. The cost of end of life care was included for the last cycle that patients were alive in the model and for both intervention and comparator.	Section B.3.5 p115-126

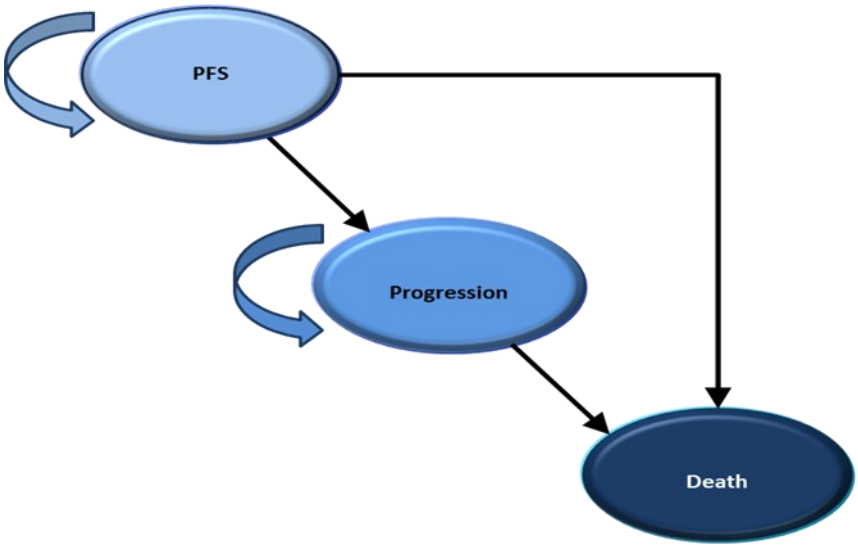
	Approach	**Source / Justification**	**Location in CS**
Discount rates	Costs and benefits were discounted at 3.5% per annum	In accordance with the NICE reference case.	Section B.3.2.2; p93
Sensitivity analysis	Probabilistic sensitivity analysis was performed. Deterministic univariate probabilistic analysis was performed on a series of model parameters. A series of scenario analyses was also performed.	In accordance with the NICE reference case.	B.3.8; p133-138

5.2.1 Model structure

The *de novo* analysis presented by the company compares entrectinib with established management. Established management consisted of a composite comparator represented through a weighted average of comparators from the tumour types represented in the integrated analysis for entrectinib, see Section 5.2.4

The cost-effectiveness analysis presented by the company is based on a partitioned survival model (PSM) or "area under the curve" analysis, depicted in Figure 11. It comprises three mutually exclusive health states: (i) PFS (progression free), (ii) progressive disease (PD; progression), and (iii) death. Within the PFS and PD health states, the model distinguished between patients who are receiving treatment and those who are not. The model predicted the total costs and QALYs separately for the entrectinib arm and the pooled comparator arm in order to estimate a single ICER.

Figure 11 Model structure (Figure 16 in CS)



Transitions between states are not explicitly incorporated into the analysis using probabilities. Instead the distribution of patients in each health state is determined by using estimates of PFS and OS.

For entrectinib, transitions were based on extrapolated KM data from the *NTRK* efficacy evaluable analysis set (Section 4.2.6). Time-to-event data for OS was used to determine the proportion of patients alive, while the proportion of patients in the PD state was calculated as the difference between OS and PFS. In scenario analysis time to off treatment (TTOT) was also used to determine time to discontinuation of treatment; in the base-case this was assumed to coincide with progression.

Transitions between states in the comparator arm were modelled using a different approach to that used to model the entrectinib arm. Time-to-event data were not used, instead estimates of mean OS

and PFS for each tumour type were modelled to estimate time in each health state. For each tumour type, time alive was estimated using mean OS, while time spent in the PD health state was estimated as the difference between the mean OS and mean PFS. These estimates of time in state were then used to estimate total costs and QALYs for each tumour type. Total costs and QALYs for the comparator arm were then estimated as weighted averages using the distribution of tumours in the integrated analysis of entrectinib (See Section 4.2.6) to determine the appropriate weight. As for the entrectinib arm, time on treatment for the comparator arm was assumed to align with PFS.

The cycle length used in the model was one week. Transitions between health states were assumed to occur at any time within the cycle. To account for the over- or under-estimation of transitions occurring at the beginning or end of the cycle, half-cycle corrections were applied to each time interval.

ERG comment

While the model structure is consistent with previous technology appraisals in advanced cancers, the ERG notes a number of issues, regarding the selection of endpoints from the clinical trials to define transitions between health states and the estimation of a single ICER for the full population covered by the marketing authorisation.

Choice of clinical endpoints to model health state transitions

As described above, the company's approach to estimating the cost-effectiveness of entrectinib follows the typical approach adopted in cancer appraisals of directly using extrapolated PFS and OS to populate a partitioned survival model. This approach, however, may not be a suitable model structure to leverage the available data for the present decision problem as it requires the availability of reliable, mature PFS and OS data for both the intervention assessed and the comparator. As outlined in Section 5.2.6, the available PFS and OS data for entrectinib and the reliability of the PFS and OS for the constructed comparator data set are severely limited, with concerns raised regarding the representativeness of the recruited population, uncertainties around positioning of entrectinib and potential confounding due to secondary therapy received.

The OS data from the efficacy evaluable analysis set is also immature with median OS not yet reached and as such, a significant proportion of the predicated benefit of entrectinib is based on the survival extrapolation. This issue is further exacerbated by the difficulty in validating predictions from the survival extrapolation analysis of entrectinib. It is good practice to assess the plausibility of the extrapolated portions of parametric survival models through the use of external data and clinical validity informed by clinical expert opinion and biological plausibility.[53] This is of particular importance in the present appraisal given the limitations of the OS data and lack of any external

datasets characterising the long-term prognosis of *NTRK* fusion positive patients. The context of this appraisal, however, makes this particularly challenging because the survival analysis is based on a population of patients with many different tumours types, which makes elicitation of clinical opinion particularly complicated.

Furthermore, while entrectinib has not yet been assessed by any regulatory bodies at the time of the ERG report, it is anticipated that the evaluation by these bodies will be based on a similar profile of evidence to that of larotrectinib,[54] which uses response outcomes (ORR and DOR) as the main regulatory endpoints. As the main regulatory endpoints in the larotrectinib evidence, it is these that formed the basis of the FDA decision to approve larotrectinib and is from these outcomes that they inferred the likelihood of clinical benefit, rather than using PFS or OS. The company's approach, however, ignores these outcomes and asks us to infer clinical benefit on the basis of PFS and OS, which appear to have been considered unsuitable for such a purpose in a regulatory setting, at least for larotrectinib. Alternative model structures built around response may therefore have been more suitable to address this decision problem and could represent a more robust approach upon which to predict long-term outcomes. Such an approach may also better lend itself to characterise uncertainty resulting from any heterogeneity in the treatment effect and therefore increase the opportunity to identify cost-effective subgroups as well as help focusing future data collection activities. In section 6 the ERG explores a response based model as alternative to the company's PFS and OS based approach.

The estimation of a single "full population" ICER

The model was designed to provide an estimate of a single "full population" ICER. This approach does not capture the heterogeneity in the patient population, and diverges from the Committee preference to date for a tumour type-specific treatment recommendation. The preference for making tumour type-specific decisions has been demonstrated in two previous NICE appraisals of interventions with a broad marketing authorisation, in which the Committee preferred to make tumour type-specific recommendations. In two appraisals of neuroendocrine tumours [55, 56] and of bone metastases from solid tumours and multiple myeloma [57], the NICE scope specified the consideration of the location of tumour or type of primary cancer, and the Committee deemed it appropriate to perform separate clinical and cost-effectiveness analyses given differences in prognosis and HRQoL associated with each of the tumour types. It is notable that in both of these appraisals, either separate studies were available for different tumour types or it was more feasible to undertake subgroup analyses than in the present appraisal.

Having a single "full population" ICER is not appropriate as the ICER is likely to differ across tumour types, and will be driven by a range of factors such as differences in treatment effect and comparator

effectiveness and comparator drug acquisition costs. Further, given the amount of heterogeneity associated with a histology-independent indication, estimating the average cost-effectiveness for the full patient population covered in the scope may not provide enough information to decision-makers about whether the drug is cost-effective across all subgroups.

While it is generally the view of the ERG that an optimised decision is preferable where possible because it increases allocative efficiency, it is acknowledged that this is more challenging in the present decision problem and an analysis of outcomes within each individual tumour type would not be sufficiently robust for decision making, since they would be based on very small patient numbers In this respect, a response based model, referenced above, may also confer advantages in accommodating any heterogeneity across the population because far fewer observations are required on response outcomes to draw meaningful conclusions about differences between tumour types, than would be required by time-to-event outcome such as PFS and OS.

5.2.2 The company's economic evaluation compared with the NICE reference case checklist

Table 28 summarises the ERG's assessment of whether the company's economic evaluation meets NICE's reference case and other methodological recommendations.

Table 28 Comparison of company's economic evaluation with NICE reference case

Attribute	Reference Case	Included in CS	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Comparator(s)	Alternative therapies in the NHS, including those currently regarded as current best practice	Partly	There is significant uncertainty regarding the position of entrectinib within the patient pathway. It is therefore not clear whether the included comparators represent those patients would receive in clinical practice. Furthermore the modelled population does not cover all tumour types covered by the NICE scope and therefore comparator therapies relevant to these tumour types were not modelled.
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective - costs	NHS and PSS	Yes	NHS and PSS costs have been taken into account.
Perspective - benefits	All health effects on individuals	Yes	QALY benefits to treated individuals were considered.

Time horizon	Sufficient to capture differences in costs and outcomes	Yes	The economic model uses a lifetime horizon (30 years). Less than 0.001 % of patients are expected to survive beyond this period.
Synthesis of evidence on outcomes	Systematic review	Yes	
Outcome measure	QALYs	Yes	EQ-5D-3L was collected in the STARTRK-2 trial and used to populate utilities for patients receiving entrectinib in the pre-progression health state. Quality of life for patients receiving comparator therapies and either therapy in the post progression health state were based on a weighted average of utilities reported in previous technology appraisals
Health states for QALY measurement	Described using a standardised and validated instrument	Yes	Derived from EQ-5D
Benefit valuation	Time Trade Off or Standard Gamble	Yes	Time Trade Off
Source of preference data	Representative sample of the public	Yes	Societal tariffs from EQ-5D.
Discount rate	3.5% on costs and health benefits	Yes	Costs and benefits have been discounted at 3.5% per annum.
Equity weighting	No special weighting	Yes	No special weighting undertaken.
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Probabilistic sensitivity analysis was undertaken.

5.2.3 Population

The modelled population considered in the company's base-case was assumed to represent the population in the integrated efficacy analysis, with clinical evidence on the effectiveness of entrectinib drawn from this analysis, see Section 3.1 for details. The modelled population therefore includes the 13 tumour types represented in this analysis. The distribution of the tumour types in the integrated efficacy analysis is also used in the comparator arm of the model, as comparator effectiveness, utilities and costs for each tumour type are all weighted according to their distribution in the integrated analysis.

Modelled patient characteristics were mean body weight and mean height which were drawn from the integrated analysis. These were used to estimate mean body surface, used in the dose calculation for some chemotherapy regimens.

ERG comment

The ERG's concerns regarding the model population centre around a key, but implicit assumption of the company's economic analysis that the modelled population represents the wide, histology-independent, anticipated market authorisation. As outlined in Section 3.1, the ERG has several substantive concerns regarding the population recruited to the integrated efficacy analysis and the degree to which it represents the population potentially eligible to receive entrectinib therapy. These limitations include the distribution of tumour types, the exclusion of available evidence on CNS and paediatric patients, unrepresented tumour types and the underrepresentation of the *NTRK2* gene fusion population. Each of these issues is discussed in turn in the sections below, with specific focus on the implications for the economic analysis presented.

Distribution of tumour types

The company's approach to constructing an established management comparator was to produce an average of clinical outcomes derived from NICE appraisals, weighted by the proportion of tumour types represented in the integrated analysis population. See Section 3.3 for a discussion of the comparator. The proportions used in the company's base case can be seen in Table 29.

The ERG is concerned that the estimated cost-effectiveness of entrectinib is being driven by the proportion of tumour types seen in the integrated efficacy analysis. In applying the distribution of tumour types observed in the trial, it is assumed that the trial population is reflective of practice, but this is unlikely to be the case. An alternative distribution is provided in Table 29. This is important because the prognosis and costs vary substantially across tumour types, in particular, the costs associated with screening for *NTRK* fusions can vary significantly.

The impact of alternative distributions of tumour types is also illustrated in scenario analyses whereby 100% weighting is given to a single tumour type. These show that the ICER varies significantly by tumour type. The result of this scenario show the ICER can range from £114,524 if 100% weighting is given to pancreatic cancer to £31,064 if 100% weighting is given to MASC. For further details, see Section 5.2.9.3.

Reflecting these concerns, the ERG queried the company regarding the representativeness of the distribution of patients across tumour types in the integrated efficacy analysis. The CS stated in their response that given [redacted] patients were screened for *NTRK* gene fusions, it is reasonable to expect that the

proportions of tumour types used in the base-case, may reflect the population seen in clinical practice, with the exception of MASC, which the company agreed is over represented.

Consideration of alternative data sources regarding the likely distribution of tumour types, however, undermines the company's stated position. A comprehensive data set of over 166,000 tumour samples along with the observed frequency of *NTRK*-gene fusions in specific tumour types was provided to the ERG in response to clarification questions. This information was used to estimate the number of patients in the population eligible for entrectinib, which in turn represents an alternative distribution of tumour types. The method used to estimate this can be found in Appendix B. This alternative distribution can be seen in Table 29.

Table 29. Distribution of tumour types in the entrectinib integrated efficacy evaluable analysis set and an alternative ERG distribution

Tumour Type	Proportion in CS	ERG
Sarcoma	24%	[redacted]
NSCLC	18%	[redacted]
MASC	13%	[redacted]
Breast	11%	[redacted]
Thyroid	9%	[redacted]
CRC	7%	[redacted]
Neuroendocrine	6%	[redacted]
Pancreatic	6%	[redacted]
Gynaecological	4%	[redacted]
Cholangiocarcinoma	2%	[redacted]
CS, company submission; FMI, Foundation Medicine Inc.; NSCLC, non-small cell lung cancer; MASC mammary analogue secretory carcinoma; CRC colorectal cancer		

To explore the impact of this alternative distribution on the results of the economic model, a scenario analysis is implemented in Section 6.

Exclusions from available evidence

The patient population considered in the base-case failed to encompass the entire population as defined in the NICE scope and was limited to the 10 categories of tumour types present in the entrectinib efficacy evaluable analysis set. The base-case population also excluded patients with primary CNS tumours and paediatric patients from the efficacy evaluable analysis set despite their eligibility for inclusion. The patient population considered in the base-case failed to encompass the entire population as defined in the NICE scope and was limited to the 10 categories of tumour types present in the entrectinib efficacy evaluable analysis set. The base-case population also excluded

patients with primary CNS tumours and paediatric patients from the efficacy evaluable analysis set despite their eligibility for inclusion.

The ERG believes patients with primary CNS tumours and paediatric patients should be included in the analysis as they fall within the population in which the company is seeking a recommendation and as such, requested a scenario analysis including these patients. The company response to this request was to highlight that the inclusion of paediatric patients in the economic analysis is challenging due to the absence of a counterfactual or at least any robust comparator data. The company also noted that patients with primary CNS tumours were excluded for a number of reasons including the lack of follow-up and that response was measured using Response Assessment in Neuro-Oncology Criteria (RANO). This is discussed in Section 4.2.6.1. However, the company reiterated that a NICE recommendation in accordance with the proposed license is anticipated, which includes all paediatric and adult patients harbouring *NTRK* fusion including primary CNS tumours.

The ERG is concerned that the company is seeking a recommendation in patients with primary CNS tumours and paediatric patients, yet despite data being available at the original CCOD for these patient groups, the company has decided to omit information from the base case provided in the CS due to differing response measurements when in fact response outcomes are not used in the economic model.

The company provided an updated economic model in the updated response to clarifications with the inclusion of the five efficacy-evaluable adult primary CNS tumour patients and the seven paediatric patients added to the model. The ERG welcomes the inclusion of these two additional patient populations into the economic analyses. See Section 5.2.9.4 for a discussion of the impact of the inclusion of these populations on the company's base case ICER.

Unrepresented tumour types

As highlighted in Section 3.1 an important limitation of the integrated efficacy analysis is that it does not include all tumour types covered by the anticipated marketing authorisation. For a list of those tumour types identified, see Table 2. This issue of unrepresented tumour types is potentially significant as based on current knowledge of which tumour types exhibit an *NTRK* fusion, a [redacted] will be covered by the anticipated marketing authorisation based on those tumour types in which an *NTRK* fusion has been identified. Furthermore, clinical advice received by the ERG has suggested that it is plausible that *NTRK* gene fusions could potentially be present in 400+ possible tumour types.

The impact of this omission in the modelled population is potentially very significant particularly given the weak support for the assumption of homogeneous response rates for the different tumours (Section 4.3.1), which suggests that different response rates may be observed in the missing tumour

types. Furthermore, this issue persists beyond the treatment effect and also impacts upon costs and utilities of both entrectinib and the comparator as the 10 tumour types present in the trial are being assumed to represent all of the potential tumour types in the population, which as outlined above is very unlikely to be the case. See Section 5.2.6 for a discussion of how the tumour types may impact on the treatment effect and costs respectively.

Underrepresentation of NTRK2

The efficacy evaluable analysis set includes 54 patients, with only one of these patients harbouring an *NTRK2* gene fusion. The ERG is concerned about the low representation of patients with *NTRK2* fusions in the trial and that the population used in the underrepresents this specific fusion type. At the points for clarification stage, the ERG queried the underrepresentation of these patients, with the company responding that the low number of recruited patients reflects the low prevalence of this fusion within the wider population. However, based on The Foundation Medicine Inc. data provided by the company in their clarification response, *NTRK2* patients may make up [redacted] of the *NTRK*-fusion positive tumours, much higher than the 2% included in the entrectinib integrated efficacy analysis.

There is insufficient evidence to establish whether patients with an *NTRK2* gene fusion have different prognoses to patients with *NTRK1* and *-3* gene fusions or whether there is potential for different responses to entrectinib based on *NTRK* fusion type. Clinical advice given to the ERG suggests that both of the scenarios are plausible and it was explained that it will depend upon the role *NTRK2* fusions play in the tumour growth within the tumour types they occur in. Further, data presented to the FDA as part of an NDA Multidisciplinary Review and Evaluation of larotrectinib in patients with *NTRK*-gene fusions, suggests that patients with *NTRK2* gene fusions had a lower overall response rate than those with *NTRK1* and *-3* gene fusions, which may suggest differential response to TRK inhibitors in this population.[54]

The under representation of this fusion type therefore presents additional uncertainty and a further problem with the representativeness of the population that makes up the integrated efficacy analysis. The size and direction of the consequences of this in the economic model are not fully apparent, but if, as suggested by the FDA, *NTRK2* positive patients are less likely to respond, it may lead to an overestimation of the treatment effect and consequently an underestimation of the ICER.

5.2.4 Interventions and comparators

The economic model presented in the CS compares entrectinib with established management which was assumed to consist of a blended comparator of chemotherapy regimens and BSC.

The modelled dose of entrectinib was assumed to align with the anticipated recommended dose of entrectinib, which is detailed as 600 mg once daily, with each 600 mg dose administered as 3x 200mg capsules. Duration of treatment for entrectinib was assumed to be aligned with the anticipated marketing authorisation, i.e. [redacted] and in the base-case analysis was set equal to progression free survival. Scenario analysis was also presented where time on entrectinib treatment was based on observed time on treatment in the integrated efficacy analysis.

The modelled blended comparator consisted of chemotherapy regimens and BSC. This blended comparator was based on previous NICE TAs identified as providing relevant effectiveness data. As outlined in Section 3.3 comparator effectiveness data for each tumour type was generated from multiple TAs therefore the modelled comparator was blended both at the individual tumour type level as well as at the across tumour types. The active comparators consisted of combination of alkylating agents (oxaliplatin, trabectedin), antimetabolites (capecitabine, fluorouracil, gemcitabine), anti-tumour antibiotics, multi-kinase inhibitors (Nintedanib), topoisomerase inhibitors (irinotecan), mitotic inhibitors (docetaxel, paclitaxel) and therefore cover a wide range of agents. Dosing of comparator therapies was based on their Summary of Product Characteristics (SmPC) guidance with duration of therapy based on PFS, i.e. treatment until either progression or death.

A list of comparators for each tumour type is presented in Table 30. Note that the comparators listed reflect those used to generate comparator effectiveness as comparator costs were based solely on active comparators. The modelled effectiveness data was therefore inconsistent with the modelled comparator costs, see Section 5.2.8 for further discussion of comparator drug acquisition costs.

Table 30 Summary of comparators modelled and data sources

	Therapy	**Reference**
Non Small-cell Lung Cancer	Docetaxel	Average of values from NICE TAs 520, 428, 483, 484, 403, 347, 124
	Docetaxel + nintedanib	NICE TA347
Colorectal Carcinoma	FOLFIRI	NICE TA307
	Irinotecan	NICE Guideline CG121 - Kim et al 2009
	Trifluridine-tipiracil	NICE TA405
	Trifluridine/tipiracil	NICE TA405
	Best supportive care	NICE TA405
	Best supportive care	NICE TA405
Breast Cancer incl. secretory breast	Capecitabine	NICE TA515
	Eribulin	NICE TA423
	Vinorelbine	NICE TA423
	Gemcitabine + paclitaxel	NICE TA423
Salivary Gland Cancer (incl. MASC)	Best supportive care (Platinum Gemcitabine data used as surrogate)	Surrogate data for BSC - Laurie et al. 2010
Soft Tissue Sarcoma	Doxorubicin	NICE TA465
	Trabectedin	NICE TA185
Pancreatic	Gemcitabine & nab-paclitaxel	NICE TA476
	Gemcitabine	NICE TA476

	FOLFIRINOX	NICE Guideline NG85 - Conroy et al 2011
Thyroid (papillary), unsuitable/refractory to radioactive iodine	Best supportive care	NICE TA535 (Cross-over adjusted value from Guo et al 2015)
	Best supportive care	NICE TA535
Neuroendocrine tumours	Everolimus	NICE TAs 449 and 539
	Everolimus	NICE TAs 449 and 539
	Best supportive care	NICE TAs 449 and 539
	Best supportive care	NICE TAs 449 and 539

In addition to the above, the economic model also allowed for subsequent therapy following discontinuation of entrectinib treatment; no subsequent therapy was assumed for patients receiving established management. Subsequent active therapy was assumed to be received by [redacted] of patients based on the proportion of patients with progressed disease who received subsequent therapy in the integrated efficacy analysis. Subsequent therapies were assumed to consist of established management as defined above. Patients receiving subsequent therapy were assumed to do so from the time of progression until death. Scenario analysis was also presented in which different rates of subsequent therapy used were assumed in the entrectinib model arm (50% and 80%) as well as a scenario in which 50% of chemotherapy patients were assumed to continue therapy post-progression. See Section 5.2.9.3 for the result of the scenario analyses.

ERG comment

As discussed in Section 3.3 significant uncertainties existing regarding the positioning of entrectinib in the patient pathway because the anticipated marketing authorisation allows entrectinib to be used at several points in the treatment pathway. Table 6, reported on page 30 of the CS provides some indication of where the company anticipate entrectinib will be positioned in UK practice. However, this table does not cover all of the tumour types represented in the integrated analysis or the additional tumour types know to harbour *NTRK* fusions, see Table 2 Tumour types included in ERG population size calculations. As outlined in Section 3.1 there are also some concerns about whether the indicated positions proposed in Table 6 of the CS, reflect the likely positioning of entrectinib. This uncertainty in the positioning of entrectinib means that it is not possible to validate the selected comparators as it is not clear at what position entrectinib will be positioned in the respective pathways.

The company's approach to identifying comparators also does not help to provide clarity regarding which are the appropriate comparator because for each tumour type, multiple TAs spanning multiple lines of therapy have been selected. The ERG considers this a logical inconsistency and that at least for the purposes of generating an externally valid comparator it would have been preferential to instead select a single line therapy to represent the anticipated positioning of entrectinib. Furthermore, while the ERG has not been able to fully validate the selected comparators for every tumour type, there are clear examples of where comparators have been selected that are rarely used in UK practice.

For example, with respect to the third line treatment of breast cancer, gemcitabine plus paclitaxel is rarely used while in colorectal cancer, irinotecan is less likely to be administered than FOLFIRI as second line therapy.

A further issue relates to the representativeness of the modelled comparators which, in principle should not only represent the trial population, but the wider population covered by the marketing authorisation i.e. all patients with *NTRK* fusions. As discussed in Section 3.10, the distribution of tumour types is not fully reflective of the eligible population with some types over/under represented in the trial population as well as there being a significant number of missing tumour types that may represent 20% of the population potentially eligible for treatment with entrectinib, based on those tumour types in which an *NTRK* gene fusion has been identified. See Table 6 for a list of these tumour types.

The validity of the modelled comparators is therefore subject to significant uncertainties, and at least for a number of tumour types, the selected comparators do not represent current UK practice, see Section 3.3. Furthermore, even ignoring these issues, the modelled comparators are only appropriate to the degree that they are representative of tumour types missing from the integrated efficacy analysis, and there is no reason to believe that this is the case.

In addition to the above issues, the ERG also has concerns regarding the modelling of subsequent therapies both with respect to the duration of subsequent therapy assumed in the company's base-case model and the mix of therapies assumed.

With respect to duration of therapy, the base-case analysis assumes subsequent therapies are received from progression until death. The ERG, however, considers that this assumption is likely to be overly pessimistic and that many patients will move to BSC before death. This may reflect either exhaustion of treatment options or lack of fitness to continue to receive therapy. The impact of this assumption is to decrease costs in the entrectinib arm of the model and therefore to substantially decrease the ICER.

At the points for clarification stage the ERG requested that the company provide further analyses in which patients were assumed to discontinue subsequent therapy prior to death. In response to the ERG's request the company provided two scenarios: one in which patients continued on subsequent therapy for three months and a second in which they continued on subsequent therapy for six months. The resulting ICERs from this analysis were £39,849 and £40,093 per QALY respectively, both of which are substantially lower than the company base-case of £52,609 per QALY. The ERG's preference is for a 6 month period of treatment following progression given that this represent roughly half the PPS, though the ERG acknowledges that this is a rather arbitrary assumption.

With respect to the mix of subsequent therapies used, the ERG considers the company's assumptions reasonable to the extent that entrectinib is likely to displace currently used therapy, but notes that this mix of therapies was very different to that received by patients in the integrated analysis which includes a wide range of therapies including several targeted therapies and immunotherapies, see points for clarification response question B10.

5.2.5 Perspective, time horizon and discounting

Consistent with the NICE methods guide [49] the company's analysis used NHS and Personal Social Services (NHS & PSS) perspective and discounted costs and benefits at a rate of 3.5%.

A lifetime horizon of 30 years was chosen as it was considered sufficient to capture all relevant differences in costs and benefits between comparators. The impact of shorter 5 year, 10 year and 15 year time horizons were also explored in a scenario analysis. The ERG considers this an appropriate time horizon, as it is very unlikely that any patients would remain alive beyond this time period.

5.2.6 Treatment effectiveness and extrapolation

As stated in Section 5.2.1, the company used a partitioned survival approach to provide a direct comparison of the timing and rates of progression and death. The main effectiveness inputs included in the company's economic model are therefore PFS and OS. For the model base case, OS and PFS survival estimates for entrectinib were drawn from the integrated analysis which pooled data from the ALKA, STARTRK-1, and STARTRK-2 trials. The integrated analysis set included 54 patients across 13 different tumour types, but excluded 6 patients with primary CNS and a paediatric patient. The CS states that the patients with primary CNS tumour were excluded from the integrated analysis because progression was measured by a different criteria, RANO rather than RECIST 1.1. The paediatric patient was excluded as the integrated analysis only includes patients from the adult studies, thereby excluding patients from the paediatric study STARTRK-NG. See Section 3.1 for further discussion of the population and critique of these exclusions.

The data-cut off used in the economic model was the 31st of May 2018; note this differs from the clinical section of the CS submission which presents data from the later [redacted] cut off. At the points for clarification stage the ERG requested that this latest data cut be integrated into the economic model which was provided by the company in their response. Results from this updated cut are presented in Section 0.

For the comparator therapies PFS and OS outcomes were sourced from previous TAs identified by the company to represent established management in the NHS for each of the tumour types modelled. A list of the source TAs used by the company to model comparator PFS and OS is presented in Table 30 (Section 5.2.4). For each tumour type median PFS and OS was extracted for all relevant TAs and then

a simple average of median PFS and OS estimated for each tumour type. Mean PFS and OS used to estimate time in pre-progression and post-progression health states was then estimated for each tumour type by assuming PFS and OS followed an exponential survival function.

Figure 12 illustrates the KM curve and extrapolated exponential OS curve for entrectinib (using the latest data cut, [redacted]) along with the average survival curves generated for patients receiving established management. Figure 13 illustrates the KM curve and various extrapolated OS curves presented in the CS using the latest data cut. Figure redacted

Figure 14 and Figure 15 present similar data for PFS.

Figure 12 Kaplan-Meier, parametric extrapolated exponential OS for entrectinib and average survival for established management ([redacted] data cut)

Figure redacted

Figure 13 Kaplan-Meier and parametric extrapolations of OS for entrectinib (▉ data cut)

Figure redacted

Figure 14 Kaplan-Meier, parametric extrapolated exponential PFS for entrectinib and average survival for established management (▉ data cut)

Figure redacted

Figure 15 Kaplan-Meier and parametric extrapolations of PFS for entrectinib (▉ data cut)

Figure redacted

As can be seen above, available PFS and in particular OS data is immature, with ▉ and ▉ of patients having experienced a PFS and OS event respectively.

5.2.6.1 Uncontrolled comparison of treatment effectiveness

Generating an appropriate comparator dataset poses a significant challenge due to the histology-independent nature of this appraisal. As described above, the company's approach focuses on generating a comparator dataset using data sourced from previous TAs, which are then weighted by the distribution of tumour types in the integrated efficacy analysis. The principal concern regarding the company's approach is the fact that it relies on an unadjusted naïve comparison between the weighted comparator and the integrated efficacy analysis which has significant scope for confounding biases resulting from differences in population characteristics.

One potentially significant source of confounding results from differences in the number of lines of therapy patients have received. As highlighted in Section 4.2.6 a significant proportion of the patients in the integrated efficacy population (37.0%) received entrectinib as a first-line systemic therapy with a further 20.4% receiving entrectinib as a second line therapy. The comparator data set however, draws predominantly from patients in later lines of therapy, with 7 of the 10 tumour types, representing 57% of patients, including no evidence from patients receiving first line therapy. Further, a feature of the integrated analysis is that patients received entrectinib at different points in the clinical pathway including within single tumour types. For example, the integrated analysis includes NSCLC patients who received entrectinib as a first, second and third or later lines of therapy. The comparator data set generated, however, makes no account for this and does not attempt to match the comparator data used to the position patients are in the integrated analysis. This difference between the comparator data set and the integrated analysis is a potentially significant source of bias as line of therapy is an important determinant of prognosis. As a result, it is very likely that estimates of PFS and OS are confounded in favour of entrectinib.

A further and potentially important source of confounding bias is the fact that only a small proportion of patients in the comparator data set are likely to be *NTRK* fusion positive. This is problematic because there is evidence to suggest that *NTRK* fusions are prognostic, though available evidence is limited. To account for the potential prognostic value of *NTRK* fusions the company presents a scenario analysis which draws on evidence from patients with colorectal cancer. This analysis suggests *NTRK* is an indicator of poor prognosis and results in much lower ICERs than the base-case analysis. In the ERG's view, however, this scenario should be interpreted with caution as literature identified by the ERG suggests that the prognostic value of *NTRK* may be different across tumour types (Table 2). This was also confirmed by the clinical adviser to the ERG, who suggested that such variability in prognosis was possible, and likely dependent on the role *NTRK* fusions play in that specific tumour type.

In addition, there is also the possibility of differences in a wide range of other patient characteristics, such as age and ECOG status which are commonly prognostic. The CS does not report any baseline characteristics for the comparator arm and interpretation of these would have been complicated by the large number of tumour types and data sources. However, one approach the company could have taken to better understand the potential for differences is to have extracted commonly reported prognostic baseline characteristics such age, ECOG status and presence of brain metastasis from the source TAs. These could then have been used to generate a weighted set of baseline characteristics in a similar way to how the effectiveness data was generated. This could then have been compared to the integrated efficacy population. Such an approach could also potentially have been extended by implementing a Matching-adjusted Indirect Comparison (MAIC) or Simulated Treatment Comparison (STC) [58] to adjust the effectiveness data for entrectinib. Implementing such an approach would, however, have been very challenging not least because of the large number of source data sets and would make strong assumptions about the prognostic value of characteristics across tumour types. Furthermore, even if a suitable adjusted comparison could be generated it would only be able to account for a small number of observed characteristics due to the small sample size in the integrated analysis and therefore there would likely be significant residual confounding bias.

In summary, while the ERG considers that the broad approach adopted of using a weighted comparator data set to be reasonable, there are significant challenges associated with implementing this successfully, as well as further issues resulting from the company's execution of this approach. As such, while rectifying the specific issues highlighted above could potentially improve upon the validity of the comparator data set, it is likely that substantive concerns regarding the suitability of the comparator data would remain. Because of this, the ERG considers that company should have also considered alternative methods of generating comparator effectiveness estimates. The company could have for example considered two approaches discussed in a recent publication by Hatswell *et al*.[1] and

described below, which, while also subject to limitations, could have provided some degree of reassurance regarding predicted comparator effect estimates.

The first approach proposed by Hatswell *et al*.[1] uses effectiveness data on non-responders as a proxy for patients not receiving an active treatment. Comparator effectiveness estimates of PFS and OS under this approach would therefore be based on observed PFS and OS amongst non-responders in the integrated efficacy analysis. The rationale behind this approach is that patients in which no response is observed represent those with a lack of treatment effect (as they have no response to treatment) and therefore are representative of a counterfactual where no effective therapy exists. The advantages of this approach are that the patients are likely to be better matched with the intervention arm because they are drawn from the same population. However, this approach has disadvantages and makes a number of strong assumptions. It assumes that response is not systematically correlated with tumour type, which is unlikely to be true as is observed in analysis presented in Section 4.3. It also assumes that lack of response is indicative of comparator treatment effects which is likely to depend on the treatment considered as a comparator, but may be reasonable given the anticipated marketing authorisation which permits use of entrectinib [redacted].

The second approach outlined by Hatswell *et al.* [1] compares the outcomes for patients on entrectinib with their outcomes on the previous line of therapy. Under this approach the average time to progression (TTP) on the previous line of therapy is compared with average patients TTP when treated with entrectinib and a ratio estimated. The inverse of this ratio would then be applied to the observed PFS data from the entrectinib integrated efficacy analysis, to estimate comparator PFS, with PPS survival for both the entrectinib arm and comparator arm assumed to be same and also sourced from the integrated efficacy analysis. As with the first method, the advantage is that effect estimates are drawn from the same population as the intervention arm and therefore better matched, however there are also disadvantages. Firstly, this can only be implemented for patients who have received a previous line of therapy. Secondly, it also assumes that the ratio of TTP across lines of therapy is indicative of the treatment effect and it is uncertain to what degree this is likely to hold true. Finally, because this method can only estimate PFS it assumes that PPS survival is the same across therapies which similarly may not hold true.

To explore the uncertainties in the estimated treatment effect, the first method is implemented using non-responders as controls using data provided by the company at the points for clarification stage. See Section 6.5.1 for the results of this responder-based approach. The ERG also considered the second approach potentially valid, but did not feel that that this could be implemented within the time and resource available as the data requirements are significantly higher, but this could be considered as a reasonable scenario to be implemented at a later date.

5.2.6.2 Heterogeneity in treatment effect

A significant issue in the context of the present appraisal is the possibility for heterogeneity in the treatment effect across tumour types, as well as across other clinical characteristics such as age (paediatric vs adults), fusion type and position in the treatment pathway. This issue is however, largely neglected in the CS, with minimal analysis devoted to exploring the potential for heterogeneity in PFS or OS or indeed other measures of effectiveness. This is important as an implicit assumption in the company's base-case analysis is that the modelled treatment effect is constant not only across the modelled tumour types, but also across all tumour types covered by the marketing authorisation, see Table 6 for a list of the tumour types in which an *NTRK* fusion has been identified in the literature.

As demonstrated in the ERG exploratory analyses on response data (see Section 4.3.1), there is evidence to suggest that the treatment effect is heterogeneous across tumour types. These analyses showed that response outcomes for entrectinib vary considerably across tumour types ranging from [redacted] to [redacted], see Table 26, Section 4.3.1.3. Further, the predictive distribution, which provides an estimate of the likely response rate in an unrepresented tumour has a credible interval of [redacted] to [redacted] (Table 25, Section 4.3.1) implying that mean response across all eligible tumour types could be very different to that estimated in the integrated analysis. It is unclear how heterogeneity in response outcomes impacts on survival outcomes, and consequently cost effectiveness estimates, but these analyses do illustrate the potential for heterogeneity. The ERG presents further analyses in Section 6 exploring this potential heterogeneity in the treatment effect using a response based model to integrate the results of this analysis into the model.

5.2.6.3 Overall survival

Entrectinib

To extrapolate OS, standard parametric models (exponential, Weibull, log-logistic, lognormal, Gompertz, and generalised gamma) were fitted to the available KM data. To determine the most appropriate model, the CS states that reference was made to fit statistics (AIC/BIC; see Table 46 of the CS), visual fit to the observed KM curves, and clinical plausibility of survival estimates. Figure 16 provides a graphical summary of each curve and their fit to the observed KM data.

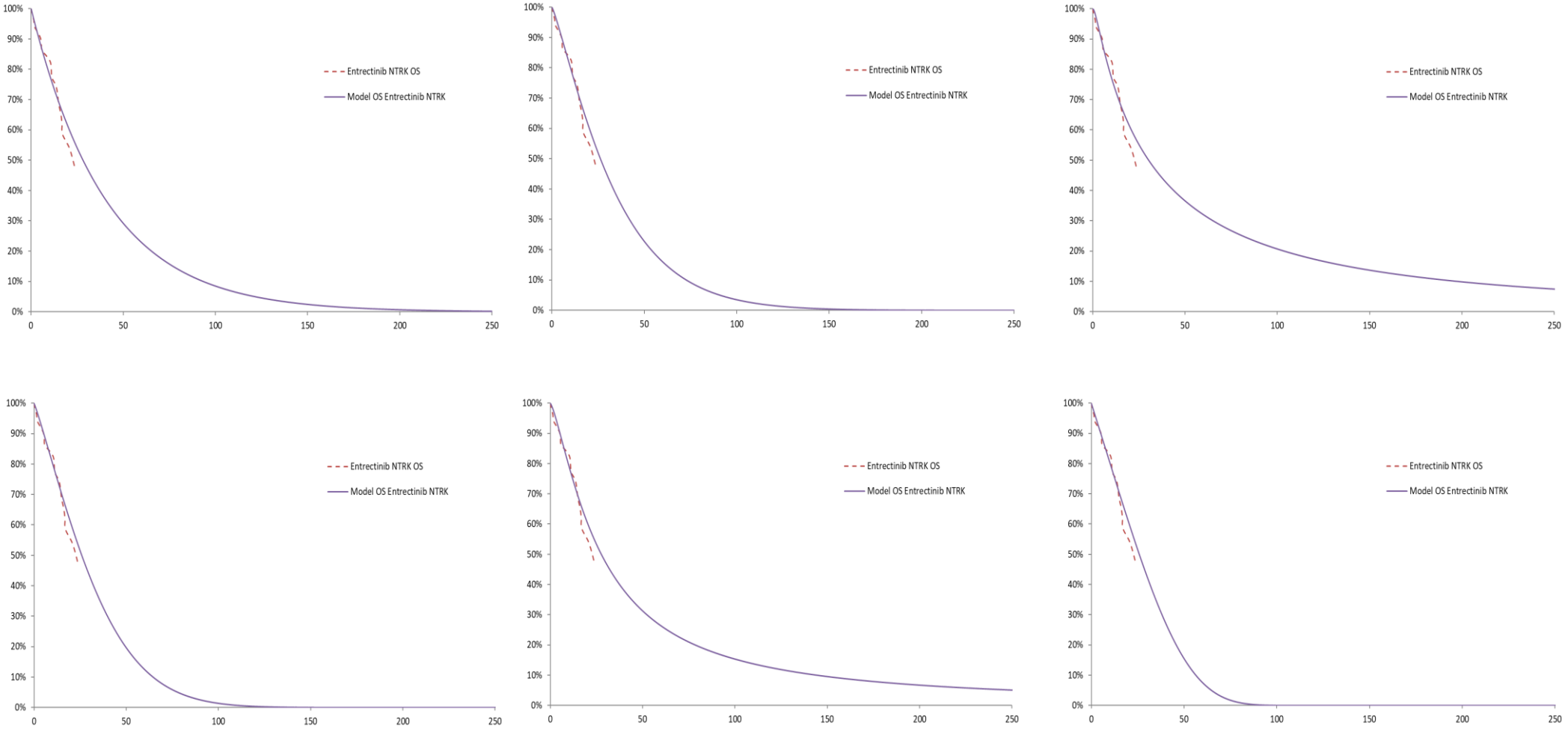
Consideration of clinical plausibility made reference to a landmark analysis of OS, which considered predicted proportion of patients alive at 2, 5, 10, 15, 20 and 30 years. This analysis is summarised in Table 47 of the CS. On the basis of the landmark analysis, only two of the curves were considered by the company to produce clinically plausible survival predictions. These were the exponential and Weibull curves. The base-case survival model presented in the CS selected an exponential curve and was justified on the basis that this had the best statistical fit, with all six other parametric curves considered in scenario analyses. The updated base-case based on the latest data cut ([redacted]) and presented

as part of the company's points for clarification response also retained the exponential function as company's preferred extrapolation. In their response the company cited statistical fit as the justification for selecting the exponential functions stating that it consistently had the best statistical fit across all scenarios.

ERG Comment

In the context of the present appraisal which combines so many tumour types, consideration of clinical plausibility of alternative extrapolations of OS is challenging. However, the ERG considers the exponential function selected to be a plausible extrapolation of OS, with good statistical fit to the observed OS data. The exponential function further makes reasonable predictions regarding long-term survival with most patients predicted to have died after five years and nearly all at 10 years. The ERG, however, notes that other survival functions similarly have good statistical fit while also making reasonable predictions about long term survival extrapolations, including the Weibull, Gompertz and Gamma functions. In considering the appropriateness of these individual functions the ERG notes that the exponential function is the only one to predict that post progression survival is longer than pre-progression survival. The ERG questions the clinical plausibility of this given that entrectinib therapy is discontinued on progression and that only [redacted] of patients received any subsequent therapy. Further, the positioning of entrectinib as a therapy of last resort suggests that few effective treatment options would remain to regain tumour control after progression, with consequences for post-progression mortality. The ERG therefore considers the Weibull, Gompertz and Gamma to represent a more reasonable extrapolations of OS with the ERG favouring the Weibull function due to its marginally better statistical fit over the other two functions.

Figure 16 Alternative entrectinib OS parametric curves (1st row, left to right: Exponential, Weibull, Log-normal; 2nd row, Gamma, Log-logistic, Gompertz, Exponential)



Comparator therapies (established management)

As described in Section 5.2.4, OS for comparator therapies was drawn from previous TAs and extrapolated assuming an exponential survival function. Because the company extracted only median OS values and not KM data, no other survival functions were explored. Survival predictions were, however, validated by clinical experts who "endorsed" the model predictions.

ERG Comment

The ERG has two main concerns about the modelling and extrapolation of the comparator data.

The ERG's first concern relates to the method used to pool median OS values from the NICE TAs selected to represent established management. Specifically, the ERG considers the approach of averaging median OS estimates at the individual tumour level to be inappropriate and mathematically incorrect. The company should instead have estimated mean OS for each TA and then pooled these. The impact of this calculation error is small, but is corrected in Section 6.

The ERGs second concern relates to the use of the exponential function to extrapolate OS. The ERG notes that this is a consequence of the approach taken by the company to identifying relevant effectiveness evidence, but considers it less than ideal. Examination of the source TA reveals that the exponential curve was rarely favoured by the committee in the considered appraisals, with the consequence that comparator OS is likely overestimated for some tumour types and underestimated for others. Furthermore, the estimates of post progression survival appear excessively long, with mean survival time post progression twice that of survival time prior to progression. In the context of the comparator data it is, however, unclear whether this is driven by underestimation of PFS or overestimation of comparator OS, though both of these will result in the ICER being overestimated.

The ERG further considers that other methods could have been adopted by the company to develop a comparator dataset, which would have greater face validity and flexibility. For example, the company could have extracted estimated life years gained from the committee's preferred scenario which would have accounted for the committee's preferred extrapolation. Alternatively, the company could also have extracted reported KM data from each TA, which would have given the company the flexibility to fit the best parametric extrapolation.

5.2.6.4 Progression free survival

Entrectinib

In common with the approach used for OS, PFS was extrapolated by fitting standard parametric functions to the available KM data with selection of an appropriate parametric function similarly

based on references to the statistical fit, reference to the hazard trend and to landmark analysis considering the clinical plausibility of predicted PFS at 2, 5, 10, 15 and 20 years.

On the basis of the landmark analysis depicted in Table 45 of the CS four of the parametric functions were considered clinically plausible. These were the exponential, Weibull, Gamma and Gompertz functions. The company base-case presented in the CS selected the exponential curve on the stated grounds that this represented a "conservative, but statistically and clinically plausible estimate of progression-free survival for entrectinib patients". As with OS, the updated base-case using the latest data cut (█) retained the exponential function as the company's preferred extrapolation, with statistical fit cited as the main reason for selecting this curve.

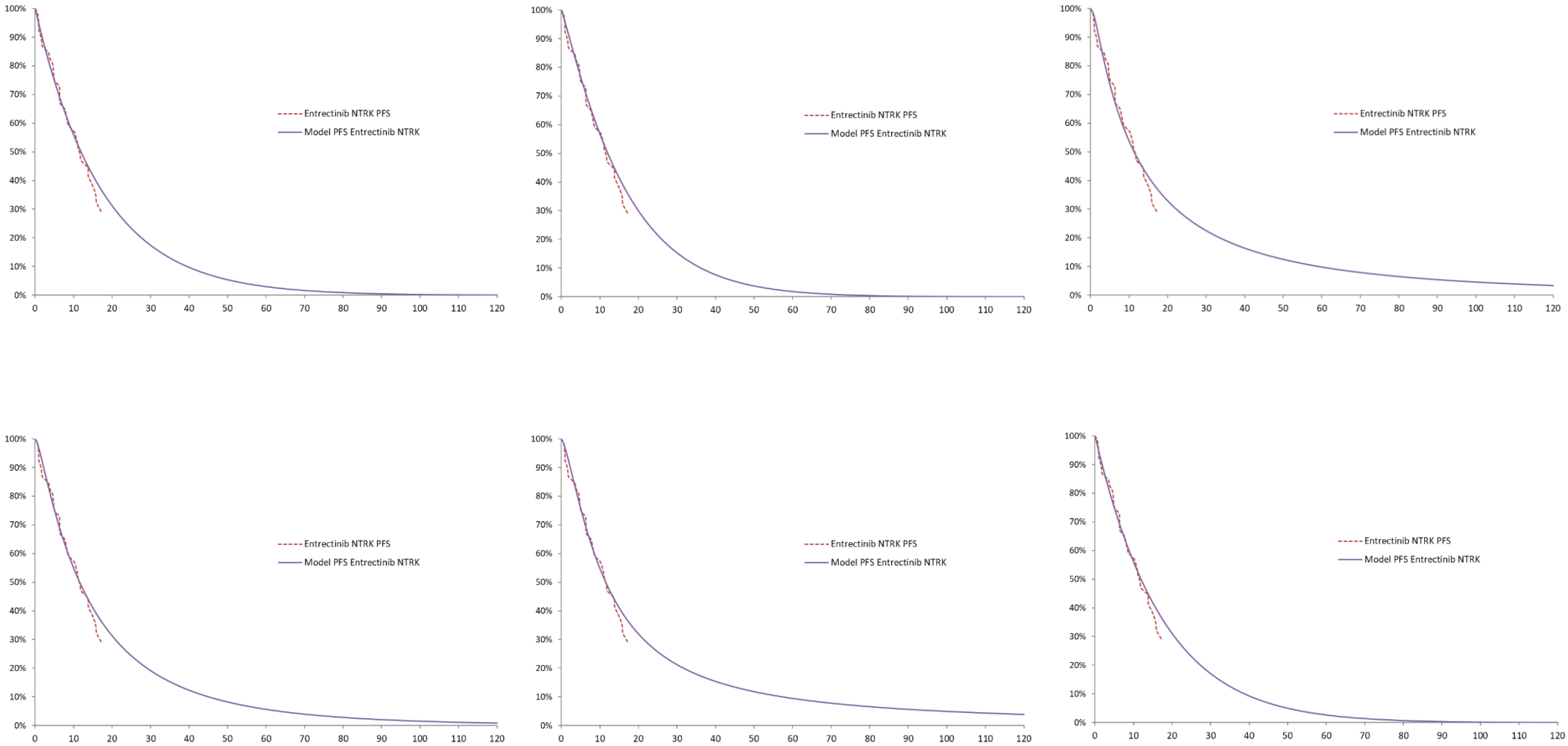
A graphical comparison of the extrapolations of PFS using the base case and alternative parametric function is presented in Figure 17 below.

The ERG considers that the exponential, Weibull, Gamma, and Gompertz functions all represent reasonable extrapolations of PFS and produce predictions that are consistent with the OS evidence when an exponential function is used; the ERG notes that after a certain time point, the Log-normal and Log-logistic functions yield estimates of progression that were higher than any of the plausible OS survival curves. The ERG, further notes that all four of these curves produce very similar estimates of mean PFS ranging from 15.85 months using the Weibull function to 17.65 months using the Gamma function. This small variation in predicted mean PFS and the relative insensitivity of the model to this input means that the ICER is relatively robust to the function adopted, changing by less than £1,000 per QALY across all four plausible extrapolations. The ERG considers the Weibull function as the preferred extrapolation function, as it is likely the most appropriate given its good statistical fit and for consistency with the ERG's preferences regarding the extrapolation OS; combining the Weibull function for OS with an exponential function for PFS implies a decreasing hazard for progression events which is clinically unlikely.

Comparator therapies (established management)

The company's approach to modelling PFS for patients receiving comparator therapies was the same as for OS and used median PFS values drawn from relevant previous TAs. These were then extrapolated assuming an exponential function with no other functions considered.

Figure 17 Alternative entrectinib PFS parametric curves (1st row, left to right: Exponential, Weibull, Log-normal; 2nd row, Gamma, Log-logistic, Gompertz, Exponential)



ERG Comment

The ERG's concerns regarding the extrapolation of PFS are largely similar to that of OS, with issues relating to both the method of calculation and limitations of assuming that PFS follows an exponential function. The impact of this uncertainty in comparator PFS is, however, not as significant as the uncertainties relating to OS as it is a less significant driver of the model. Importantly, however, because PFS determines time on treatment, increasing PFS actually decreases the ICER rather than increasing it. This is because the relatively small increases in QALY results from extending PFS are outweighed by increased drug acquisition costs.

5.2.6.5 Adverse events

Adverse events from treatment with entrectinib and established management were considered in the economic model to capture associated costs. Scenario analysis also considered the impact of AEs on quality of life by including disutilities. Only Grade 3-4 events were modelled and only if they occurred in >5% of patients. Four AEs were modelled: anaemia, fatigue, neutropenia and weight increase. Event rates for both the entrectinib and established management arms of the model were drawn from the integrated efficacy analysis, see Table 65 of the CS Page 123 for event rates. Because of the lack of data on event rates for AEs occurring in patients receiving established management, the model assumed identical event rates for both entrectinib and established management arms with the exception of weight increase which was assumed to only occur in patients who received entrectinib. The AE weight increase, was, however, associated with zero cost and therefore the assumption of differential event rate had no impact on the results of the base-case analysis. In effect, therefore the base-case analysis assumes no difference in AE rates between entrectinib treatment and established management.

The company considered this assumption to be conservative with respect to entrectinib as the company noted that many of the chemotherapy regimens patients would receive as part of established management have a poor toxicity profile.

ERG Comment

The ERG considers the use of the integrated analysis to model AE rates in patients receiving entrectinib to be reasonable and recognises the difficulty of modelling AE rates for patients receiving established management. However, the ERG considers the approach taken by the company to be less than ideal. Alternative approaches could have been considered for generating adverse event rates, such as using the source TAs to identify AE rates for the comparator therapies. Further, the company's assertion that the assumption of largely equal AE rates is a conservative one is not a certainty and will depend significantly upon the comparator being considered. For example, a number of the comparators listed by the company, see Table 30 above, consist of BSC and therefore patients

are receiving no active therapy. In such cases we may expect adverse event rates for patients receiving entrectinib to exceed those of established management. The impact of these simplifying assumptions is, however, likely to be minimal as even in an extreme scenario where 0% event rates are assumed for patients receiving established management the ICER falls only very slightly.

5.2.7 Health related quality of life

The company estimated health state utility values for entrectinib based on EQ-5D-3L data collected in the STARTRK-2 study. For the comparator arm, these were identified via a review of published literature. A summary of utility values is presented in Table 31. Utilities were not adjusted for age, and disutilities relating to adverse events were not applied in the base case analysis.

Table 31 Utility values used in the cost-effectiveness analysis (adapted from Table 53 of CS)

State	Utility value	95% confidence interval	Justification
Progression-free survival			
Entrectinib	[redacted]	[redacted]	Utility derived from clinical trial and valued according to UK societal preferences
Established management	0.73	Applied at individual tumour level	Weighted average of tumour-specific utilities
Progressed disease			
Entrectinib	0.59	Applied at individual tumour level	Assumption of equivalent progressed utility to comparator
Established management	0.59	Applied at individual tumour level	Weighted average of tumour-specific utilities

5.2.7.1 Utility values for entrectinib

EQ-5D-3L data were collected in the STARTRK-2 trial, from which 51 of the 54 integrated analysis population patients came. EQ-5D assessments were collected from 44 of the 51 STARTRK-2 patients across nine tumour types (the specific tumour types included in the analysis not reported in the CS). Patients completed the EQ-5D-3L questionnaires at baseline, on Day 1 of each subsequent treatment cycle of 28 days thereafter, at the end of treatment visit, and in the period after treatment.

The company estimated the mean utility value for the progression-free and progressive disease health states, based on 409 and 44 observations respectively (Table 32). Completion rates of the questionnaire reduced over the course of the trial: completion rates fell below 50% after cycle 10 of the trial, and 22.9% of patients completed the End of Treatment questionnaire. A model was not fit to the EQ-5D data available post progression given the limited number of observations, and the data was

not felt to be plausible as it provided a higher health state utility than that of the pre-progression health state.

Table 32 Mean utility estimates for entrectinib (Table 50 in CS)

State	Number of Observations	Mean	Minimum	Maximum	Median
Baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pre-progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

A linear mixed model was fitted to the pre-progression EQ-5D data, adjusting for sex, tumour type and age, and accounting for the repeated observations per subject. The final model resulted in an estimate for utility of 0.8119 (0.76, 0.86). Results from a model with a nested random effect by patient within tumour type were used in the model base-case, and resulted in a utility of [redacted]. This makes the assumption that tumours were randomly sampled from a population of possible tumours and that patients were then sampled randomly from within this tumour pool.

Due to the small sample size and associated uncertainty, the post-progression utility from the integrated efficacy analysis was not used in the economic analysis. The predicted utility was also estimated as being higher than for the progression free health state, which was not considered to be plausible. The company therefore assumed that utility in the PD health state was equal to that of established management.

5.2.7.2 Utility values for established management

A systematic review of utility values undertaken by the company did not identify any studies of utility values specific to an *NTRK* population. The company therefore undertook a search of relevant NICE TAs for appropriate utility values, similar to the approach taken to identify clinical outcomes (see Appendix H in CS), In contrast with the approach taken for the comparator efficacy, where a range of estimates for each tumour type were pooled, the utility values extracted for each tumour type were obtained from a single selected TA. Table 33 reports the selected utility value for each tumour type along with associated uncertainty parameters. Note the company reported that the standard error for each utility estimate was often not available in the source document. Where no standard error was reported, a common arbitrary standard error of 0.14 was used for these estimates (i.e. for colorectal cancer). For PPS, where a pooled utility was estimated, a common standard error was estimated by averaging the standard error for each tumour type.

Table 33 Utility sources for comparator tumour types (Table 51 in CS)

Tumour type	N	Utility estimate – PFS	Measure of uncertainty (SE)	Utility estimate – PPS	Measure of uncertainty (SE)	Source
Colorectal cancer	4	0.73	0.14	0.64	0.14	TA40[59]5
MASC	7	0.725	0.14	0.60	0.14	Assumption: average of known
Thyroid cancer (papillary and anaplastic)	5	0.72	0.14	0.64	0.14	TA535[60]
Non-small-cell lung cancer (squamous and non-squamous)	10	0.74	0.18	0.59	0.06	TA428 [61]
Pancreatic cancer	3	0.70	0.14	0.65	0.14	TA476 [52]
Sarcoma	13	0.72	0.14	0.56	0.14	TA465 [62]
Neuroendocrine tumours	3	0.767	0.14	0.725	0.14	TA539[55]
Breast cancer (including secretory)	6	0.705	0.14	0.496	0.14	TA515[50]
Other (average of known)	3	0.725	0.14	0.65	0.14	Assumption: average of known
Weighted average		0.73		0.59		Calculation

5.2.7.3 Adverse event disutilities

The assumption that any disutility has already been incorporated into the base case health state utilities through trial derived EQ-5D utilities, and incorporating an additional disutility may be considered double counting.

The impact of including disutilities of selected adverse events was explored in a scenario analysis conducted by the company, and found that these had a minimal impact on model results.

ERG Comment

Mapping of EQ-5D data

The ERG considers that the use of a linear mixed model is appropriate for analysing the EQ-5D data and it makes assumptions regarding the random sampling from a population of possible tumours that are consistent with those used in the Bayesian hierarchical modelling presented in Section 4.3.1. However, it was not clear how the variables in the linear mixed model were selected and it was considered that the company could have adjusted for additional characteristics that are likely to

impact on HRQoL, such as CNS metastases or line of therapy, as well as the baseline utility to control for differences between patients.

HRQoL benefit for entrectinib patients

The company base-case analysis applies a higher health state utility value for patients receiving entrectinib in the PFS health state compared with those receiving established management. The company justify this difference stating that "entrectinib is an oral TKI therapy with a more convenient administration and relatively tolerable safety profile when compared with traditional cytotoxic chemotherapies, which form the majority of comparator products". However, as discussed in Section 5.2.7, many of the comparators consist of BSC (no active therapy); in such cases, adverse event rates for patients receiving entrectinib may exceed those of established management and no HRQoL impacts of drug administration would apply.

On balance the ERG considers the assumed quality of life benefit in the pre-progression state to be plausible and reasonable given the safety profile of entrectinib, but is concerned about the lack of evidence to justify the assumption of differential quality of life pre- progression and considers there to considerable uncertainty regarding the magnitude of any difference. The ERG further notes that this assumption is not a significant driver of cost-effectiveness as demonstrated in a scenario presented by the company where the utility value was lowered to that of the comparator arm.

HRQoL of patients on established management

With respect to comparator utilities, the ERG is satisfied that the estimates appear reasonable and comparable with other advanced cancers, but is unable to individually verify each individual utility estimate used given the limited time and resource available to the ERG. The ERG, however, does consider there to be a degree of uncertainty in the provided estimates and notes limitations with the company's approach to selecting utility values. Specifically, the ERG notes the inconsistency in approach between data used to populate the effectiveness of comparator therapies and that used to identify utilities. The impact of this inconsistency is not fully clear, but may impact on estimated utilities as the selected utilities will reflect a specific line of therapy, which may not be directly comparable to the equivalent entrectinib patient. In this respect, the ERG notes that data provided by the company in response to the ERG's clarification request highlighted that entrectinib was often given at earlier lines of therapy compared with the sources of data selected to represent established management (see Section 5.2.6). Given that patients in earlier lines of therapy may have better HRQoL, this may lead to the difference in HRQoL between the two arms being overestimated, biasing the cost-effectiveness analysis in favour of entrectinib.

With regards to the decision not to model AE-related disutilities, the ERG considers that it is a reasonable assumption for entrectinib, since the HRQoL data collected in the STARTRK-2 trial is likely to capture any effects relating to events. It is not possible to determine whether this is the case in the comparator arm without reviewing each individual utility estimate in detail, but the impact of inappropriately excluding a disutility is unlikely to be a driver of the economic analysis.

5.2.8 Resources and costs

The CS (Appendix D, Page 6) describes the search strategies used to identify studies of resource use and treatment costs. The costs included in the model comprised drug acquisition, administration and monitoring for entrectinib and the estimated comparator. Unit costs were sourced from the British National Formulary, NHS reference costs 2017-2018 and the Personal Social Services Research Unit (PSSRU).[63] Costs also included *NTRK* fusion screening costs obtained from previous NICE technology appraisals[64], The Scottish Science Advisory Council [65]and inputs from NHS genomic laboratories.

5.2.8.1 Treatment acquisition cost – entrectinib

Table 34 presents the treatment acquisition costs and drug dosing schedules included in the company's base case analysis. The acquisition cost for entrectinib includes the agreed simple PAS. The dosing intensity applied in the model was 100%, which is higher than the observed dosing intensity taken as an average across all of the trials (96.6%).

Table 34 Entrectinib drug acquisition costs

Drug	Pack concentration	Pack volume	Dose per pack	Cost (£)/pack	Source
Entrectinib	100 mg	30	3,000 mg	[redacted](860.00)	[redacted] (and list) price
Entrectinib	200 mg	90	18,000 mg	[redacted](5,160.00)	[redacted] (and list) price

ERG Comment

The ERG accepts the dosing intensity assumed by the company. Although this is an overestimate of the dosing intensity as observed in the trial, the ERG considers that the applied 100% dose intensity is a reasonable though potentially conservative assumption. The ERG, however, notes that the base-case model presented in the CS fails to account for drug wastage due to discontinuation of therapy. The ERG considers this to be unrealistic because once a pack of tablets has been started these would not be reused should the patient discontinue therapy part-way through a pack. The impact of adding drug wastage for entrectinib is to increase drug acquisition for entrectinib and hence to increase the ICER.

5.2.8.2 Treatment acquisition cost – established management

Drug acquisition costs, dosing frequency, and route of administration information for the interventions forming the established management comparator were obtained from the British National Formulary (BNF), see Table 36 below. The CS does not include the confidential PAS schemes, which have been approved for eribulin, everolimus, nab-paclitaxel, nintedanib, trabectedin and trifluridine-tipiracil. Details of these confidential PAS schemes were made available to the ERG and have been incorporated into the analysis presented in a confidential appendix.

Due to the weighted average approach used to construct the established management arm (as discussed in Section 5.2.4), the CS provides an average of the monthly acquisition costs for each identified comparator for the given tumour type. This can be seen in Table 35.

Table 35 Tumour-specific monthly drug acquisition – average by tumour type

Tumour type	Cost per month
Colorectal cancer	£1,878.09
MASC	£0
Thyroid cancer (papillary and anaplastic)	£0
Non-small-cell lung cancer (squamous and non-squamous)	£1952.05
Pancreatic cancer	£1,507.37
Sarcoma	£3,096.16
Neuroendocrine tumours	£1,354.32
Breast cancer (including secretory)	£1,178.76
Other (average of known)	£1,281.60
MASC, mammary analogue secretory carcinoma	

ERG's comment

Table 36 presents the drug acquisition cost data and drug dosing schedules included in the company's base-case analysis. Table 36 also includes drug acquisition costs obtained from the electronic market information tool (eMIT). [66] This provides information on the average price paid by the NHS for pharmaceuticals, which can differ from the list prices listed in the BNF and is seen as a more accurate and up to date indicator of costs. The ERG considers eMIT to be a more appropriate source of drug acquisition costs and conducts scenario analysis presented in Section 6 using these values.

Table 36 Individual comparator acquisition costs (adapted from Table 55, pg. 115 of CS)

Treatments	Composition (mL or tabs)	Cycle length	Dose per cycle	*BNF Cost*	*eMIT cost*

Capecitabine	150mg/tablet	2 weeks	1250 mg/m^2	*£30.00*	£8.15
Eribulin	0.88mg/2ml	3 weeks	2.26 mg/m^2	£361.00	NR
Vinerolbine	10mg/1ml	Weekly	25-30 mg/m^2	£29.00	£35.83
Gemcitabine	1g/10ml	3 weeks	2500 mg/m^2	£13.09	£8.66
Paclitaxel	100mg/16.7ml	3 weeks	175 mg/m^2	£200.35	£9.49
Docetaxel	20mg/ml	3 weeks	75 mg/m^2	£91.51	£11.61
Irinotecan	40mg/2ml	2 weeks	180 mg/m^2	£39.38	£3.19
Folinic acid	50mg/5ml	2 weeks	500 mg/m^2	£20.00	NR
Fluorouracil	500mg/10ml	2 weeks	12 mg/kg	£6.08	£0.97
5FU	2.5g/50ml	2 weeks	600 mg/m^2	£32.00	NR
Oxaliplatin	50mg/10ml	2 weeks	85 mg/m^2	£155.00	£4.32
Trifluridine-tipiracil	15mg	4 weeks	700 mg/m^2	£500.00	NR
Everolimus	10mg/tablet	Daily	10 mg	£2,673.00	NR
Nab-paclitaxel	100mg	4 weeks	375 mg/m^2	£246.00	NR
Gemcitabine (combination)	1g/10ml	3 weeks	*1000 mg/m2*	£13.00	£8.66
Leucovorin	100mg/10ml	2 weeks	200 mg/m^2	£37.50	NR
Lenvatinib	24	Daily	*24 mg*	£47.90	NR
Sorafenib	200	Daily	*800 mg*	£3,576.56	NR
Doxorubicin	200mg/100ml	3 weeks	60 – 75 mg/m^2	£391.40	£15.59
Ifosfamide	1g	3 weeks	*5-6g/m2*	£115.79	NR
Trabectedin	0.25mg	3 weeks	1.5 mg/m^2	£363.00	NR
Pegylated liposomal doxorubicin	20mg/10ml	*4 weeks*	*50mg/m2*	£360.23	NR
Carboplatin	50mg/vial	*3 weeks*	AUC 5–6 IV	£20.00	£3.59
Nintedanib	100mg/tablet	3 weeks	8000 mg	£2,151.10	NR
BNF, British National Formulary; eMIT, electronic market information tool; 5FU, 5-fluorouracil; NR, not reported					

The ERG is concerned that the monthly drug acquisition costs of the current standard of care for gynaecological cancers and cholangiocarcinoma have not been estimated for the individual tumour types. Rather, the costs associated with these tumour types are an average of colorectal cancer, MASC, thyroid cancer, NSCLC, pancreatic cancer, sarcoma, neuroendocrine tumours and breast cancer. It is unclear whether this reflects costs of relevant therapies for these tumour types, drug

acquisition costs of the comparator therapies were, however, not a major driver of cost effectiveness and therefore the ERG does not consider this issue further.

5.2.8.3 Treatment administration costs

The company state that in order to estimate health state costs across a range of chemotherapy types, it was necessary to apply a simplifying assumption that treatments with similar routes of administration are likely to be associated with similar routine healthcare costs across the different tumour types. The company grouped the interventions into three administration classes: oral, simple IV and complex IV. Each class is associated with an average monthly administration cost, which is applied to each intervention.

The NICE TAs used to inform the costs for each of these classes were TA515 (Eribulin for treating locally advanced or metastatic breast cancer after 1 chemotherapy regimen) [50] for oral chemotherapy, TA520 (Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy)[67] for simple IV chemotherapy and TA476 (Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer) [52] for complex IV chemotherapy.

The oral therapies, including entrectinib, have a monthly cost of £14.59, the simple IV interventions have a monthly cost of £331.69 and the complex IV therapies have a monthly cost of £488.12. The majority of comparators were simple IV chemotherapy or complex IV chemotherapy.

ERG Comment

The ERG is concerned about the appropriateness of the simplifying assumptions made by the company and notes that within categories that infusion time varies significantly. For example, eribulin is classed as a simple IV therapy with infusion time of 2-5 minutes, while trabectidin, also classed as a simple IV therapy, is administered over a period of 24 hours. Similar inconsistencies are seen in the complex category. For example, vinorelbine is infused over a period of 6-10 minutes whereas FOLFIRI is administered as irinotecan infused over 60-90 minutes, folinic acid infused over 2 hours, fluorouracil infused as a bolus and 5FU infused over 46 hours.

The ERG considers that the simplifying assumptions made by the company were not necessary and that individual administration costs for each of the comparator interventions could have been sought. It is unclear whether this approach under or overestimates the costs, however the result is increased uncertainty in the administration costs. Due to the resource required to implement this, the ERG was unable to address this issue in our additional analyses.

5.2.8.4 Health state costs

The three health states in the model are: progression free, progressed disease and death – the model includes those costs associated with patients being in each of these health states.

Progression-free costs

The costs applied in the model in the PFS state are differentiated by the same three classes as the treatment administration costs. These are oral, simple IV and complex IV. The modelled unit costs and resources in each of these classes can be seen in Table 37.

Table 37 Progression free health state costs

Progression free health state costs: Oral treatment					
Item	**Number used**	**% of patients**	**Unit cost**	**Monthly cost**	**Reference***
Medical oncology, outpatient visit	1	100	£162.05	£162.05	Total Outpatient Attendances -row 370 (Outpatient, consultant-led)
GP surgery visit	1	100	£37.40	£37.40	PSSRU 2018 [63]: 10.3b GP unit costs (9.22 minutes patient time)
CT scan	1	33	£132.75	£44.21	RD22Z (CT scan of one area, pre- and post-contrast)
Full blood count	1.55	100	£2.51	£3.89	DAPS05 (haematology)
Liver function tests	1.66	100	£1.11	£1.84	DAPS04 (clinical biochemistry)
Progression free health state costs: Simple IV treatment					
Medical oncology, outpatient visit	1.33	100	£162.05	£215.53	Total Outpatient Attendances -row 370 (Outpatient, consultant-led)
GP surgery visit	1	100	£37.40	£37.40	PSSRU 2018[63]: 10.3b GP unit costs (9.22 minutes patient time)
CT scan	1	33	£132.75	£44.21	RD22Z (CT scan of one area, pre- and post-contrast)
Full blood count	1.55	100	£2.51	£3.89	DAPS05 (haematology)
Liver function tests	1.66	100	£1.11	£1.84	DAPS04 (clinical biochemistry)
Progression free health state costs: Complex IV treatment					
Medical oncology, outpatient visit	1.33	100	£162.05	£215.53	Total Outpatient Attendances -row 370 (Outpatient, consultant-led)
Medical oncology,	1	50	£104.00	£52.00	Total Outpatient Attendances -row 370 (Outpatient,

outpatient, nurse-led					non-consultant-led)
GP surgery visit	1	100	£37.40	£37.40	PSSRU 2018 [63]: 10.3b GP unit costs (9.22 minutes patient time)
Nurse community visit	1	50	£42.00	£21.00	PSSRU 2018 [63]: 10.2 GP practice nurse unit costs (assumes 1 hour patient contact)
CT scan	1	33	£132.75	£44.21	RD22Z (CT scan of one area, pre- and post-contrast)
Full blood count	1.55	100	£2.51	£3.89	DAPS05 (haematology)
Liver function tests	1.66	100	£1.11	£1.84	DAPS04 (clinical biochemistry)

As with administration costs, the resources used in each of the classes were taken from three TAs identified in the company's search for comparator data. These appraisals are TA515 (eribulin for treating locally advanced or metastatic breast cancer after 1 chemotherapy regimen) [50] for oral chemotherapy, TA520 (atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy)[68] for simple IV chemotherapy and TA476 (Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer) [52] for complex IV chemotherapy. Monthly HCRU costs for entrectinib were assumed to be entirely associated with those of an oral therapy. The company states that a clinical expert validated the resources used in these appraisals as being generalisable to the tumour types covered in this appraisal. Unit costs were sourced from the 2017-18 NHS Reference Costs and the PSSRU 2018.[63]

ERG Comment

The ERG has concerns about the company's approach to modelling costs in the PFS state based on the type of therapy received by the patients. The company's approach is an oversimplification of the costs associated with the care in different tumour types and as a result, there is significant uncertainty in the modelled cost inputs. The approach taken by the company would suggest that the cost associated with treating a patient with NSCLC and neuroendocrine with oral therapies in the PFS health state are identical. As with the administration costs, the ERG considers that the simplifying assumption of grouping therapies into one of three classes was not necessary and that tumour specific health state costs could have been sought. As with administration costs it is unclear whether this approach under or overestimates the costs, however, the result is increased uncertainty in the administration costs.

However, as the influence of administration costs on the ICER is minimal, the ERG does not consider this issue further.

Progressed disease costs

The costs applied in the model for the PPS state are the same across both entrectinib patients and comparator patients. These costs are presented in Table 38.

Table 38 Progressed disease health state costs (Table 63, page 122 of the CS)

Item	Number used	% of patients	Unit cost	Monthly cost	Reference
Medical oncology, outpatient visit	1	100	£162.05	£162.05	Total Outpatient Attendances - row 370 (Outpatient, consultant-led)
Medical oncology, outpatient, nurse-led	1	100	£104.00	£104.00	Total Outpatient Attendances - row 370 (Outpatient, non consultant-led)
GP home visit	1	100	£37.40	£37.40	PSSRU 2018[63]: 10.3b GP unit costs (9.22 minutes patient time)
Nurse community visit	1	67	£42.00	£30.15	PSSRU 2018 [63]: 10.2 GP practice nurse unit costs (assumes 1 hour patient contact)

ERG Comment

The ERG has no major concerns with the unit costs of the items included in the progressed-disease health state included in the model. The ERG, however, notes that some additional costs could have been included to accurately reflect the care received by patients with cancer in a progressed disease state. These include medication costs (e.g. steroids, NSAIDS, morphine, bisphosphonates and dietary supplementation), and tests and procedure costs (e.g. full blood count, serum chemistry, CT scan, home oxygen and x-ray). The omission of these costs is likely to be small, but will lead to an underestimation of the ICER.

5.2.8.5 End of life costs

The CS model calculated a one-off cost to account for terminal care incurred. The model applied this at the transition from progressed disease to death. The costs were obtained from Georghiou and Bardsley [69], adjusted for inflation to 2017-2018. These costs can be seen in Table 39.

Table 39 Summary of components of end of life costs

Component	Mean cost, last 3 months (2017 – 2018)	Mean cost/month, last 3 months (2017 – 2018)
Emergency inpatient admission	£4049.29	£1349.76
Non-emergency inpatient admission	£1352.75	£450.92
Outpatient attendance	£375.98	£125.33
A&E visits	£79.57	£26.52
Social care	£441.63	£147.21
District nursing care	£584.86	£194.95
GP visits	£363.05	£121.02

ERG Comment

The ERG has no major concerns with the end of life costs as these are common to both groups, and because virtually all participants die within the time horizon, the only differences in these costs between the two treatment groups are as a result of discounting.

5.2.8.6 Adverse event costs

The company only included those adverse events occurring at a rate of ≥5% in the model. All adverse events except increased weight were considered to occur at the same rate for both entrectinib and comparator patients. The adverse events that were included in the model were: anaemia, fatigue, neutropenia and weight increase.

ERG Comment

The ERG did not identify any areas of concern regarding the company's choice of adverse events to include in the model. The ERG, however, notes that event rates were assumed equal for all AE except weight gain which had a zero cost. As such, effectively no costs of AEs were included in the model. As outlined, in Section 5.2.6.5, the ERG considers that the company could have modelled AE for comparator therapies, though it is acknowledged that this is likely to have only a small effect on the estimated ICER.

5.2.8.7 *NTRK*-fusion screening costs

The company included a projected cost for screening eligible patients in the base-case analysis. In the model, the company proposed a hierarchical approach in which IHC is conducted to identify patients with tumours expressing *NTRK* protein, followed by confirmatory testing with an NGS panel to establish whether these patients have specific *NTRK* gene fusions.

To calculate the *NTRK*-fusion screening costs, the company estimated the number-needed-to-screen for each of the tumour types represented in the integrated efficacy analysis (Table 1). The company then identified the screening tests conducted in current practice within the NHS based on the NHS Genomic Testing Directory (Table 40).

Table 40 Frequency of NTRK fusions in enrolled tumour types (Adapted from CS, Table 66, p 125)

Tumour type	NTRK-fusions rate	Number needed to screen	Current Testing
CRC	[redacted]	[redacted]	Other biomarker screening
NSCLC (squamous and non-squamous)	[redacted]	[redacted]	Other biomarker screening
Pancreatic	[redacted]	[redacted]	No molecular testing within directory
Non-secretory breast cancer	[redacted]	[redacted]	Other biomarker screening
Secretory Breast Carcinoma (0.02% HER2-)	[redacted]	[redacted]	Other biomarker screening
Thyroid (papillary/anaplastic)	[redacted]	[redacted]	Other biomarker screening
Neuroendocrine tumours	[redacted]	[redacted]	No molecular testing within directory
Sarcoma (non-paediatric)	[redacted]	[redacted]	WGS
MASC	[redacted]	[redacted]	NTRK-fusion testing
Other	[redacted]	[redacted]	No molecular testing within directory
Paediatric cancers	[redacted]	[redacted]	WGS

CRC, colorectal cancer; NSCLC, non-small cell lung cancer; MASC, mammary analogue secretory carcinoma; WGS, whole genome sequencing; NR, not reported.

In tumour types where a genetic test is already conducted in clinical practice ('other biomarker screening' in Table 40), the company assumed the unit cost of a standard IHC test to be £75.00. In the base-case analysis, the company assumed the cost of screening for the comparators in which genetic testing already occurs, to be the number-needed-to-screen (NNS) (Table 40) multiplied by the cost of the IHC test. In the entrectinib arm, the screening cost was assumed to be the same screening cost calculated for the comparator arm with an additional cost of adding a confirmatory NGS test for 11%

of the patients IHC tested. The unit cost of an NGS test applied in the base case was [redacted] per test. The proportion of patients receiving NGS was based on clinical data provided by an investigator involved in the entrectinib clinical development programme, which suggested that the IHC testing approach will remove 89% of *NTRK*-fusion negative samples. In those tumour types where no genetic testing is currently conducted in clinical practice ('no molecular testing in clinical practice' in Table 40), no screening costs were attributed to the comparator arm in the base case. In these tumour types, the screening costs in the base case for the entrectinib arm was the cost of an IHC test multiplied by the NNS, plus an additional cost a confirmatory NGS test for 11% of IHC tested patients.

In their base case, the company assumed that MASC patients were diagnosed by IHC alone in line with current testing for these patients. As *NTRK* fusions are already included in the Genomic Testing Directory for MASC, this cost was applied in entrectinib and established management arms of the model. For those tumours in which WGS is reimbursed for specific tumour types (paediatric tumours and sarcoma), the company assumed that *NTRK* fusion positive patients would be identified via current testing practice. A unit cost of screening for *NTRK*-fusions is £800.00 per test per patient tested was therefore applied in both entrectinib and established management arms of the model.

The costs of screening each tumour type to identify one entrectinib-eligible patient are shown in Table 41. The model used a weighted average of the costs for each tumour type, weighted by the number of patients with that specific tumour type in the efficacy evaluable population. In the model, the average incremental cost of screening for the entrectinib arm over the comparator arm was an additional £15,828.

Table 41 Costs of screening by tumour type to identify one patient used in the base case (Adapted from CS, Table 69, p 127)

Tumour type	Base case: entrectinib	Base case: comparator
CRC	[redacted]	[redacted]
NSCLC (squamous and non-squamous)	[redacted]	[redacted]
Pancreatic	[redacted]	[redacted]
Non-secretory breast cancer	[redacted]	[redacted]
Secretory Breast cancer	[redacted]	[redacted]
Thyroid (papillary/anaplastic)	[redacted]	[redacted]
Neuroendocrine tumours	[redacted]	[redacted]
Sarcoma (non-paediatric)	[redacted]	[redacted]
MASC	[redacted]	[redacted]
Other	[redacted]	[redacted]

Weighted average within integrated analysis	[redacted]	[redacted]

CRC, colorectal cancer; NSCLC, non-small cell lung cancer; MASC, mammary analogue secretory carcinoma

In addition to the above the, company also presented scenario analyses in which 50% and 25% of the costs of screening are applied to entrectinib to represent scenarios in which 2 or 4 *NTRK* fusion-targeting medicines are available, respectively.

ERG Comment

There are a significant number of uncertainties in estimating appropriate testing costs. These relate to the testing strategy adopted, unit costs applied, feasibility and provision of current services.

Testing Strategy

The ERG considers that the company's proposed, hierarchical approach to testing to be a plausible strategy to identify individuals with an *NTRK* fusion. IHC is high-throughput and is inexpensive, making it a practical screening tool to use in a large population. The diagnostic accuracy of IHC is, however, variable. IHC has a low sensitivity for tumours expressing the *NTRK*3 fusion and high rates of false negatives in smooth muscle and neural tumours. Implementation of this strategy would therefore mean that a proportion of *NTRK* fusion positive patients are likely to be missed. See Section 2.2.2.2 for further discussion of limitations of this approach.

The company's assumption that *NTRK*-fusion positive paediatric and adult sarcoma patients would be identified under established pathways as WGS is disputed by the ERG. The ERGs clinical advisers stated RNA-based NGS would be needed after WGS to confirm an *NTRK*-fusion positive tumour. This will require the entrectinib testing costs to include an additional RNA-based NGS cost for all of the paediatric and sarcoma patients identified through WGS. The effects of this additional cost on the company's base case will increase the costs associated with the entrectinib arm and is also explored in scenario analysis presented in Section 6.3.2.

As noted in Section 2.2.2, RNA-based NGS fusion panels are available on the NHS for a specific subgroup of patients with NSCLC, targeting a range of genes including *EGFR*, *ALK*, and *ROS1*. Whilst this panel does not currently target *NTRK1-3* rearrangements, genomic advisers informed the ERG that the costs of adding additional gene targets to an RNA-based NGS panel are nominal. Incremental costs associated with the identification of *NTRK* fusion patients with NSCLC may therefore be close to zero. Scenario analysis is implemented in Section 6.3.2 evaluating the impact of removing testing costs for NSCLC patients.

As described in section 2.2.2, there are several other strategies that could be adopted to detect *NTRK* fusions. One potential approach would be to offer NGS as a first line test to identify *NTRK* fusion positive patients, with or without confirmatory IHC. Given the low prevalence of *NTRK* fusions in most tumour types, and the large population of individuals to be tested, a primarily NGS-based testing strategy may be impractical because NGS is more expensive and time-consuming. Costs of NGS and resources required to implement it are falling over time, potentially making more plausible in the future. The advantage of this approach is that NGS has high diagnostic accuracy, consequently, it is less likely that patients with an *NTRK* fusion will be missed.

An alternative strategy outlined is also outlined in recent guidelines published by the ESMO Translational Research and Precision Medicine Working Group.[17] This approach suggests that the testing pathway for detecting *NTRK* fusion positive patients should vary depending on the frequency of *NTRK* fusion and current availability of testing for each tumour type.

- For tumours with a high frequency of *NTRK* positive patients (e.g. MASC & infantile fibrosarcoma), FISH or a targeted NGS panel should be used.
- In the tumour types where genomic testing is currently available (e.g. NSCLC and colorectal cancer), NGS should be used as a first line testing approach.
- For the tumour types where no genomic testing is available (e.g. pancreatic cancer), IHC followed by confirmatory NGS is recommended.

The ESMO recommendations take advantage of current testing availability, and therefore may be an efficient approach with potentially lower incremental testing costs. In the future, the expansion of genomic testing will allow first-line NGS to be used for a wider range of tumour types, with advantages in terms of testing sensitivity.

A limitation of this approach is that current testing may not be widely available for patients within each tumour type, and particular eligibility criteria is likely to limit the number of individuals who would be able to access testing. This may limit the range of tumour sites where first-line NGS will be viable option. A further issue with this approach is that NGS tests currently used on the NHS are primarily DNA-based, which as outlined in Section 2.2.2 have limited sensitivity to structural rearrangements. Implementing the ESMO guidelines would therefore require switching to RNA-based NGS. This might have implications on costs because RNA-based NGS is more expensive. Nonetheless, with the increased use of RNA-based NGS in clinical settings, the ESMO approach for using NGS will become more practical in the future, as fusion testing can easily be added to current testing panels for nominal costs.

Comparator Testing Costs

The ERG considers the inclusion of screening costs included in the comparator arm, as seen in Table 41, to be inappropriate. The focus of testing costs should be on the incremental testing costs associated with identifying *NTRK* fusion positive patients and therefore the ERG suggests the comparator testing costs should be removed unless current testing is able to identify *NTRK* fusions. In current practice, the NHS only reimburses *NTRK*-fusion screening assays for MASC patients. This may be subject to change over time, as molecular testing is expanded, however it is the ERG's view that decisions should be made on current practice not a possible hypothetical future. The ERG has concerns regarding the validity of the scenarios presented by the company in which testing costs are shared across two and four *NTRK* targeting therapies. The base-case analysis should represent the incremental costs of implementing entrectinib in the NHS, not the implementation of a range of hypothetical agents that may or may not be available in the future. Testing costs are significant drivers of cost-effectiveness and the implications of removing these testing costs from the comparator arm will be explored further in Section 6.3.2

Unit Costs Applied

The IHC testing unit cost was assumed to be £75.00 based on a pathologist's input obtained during a NICE committee meeting to discuss the appraisal of crizotinib for the treatment of NSCLC. In this appraisal, the cost of IHC was identified to be between £50 and £100 excluding laboratory costs. The ERG is concerned that the £75.00 unit cost applied in the CS underestimates the unit cost of IHC testing as this estimate does not include laboratory costs. Further, the ERG considers that the marginal cost of implementing IHC screening is likely to vary depending upon whether current service provision already supports the regular use of molecular testing (for further information on tumour types in which molecular testing is used see Section 2.2.2). This reflects the fact in tumour types where no testing is currently implemented will require greater investment in infrastructure, as well as additional marginal costs associated with the administration of testing. This may include costs of obtaining tissue samples, postage, and clinician time associated with interpretation of test results.

The company assumes that the cost of NGS testing is [redacted]. The ERG considers the cost reasonable and in line with clinical advice received by the ERG. The ERG, however, notes that the majority of NGS currently available on the NHS is DNA-based, which is unsuitable for testing *NTRK* fusions due to poor diagnostic accuracy. Implementation of testing regimens based around a first line NGS would therefore require the adoption of either RNA-based or hybrid DNA/RNA based NGS which is more expensive than DNA tests. Implementation of an NGS-centred testing strategy would therefore potentially incur additional costs, even where current testing includes NGS.

The ERG further notes the testing regimen proposed by the company requires the implementation of either RNA-based or DNA/RNA hybrid NGS panels as a confirmatory test. Due to its labile nature,

RNA can be easily damaged during tissue handling and preparation of the sample, most damaged samples can be detected in pre-screening tests conducted before administration of the NGS panel. However, in highly contaminated RNA, damage is not detected until the after assay has been used. Therefore, a new test is required to detect an *NTRK* fusion, and hence, increasing the cost. This additional potential cost is not accounted for in the company's calculations and therefore applied costs are likely to be underestimates.

Number Needed to Screen

In order to calculate the costs associated with testing, the company calculated the number of individuals that would need to be screened in order to identify one individual with an *NTRK* fusion. The CS reports that the NNS was estimated based on prevalence of *NTRK* figures reported in Amatu *et al* [2]and data on file. [70] The ERG, however notes that the reported estimates of NNS differ from the ERG's preferred estimates based on the FMI database which recorded the frequency of *NTRK* fusion positive patients from a sample circa 166,000 samples. The NSS to screen for each tumour type estimated by the ERG and by the company is presented in Table 42.

As can be seen from Table 42, there are a significant number of inconsistencies in the estimated NNS between the CS and ERG. For example, in pancreatic cancer the CS estimates the NNS as [redacted], while the ERG estimates a figure of [redacted]. The impact of these differences in NNS is significant, affecting both the average NNS across all sites, as well as tumour type-specific testing costs.

In addition to the above, the ERG notes that the company's estimates of the NNS, and by extension average testing costs are based upon the distribution of the 13 tumour sites represented in the integrated efficacy analysis. This is problematic for two reasons. Firstly, as discussed in section 5.2.3 the distribution of tumour types in the integrated efficacy analysis is unlikely to represent the distribution in practice with some sites over represented and other underrepresented. Secondly, the modelled population includes only 13 tumour types and excludes a number of tumour types in which it is known that *NTRK* fusions occur; NNS for unrepresented tumour types ranges from [redacted] (Infantile Fibrosarcoma) to [redacted] (Cervix Cancer). Estimated costs of testing informing the company's base case analysis are therefore unlikely to represent testing costs for the population eligible for entrectinib. The ERG implements scenario analysis in section exploring both of these issues using the ERG's preferred estimates of NNS.

Table 42 Number needed to screen by tumour type

Tumour Type	Prevalence of *NTRK* fusion (ERG)	Number Needed to Screen (ERG)	Number Needed to Screen (Company)
Salivary gland (MASC)	[redacted]	[redacted]	[redacted]
NSCLC	[redacted]	[redacted]	[redacted]
Breast cancer (not specified)	[redacted]	[redacted]	[redacted]
Secretory breast carcinoma	[redacted]	[redacted]	[redacted]
Papillary thyroid tumour	[redacted]	[redacted]	[redacted]
Thyroid Tumour (NOS)	[redacted]	[redacted]	[redacted]
Colon/colorectal	[redacted]	[redacted]	[redacted]
Neuroendocrine (NOS)	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]
Pancreatic	[redacted]	[redacted]	[redacted]
Uterine	[redacted]	[redacted]	[redacted]
Ovarian	[redacted]	[redacted]	[redacted]
Cervix	[redacted]	[redacted]	[redacted]
Soft tissue sarcoma	[redacted]	[redacted]	[redacted]
High grade glioma	[redacted]	[redacted]	[redacted]
Paediatric high grade glioma	[redacted]	[redacted]	[redacted]
Congenital mesoblastic nephroma	[redacted]	[redacted]	[redacted]
Paediatric melanoma	[redacted]	[redacted]	[redacted]
Infantile fibrosarcoma	[redacted]	[redacted]	[redacted]
Paediatric low grade glioma	[redacted]	[redacted]	[redacted]

The impact of different testing approaches and different estimates of the population on the company's base-case ICER will be explored in Section 6.3.2.

Feasibility

The company acknowledge that the requirements for screening for *NTRK* will take into account the likely large economic impact it is expected to have. However, the company do not fully consider the feasibility of implementing additional testing to a large population.

In order to determine the impact of testing for *NTRK,* the ERG calculated the number of individuals that would require IHC and NGS testing, as proposed in the company's testing strategy. As outlined in Section 2.2.2, there are two approaches that can be used to calculate the number of individuals who would require screening. Using a top-down approach, based on the total annual incidence of cancer in England with stage 3/4 tumours, 92,524 individuals would require IHC screening every year. A

further 10,178 would also require confirmatory NGS tests assuming, 11% of individuals who receive IHC will require NGS. The ERG also used a conservative, bottom-up approach to calculate the number of patients who require testing, based on the tumour sites in which there is a known *NTRK* fusion. Under this approach, the ERG estimates that an additional 52,782 IHC would be undertaken each year along with a further 5,806 confirmatory NGS.

These figures represent a significant increase in the number of molecular tests that would be untaken annually and therefore the ERG is concerned that additional investment in current genomic services will be necessary to provide *NTRK* testing across the NHS. This may include expansion of current infrastructure that would be required to meet an increasing number of referrals, additional requirements for a larger workforce to prepare samples and process tests, as well as the need to employ and train additional clinical geneticists and bioinformaticians specialised in genetic fusions and targeted medicines. [32, 34] These costs are, however, not considered in the economic model and as described in Section 2.2.2 may have long term implications on the viability of molecular testing services which are predicted to become increasingly overwhelmed by demand for testing services. See Section 2.2.2 for further discussion of the feasibility of expanding testing services.

5.2.9 Cost effectiveness results

Entrectinib has a confidential patient access scheme (PAS), comprising a simple discount of [redacted]. A number of the interventions that comprise established management also have PAS available. The results presented below include the PAS for entrectinib, but do not include PAS available for comparators, with results including these PAS presented in a confidential appendix to this report.

Table 43 presents the base-case deterministic analysis of entrectinib. It shows that entrectinib was associated with increased costs (cost difference of [redacted]) and was more effective (gain of [redacted] QALYs), compared with established management. The company's base-case ICER was £54,646 per QALY.

Table 43 Base-case results (Adapted from CS, Table 72 and Table 73, p 132)

Technologies	Total Costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£54,646
Established management	£62,931	1.74	1.12	-	-	-	-
CS, company submission; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALYs, quality-adjusted life-years; PAS, patient access scheme							

5.2.9.1 Updated base-case

Following response to clarification questions, the company presented an updated deterministic base case ICER using the latest trial data cut-off for PFS and OS for the integrated analysis population. The previous data cut was from 31st May 2018 and the updated data cut is from [redacted]. The update was implemented along with other requested model corrections suggested by the ERG. The updated results show the ICER of entrectinib compared to the established management comparator is £52,609 and is presented in Table 44.

Table 44 Updated base-case results

Technologies	Total Costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,609
Established management	£63,028	1.74	1.12	-	-	-	-
CS, company submission; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALYs, quality-adjusted life-years; PAS, patient access scheme							

Table 44 shows that entrectinib was associated with increased costs (cost difference of [redacted]) and was more effective (gain of [redacted] QALYs), compared with established management.

5.2.9.2 Sensitivity analyses

Probabilistic sensitivity analysis

The company performed a probabilistic sensitivity analysis (PSA) using Monte-Carlo simulation with 2,000 iterations. In each iteration, the model drew inputs from defined distributions for selected parameters (CS Table 70, Pages 128-129). The probabilistic ICERs were lower than those in the deterministic analysis. Table 45 presents the results of the probabilistic sensitivity analysis.

Table 45 Company probabilistic cost-effectiveness results (including Entrectinib PAS)

Technologies	Total Costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,052
Established management	£64,128	1.74	1.12	-	-	-	-
CS, company submission; ICER, incremental cost-effectiveness ratio; LYG, life-years gained; QALYs, quality-adjusted life-years; PAS, patient access scheme							

The mean probabilistic ICER of entrectinib was £52,052 per QALY gained versus established management. The probability that entrectinib is the most cost-effective treatment option at WTP threshold of £30,000 is [redacted], and [redacted] at £50,000. The cost-effectiveness acceptability curve for all comparators is provided in Figure 18.

Figure 18 Cost-effectiveness acceptability curve for entrectinib and established management (including entrectinib PAS), (CS, Figure 25, pg. 134)

Figure redacted

QALY, quality-adjusted life year; SoC standard of care

The results show there is little difference between the deterministic and probabilistic ICERs. The average incremental QALYs gained with entrectinib compared to established management was [redacted], which was [redacted] QALYs more than in the updated deterministic analysis.

ERG Comment

The ERG has concerns about the uncertainty of the probabilistic ICER included in the CS. The narrow distributions of the comparator costs, total life years gained, and total QALYs appears to be unrealistic and result in a misleading level of confidence in the comparator results. The narrow confidence intervals around the comparator effectiveness results stems from the company not properly accounting from the uncertainty in the comparator effectiveness estimates.

The ERG also has concerns about the standard errors around the survival estimates used to construct the established management comparator. The standard errors are assumed, and have not been extracted from the original sources of the comparator effectiveness, utilities, and costs. For a discussion of the methods used to construct the weighted average comparator, see Section 3.3. In response to clarification questions, the company stated that for the published survival estimates, due to lack of covariance matrices and correlations reported and the use of an exponential model for the extrapolation, the extrapolated mean is varied around a normal distribution. This was to avoid any assumptions on skewness and allow for a normal range of assessments of the uncertainty around these estimates. The ERG is concerned this approach underestimates the uncertainty around the cost-effectiveness results of the comparator, particularly given the issues involved in the method used to construct the weighted average comparator.

Deterministic sensitivity analysis

The company presented a series of deterministic sensitivity analyses (DSA) in the form of univariate sensitivity analyses to assess the impact of varying key model input parameters upon the ICER.

Selection of parameters for inclusion in the analysis was conducted *a priori.* Unless otherwise stated, base case values were adjusted across a +/- 20% range. The DSA inputs are summarised in CS Table 75.

Tornado diagrams summarising the twelve most influential parameters as reported by the company are presented in Figure 19. The results indicate that varying the median OS of the comparator and the weighted screening costs had the greatest impact upon the ICER. The utility of the PFS state in the entrectinib arm was also a driver of the model's results. The DSA did not produce any ICERs less than £40,000/QALY.

Figure 19 Univariate sensitivity analysis for entrectinib vs comparator (CS, Figure 26, pg. 135)

Figure redacted

OS, overall survival; PFS, progression-free survival; HCRU, health care resource utilisation; PPS, post-progression survival

5.2.9.3 Scenario analyses

The submission and clarification response included an extensive series of scenario analyses to assess the robustness of the model results and the impact of the assumptions included in the base-case analysis. The results of the scenario analyses performed on the company's updated base case are presented in Table 46. The results were most sensitive to variations in the parametric function used to extrapolate OS which resulted in a range of ICERs from £37,217 to £81,588 per QALY. For a discussion of the choice of parametric function, see Section 5.2.6. The results were also sensitive to the tumour weighting applied to the comparator. Reweighting the comparator data to be 100% weight applied to MASC and pancreatic comparator outcomes resulted in ICERs of £31,064 and £114,524 per QALY, respectively. For a discussion of the methods and survival data used to construct the comparator, see Section 5.2.4.

Table 46 Scenario analysis results (adapted from CS and clarification response)

Parameter	Value	Incremental Costs	Incremental QALYs	ICER (£/QALY)
Distribution Entrectinib NTRK+ OS	Exponential	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ OS	Weibull	[redacted]	[redacted]	64,149
Distribution Entrectinib NTRK+ OS	Log-normal	[redacted]	[redacted]	37,217
Distribution Entrectinib NTRK+ OS	Gamma	[redacted]	[redacted]	71,383
Distribution Entrectinib NTRK+ OS	Log-logistic	[redacted]	[redacted]	41,509
Distribution Entrectinib NTRK+ OS	Gompertz	[redacted]	[redacted]	81,588

Distribution Entrectinib NTRK+ PFS	Exponential	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ PFS	Weibull	[redacted]	[redacted]	52,463
Distribution Entrectinib NTRK+ PFS	Log-normal	[redacted]	[redacted]	53,571
Distribution Entrectinib NTRK+ PFS	Gamma	[redacted]	[redacted]	52,941
Distribution Entrectinib NTRK+ PFS	Log-logistic	[redacted]	[redacted]	53,566
Distribution Entrectinib NTRK+ PFS	Gompertz	[redacted]	[redacted]	52,570
Distribution Entrectinib NTRK+ TTD	Exponential	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ TTD	Weibull	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ TTD	Log-normal	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ TTD	Gamma	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ TTD	Log-logistic	[redacted]	[redacted]	52,609
Distribution Entrectinib NTRK+ TTD	Gompertz	[redacted]	[redacted]	52,609
Treatment duration assumption	Trial-observed treatment duration	[redacted]	[redacted]	50,838
Treatment duration assumption	According to label	[redacted]	[redacted]	52,609
Time horizon	5	[redacted]	[redacted]	68,849
Time horizon	10	[redacted]	[redacted]	54,807
Time horizon	15	[redacted]	[redacted]	53,011
Time horizon	20	[redacted]	[redacted]	52,684
Time horizon	25	[redacted]	[redacted]	52,621
Time horizon	30	[redacted]	[redacted]	52,609
Screening costs	Base case: 100% attributed to entrectinib	[redacted]	[redacted]	52,609
Screening costs	50% attributed to entrectinib	[redacted]	[redacted]	44,762
Screening costs	25% attribution to entrectinib	[redacted]	[redacted]	40,838
Screening costs	Screening costs excluded	[redacted]	[redacted]	36,914
Prognosis of comparator	Base case: aggregated trial reported outcomes	[redacted]	[redacted]	52,609
Prognosis of comparator	Adjustment to reflect poorer NTRK prognosis (HR= 2.33)	[redacted]	[redacted]	35,589
Prognosis of comparator	Incorporation of CNS	[redacted]	[redacted]	46,981

	metastases (comparator)			
Post-progression therapy	Base case: 0% active treatment for comparator patients; 35% for entrectinib	[redacted]	[redacted]	52,609
Post-progression therapy	0% active treatment for comparator patients; 50% for entrectinib	[redacted]	[redacted]	58,120
Post-progression therapy	0% active treatment for comparator patients; 80% for entrectinib	[redacted]	[redacted]	69,143
Post-progression therapy	Equivalent post-progression treatment (50% each)	[redacted]	[redacted]	54,868
PFS utility	Base case: Entrectinib PFS utility derived from trial data	[redacted]	[redacted]	52,609
PFS utility	Entrectinib PFS utility reduced to match comparator PFS value	[redacted]	[redacted]	59,390
Tumour-weighting	Base case – trial weighting	[redacted]	[redacted]	52,609
Tumour-weighting	100% weight applied to MASC comparator outcomes	[redacted]	[redacted]	31,064
Tumour-weighting	100% weight applied to pancreatic cancer comparator outcomes	[redacted]	[redacted]	114,524
OS, overall survival; PFS, progression-free survival; TTD,				

5.2.9.4 Additional scenarios requested at points for clarification

The company provided an updated economic model in the updated response to clarifications. The updated model includes a scenario whereby the five efficacy-evaluable adult primary CNS tumour patients and the seven paediatric patients have been added to the model, as per the ERG's request. This was requested as the ERG believes these patients with primary CNS tumours and paediatric patients fall within the population in which the company is seeking a recommendation. For a full

discussion of the population included in the economic analysis, see Section 5.2.3. The inclusion of the 12 patients resulted in a decrease in the company's base case ICER to £49,358. These results can be seen in Table 47.

The ERG also had some concerns regarding the company's assumption that a proportion of patients receiving second-line therapy following entrectinib will continue to receive the second-line therapy until death. The company acknowledged that this is a conservative assumption, and that the clinical plausibility of this is low. As a result, two alternative scenarios were provided in which the duration of subsequent therapy is limited to 6 months and 3 months. This reduced the ICER to £40,093 and £39,849, respectively.

Table 47 Additional scenarios following clarification questions

Parameter	Value	Incremental Costs	Incremental QALYs	ICER (£/QALY)
Inclusion of paediatric and CNS	Base case: paediatric and CNS patients excluded	[redacted]	[redacted]	£52,609
Inclusion of paediatric and CNS	Paediatric and CNS patients included	[redacted]	[redacted]	£49,358
Duration of subsequent therapy	Base case: until death	[redacted]	[redacted]	£52,609
Duration of subsequent therapy	6 months	[redacted]	[redacted]	£40,093
Duration of subsequent therapy	3 months	[redacted]	[redacted]	£39,849
ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life-years; CNS, central nervous system				

5.2.10 Model validation and face validity check

The company stated that the cost-effectiveness analysis was validated in a number of ways. The internal validity of the model processes was assessed by an external consultancy company, who undertook a technical validation of the model (including pressure testing using extreme values, formula checking, and cell references). In addition, the validation of entrectinib extrapolations, comparator choice and data, and tumour type proportions presented in the integrated analysis were also described by the company (Section B.3.3.6 of the CS).

The clinical plausibility of the survival curves for entrectinib was discussed with investigators at two UK sites from the STARTRK-2 study, through their visual inspection of all six extrapolations of the PFS and OS curves for entrectinib, with emphasis placed on OS extrapolation due to its importance in the model. No details were provided as to why certain distributions were rejected. These investigators also noted that the frequencies of the tumour types seen may reflect clinical practice, with the possible exception that MASC is overrepresented.

The specific treatment choices for each tumour type were discussed with a clinical expert in each of the following tumour types: non-small cell lung cancer, breast cancer, sarcoma, thyroid cancer, neuroendocrine tumours, colorectal cancer, and pancreatic cancer. The choices of comparator were kept broadly in line with the therapies listed in NICE Pathways, with deviations based on recommendations from clinical experts. The company also stated that their clinical experts endorsed the survival data extracted for each comparator, with the caveat that some comparator OS outcomes were confounded by crossover, and therefore exhibited better-than-expected outcomes where adjusted data could not be found.

In addition to the clinical plausibility of the extrapolated outcomes, selection of the appropriate distributions has been driven by statistical fit to the data, and the company presented a comparison of modelled and trial-based OS and PFS for entrectinib. Modelled PFS appeared to represent the clinical data well throughout the trial period up to around 12 months, after which the modelled PFS appeared higher than the trial PFS, with the degree of this overestimation varying by distribution. The OS data was less mature, and therefore it was more difficult to assess its predictive validity.

As noted in Section 5.2.6, the company's selection of an exponential distribution for modelling OS and PFS resulted in patients remaining in the progression health state longer than the PFS health state (more than twice as long, in the comparator arm). The ERG did not consider this to appear plausible, given the end stage of the pathway at which patients are treated.

5.3 Conclusions of the cost effectiveness section

The modelling of a histology independent indication such as the one covered by the present decision problems generates a number of significant challenges that impact greatly on the validity of the ICERs generated in the company's presented economic analysis. The results of the economic model are therefore subject to very considerable uncertainty and may differ significantly from those presented in the company's base-case. Further, the single ICER presented by the company conceals the potentially significant variation in the tumour specific ICERs, driven by a combination of factors, particularly variability in relative effectiveness between tumour types and testing costs. As stated in Section 5.2.1 it is the ERG's general view that optimised decisions are preferable, and while the ERG acknowledges the challenges presented by the current decision problem, it considers that the company could have gone further in justifying the use of a single ICER. In particular, the ERG suggests the company could have explored variability in the treatment effect across tumour types, and how testing costs are likely to impact on the cost-effectiveness of specific tumour types. An overview of the key uncertainties identified by the ERG are presented below.

1) *Heterogeneity in the treatment effect*

A central issue of the current appraisal is the potential for heterogeneity in the treatment effect across tumour types, as well as across other clinical characteristics such as age (paediatric vs adults), fusion type, and position in the treatment pathway. As demonstrated in the ERG exploratory analyses on response data (see Section 4.3), there is evidence to suggest that the treatment effect is heterogeneous across tumour types. Furthermore, the predictive distribution, which provides an estimate of the likely response rate in an unrepresented tumour has a credible interval of [redacted], implying that mean response across all eligible tumour types could be very different to that estimated in the integrated analysis. This has significant implications for the economic analysis and suggests that the tumour specific ICER will vary significantly.

2) Uncertainties surrounding the comparability of comparator effectiveness evidence

The ERG's has several concerns about the representativeness of the modelled population, which was based on the integrated efficacy analysis. These include concerns about the distribution of tumour types modelled, which appear to over represent some tumour types, while under-representing others. Further, the modelled population includes only the 13 tumour types included in the EEA dataset, while there is evidence to suggest that *NTRK* fusions occur in at least another 11 tumour types, representing a minimum of 20% of the eligible population. The omission of these patients has a number of implications for the model and potentially impacts upon a number of inputs used to model established management, including comparator effectiveness, comparator treatment cost, testing costs, and health state utilities. The ERG is also concerned that the analysed integrated efficacy data set excluded available evidence on patients with [redacted].

3) Uncertainties surrounding the relevance of selected comparators

There are significant uncertainties regarding whether the appropriate comparators have been modelled. The anticipated marketing authorisation for entrectinib allows it to be used at multiple points in the treatment pathway, meaning there is significant uncertainty regarding the patient group in which entrectinib may be used in practice. It is therefore unclear whether the modelled comparators represent current NHS practice. Furthermore, because the model only considers 13 tumour sites and not all tumour sites in which *NTRK* fusions may occur, there are a number of relevant comparators not covered by the model. The model therefore implicitly assumes that the modelled population is representative of the eligible population, which appears to be unlikely given available evidence on the distribution of tumour types with *NTRK* fusions.

4) Uncertainties surrounding the comparability of comparator effectiveness evidence

Because the available effectiveness evidence for entrectinib was from single arm studies, it was necessary to generate an appropriate comparator dataset. The company does this by using previous NICE TAs as a source of effectiveness data, which were then weighted by the distribution of tumour

types in the integrated efficacy analysis. While the ERG considers the broad approach adopted by the company to be reasonable, there are significant challenges associated with implementing this successfully, as well as further issues resulting from the company's execution of this approach.

The ERG's principal concerns regarding the company's approach to generating a comparator is that it relies on an unadjusted naïve comparison between the weighted comparator and the integrated efficacy analysis with significant scope for confounding bias. The ERG in particular notes that a significant proportion of the patients in the integrated efficacy population (37.0%) received entrectinib as a first-line systemic therapy, while the comparator data set draws predominantly from patients in later lines of therapy. Further, the use of NICE TAs as source of effectiveness evidence means that comparator effectiveness data is being drawn from a population who are primarily *NTRK* fusion negative. This is problematic because there is evidence to suggest that *NTRK* fusions are prognostic, with variable impact upon prognosis depending upon tumour type.

Because of these significant concerns about confounding bias and the challenges of generating a truly comparable comparator data set the ERG considers that the company should have also considered other approaches to generating a comparator data set. The company could for example have considered two approaches discussed in a recent publication by Hatswell *et al.* [1] which suggest using evidence from non-responders and on patients' time to progression on previous lines of therapy to further explore the uncertainties associated generating a comparator data set.

5) Uncertainty surrounding the extrapolation of OS data for entrectinib

The ERG highlights that the observed data for entrectinib was immature, with median OS not yet met. As such, there is significant uncertainty regarding the longer-term survival benefits of entrectinib. The company base-case fits an exponential function to the available KM which was selected from a range of standard parametric functions on the basis that the exponential function has the best statistical fit to the observed data. The ERG considers that the exponential function represents a potentially plausible extrapolation of OS, but is concerned that it implies that post-progression survival is significantly longer than pre-progression survival. The ERG questions the clinical plausibility of this given that entrectinib therapy is discontinued on progression and that only [redacted] of patients received any subsequent therapy. The ERG's preference is therefore for the Weibull function, which produces a more reasonable balance between pre- and post- progression survival, while also having good statistical fit to the observed data.

6) Uncertainty surrounding the appropriate testing strategy and applied testing costs

The ERG also has substantive concerns regarding the companies approach to modelling *NTRK* fusion testing. The ERG in particular is concerned that the company appears to have included extensive

testing costs in the comparator arm of the model. The ERG considers that the focus of modelled testing costs should be on the incremental testing costs associated with identifying *NTRK* fusion positive patients.

Furthermore, it is not clear whether the primary strategy proposed by the company of using IHC followed by NGS will reflect NHS practice should *NTRK* be recommended for use on the NHS. The ERG notes that there are a range of alternative testing strategies that have been discussed in the literature with consequences for the incremental costs of implementing *NTRK* fusion testing as well diagnostic accuracy.

7) Uncertainty surrounding broader infrastructure requirements

The implementation of an appropriate testing regime to identify patients with *NTRK* gene fusions would likely require a significant increase in molecular testing with between 50 and 92 thousand patients potentially eligible for testing. Regardless of the testing strategy adopted for *NTRK* fusions, this is likely to require a significant expansion of current testing service capacity, which would potentially require further investment infrastructure and/or training. These costs are, however, not considered in the company's economic analysis.

6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.1 Overview

This section details the ERG's further exploration of the assumptions and uncertainties raised in the review and critique of the company's cost-effectiveness analysis, presented in Section 5. This section is organised in four parts. Section 6.2 details the impact of correction of errors identified in ERG's validation of the executable model and other amendments to the company base-case analysis. Section 6.3 details a series of scenario analyses exploring the robustness of the cost-effectiveness results under specific assumptions and additional uncertainties identified by the ERG. These analyses were conducted within the company-corrected base-case analysis as presented in Table 44 in Section 5.2.9.1 . The scenario analyses presented in Section 6.3 focus on exploring the following issues:

- An alternative distribution of tumour types;
- Testing costs to identify *NTRK* fusion patients;
- Estimation of treatment-related costs.

In Section 6.4, the ERG alternative base-case is presented, which combines a number of exploratory analyses presented in Section 6.3 and alternative assumptions provided in the company exploratory analyses.

Further exploratory analyses in the context of the ERG alternative base-case analysis are presented in Section 6.5. This section presents the implementation of an alternative model structure for estimating outcomes in the established management arm. In addition, the ERG presents additional statistical analyses of the results of the economic model, including estimating the value of heterogeneity and net population benefit in the ERG's alternative base-case economic model.

Due to time constraints, ICERs based on deterministic analyses are presented throughout this section.

There are a number of treatment options in the established management arm that are associated with a confidential PAS. These include eribulin, everolimus, nintedanib, nab-paclitaxel, trabectedin, and trifluridine and tipiracil. The results of these analysis with the cPAS applied are presented in a confidential appendix to this report.

6.2 ERG corrections and adjustments to the company's base case model

The ERG identified a minor error in the executable model, pertaining to the estimation of post-progression second-line treatment costs in the entrectinib arm. The model effectively applied the discount rate twice to these costs in addition to the omission of the drug administration costs. At the clarification stage, the company provided a corrected model, which also incorporated the results of a

survival analysis for entrectinib based on a more recent data cut. The impact of this correction was minor: the company base-case ICER was reduced from £54,646 to £52,609 (discussed and presented in Section 5.2.9.1, Table 44).

Subsequent analyses in this section are based on this corrected, updated analysis.

6.3 Additional ERG analyses

6.3.1 Alternative distribution of tumour types

The company base-case analysis uses the distribution of tumour types from the integrated efficacy analysis to estimate a weighted set of outcomes for established management. The ERG presents a scenario based on a plausible alternative distribution of tumour types, which was estimated using observed *NTRK* fusion frequencies provided in the FMI data set and published cancer statistics for England. The FMI database is considered by the ERG to be more representative as it is based on a large sample of 166,000 patients. The alternative distribution used by the ERG, along with the original proportion of tumour types provided in the CS, can been seen in Table 29, Section 5.2.3 The method used to estimate this can be found in Appendix A: ERG estimates of eligible population.

The impact of incorporating this alternative distribution of tumour types resulted in a re-estimation of the weighted outcomes for the comparator arm: it was not possible to reweight the outcomes in the entrectinib arm as OS and PFS data were not available by tumour type. This approach therefore implicitly assumes homogeneous PFS and OS across tumour types.

The results of this analysis are presented in Table 48. This scenario was associated with greater incremental costs and lower incremental QALYs than the base-case analysis. The difference in the cost was mostly driven by the large decrease in the proportion of patients with sarcoma: a tumour type associated with higher total costs than the other tumour types in the established management arm.

Table 48 Results of the ERG analyses on the distribution of tumours

	Inc costs	Inc QALYs	ICER
Base case	[redacted]	[redacted]	£52,609
Scenario 1: Alternative distribution of tumour types	[redacted]	[redacted]	£69,747
Inc., incremental; QALYs, quality adjusted life years; ICER, incremental cost effectiveness ratio			

6.3.2 Testing costs to identify *NTRK* fusion+ patients

Marginal costs of testing

The ERG considers that the focus of applied testing costs should be on the incremental testing costs associated with identifying *NTRK* fusion positive patients. The ERG therefore implements a scenario in which testing costs are removed from the established management arm. As WGS is currently funded for sarcoma and paediatric patients, zero incremental costs are assumed for sarcoma and paediatric patients in this scenario. Similarly, the testing costs for MASC patients were removed from the model for this scenario as current testing already identities *NTRK* fusions in these patients. The results of this analysis are presented as Scenario 2 in Table 49.

Removal of testing costs of NGS for lung cancer patients

As discussed in Section 5.2.8, the cost of adding a new *NTRK* panel to an RNA-based NGS test is negligible. Currently, lung cancer is the only tumour type of those included in the efficacy evaluable data set where RNA-based NGS is available for a specific subgroup of patients with NSCLC. A scenario is therefore presented where no additional costs would apply for lung cancer patients. The results of this analysis is presented as Scenario 3 in Table 49. In this scenario, only marginal costs of testing are applied (as per Scenario 2).

Confirmatory NGS following WGS

The ERG received clinical advice that WGS cannot be used to confirm the presence of *NTRK* fusions at present (see Section 2.2.2) and that a confirmatory NGS test would be required for patients receiving WGS. Scenario 5 in Table 49 presents the results of including a confirmatory RNA-based NGS test in patients who already receive WGS. It is assumed WGS will remove 89% of *NTRK* fusion negative samples, reducing the requirement for RNA-based NGS confirmatory testing to 11% of the NNS population. This figure is based on the company's assumptions for IHC as the ERG were unable to identify any statistics on the performance of WGS. In this scenario (Scenario 4 in Table 49), only marginal costs of testing are applied (as per Scenario 2).

Numbers needed to screen

The discovery of the *NTRK* gene fusion is a relatively recent one and the frequency of the fusion in tumour types is still being established. As a result, there remains a degree of uncertainty regarding the exact frequencies used in the model. The number of patients who require screening to identify one individual with an *NTRK* fusion varies depending on the frequency of the gene fusion. The ERG estimated an alternative set of prevalence rates for each tumour type, with details provided in Appendix B. The results of this analysis are presented as Scenario 5 in Table 49.

Cost of testing in whole NTRK population

As outlined in Section 3.1 a number of tumour types are not represented in the model, as such the testing costs represent this population rather the [redacted]. Using data from the FMI database, the ERG implemented a scenario where testing costs are estimated based on all tumour types know to harbour *NTRK* fusions. In this scenario the distribution of tumour types is also assumed to align with ERG estimates of the NNS presented in Appendix B. In unrepresented tumour types, the ERG assumes that patients will receive IHC followed by confirmatory RNA-based NGS, unless WGS is already available on the NHS. The results of this analysis are presented as Scenario 6 in Table 49.

Removal of testing costs

The ERG also included the scenario in which all testing costs were removed (Scenario 8). This represents a future practice scenario where screening is routinely carried out on NHS. Clinical advice to the ERG, however, suggests that this is not likely to happen in the near future. The inclusion of this scenario represents a potential lower bound for the estimate of cost-effectiveness.

Identifying paediatric glioma patients

At the clarification stage, the ERG requested that the company present an analysis that includes primary CNS and paediatric patients. For the purposes of the model, the company grouped the paediatric primary CNS patients with the adult primary CNS patients for the weighted comparator costs and outcomes, since common comparators were assumed for these patients.

Following the factual accuracy check, the company highlighted that screening costs for glioma are overestimated since the costs represent a mixture of adult (five) and paediatric (four) primary brain tumours; screening costs for paediatric gliomas are significantly lower to due to inclusion in the genomic test directory.

The company analysis presented following the clarification stage applied the cost of IHC and confirmatory NGS to identify these patients (Scenario 9 in Table 50). The ERG has implemented a scenario whereby the costs of testing for paediatric glioma patients were based on WGS, which had the impact of removing the cost in these patients. As a result, the ICER was reduced from £49,358 to £48,860 per QALY.

Results

The results of the scenarios described above are presented in Table 49.

Table 49 Results of the ERG analysis on testing costs

Scenario	Inc cost	Inc QALY	ICER
Base case	[redacted]	[redacted]	£52,609
Scenario 2: Remove testing costs in comparator arm	[redacted]	[redacted]	£63,329

Scenario 3: Remove lung cancer cost of testing	[redacted]	[redacted]	£59,465
Scenario 4: Confirmatory RNA-based NGS in WGS patients	[redacted]	[redacted]	£64,608
Scenario 5: Prevalence of *NTRK* fusions (tumour types represented in the trial)	[redacted]	[redacted]	£56,914
Scenario 6: Prevalence of *NTRK* fusions (based on the whole *NTRK* population)	[redacted]	[redacted]	£65,981
Scenario 7: Cumulative impact of 2, 3, 4, 6	[redacted]	[redacted]	£64,115
Scenario 8: No testing costs	[redacted]	[redacted]	£36,914

Table 50 Results of the ERG analysis on testing costs of paediatric patients

	Inc costs	Inc QALYs	ICER (£/QALY)
Company base-case scenario including CNS patients and children (Table 47 in Section 5.2.9.4)	[redacted]	[redacted]	£49,358
Scenario 9: WGS for identifying *NTRK* tumours in paediatric patients	[redacted]	[redacted]	£48,860

6.3.3 Treatment costs

eMIT costs for therapies in established management arm

Drug acquisition costs for the comparator therapies were obtained from the BNF. However, many of these therapies are generic products that are widely available to the NHS at discounted prices. The Department of Health's eMIT database provides information on the average price paid by the NHS for pharmaceuticals, which can differ from the list prices listed in the BNF and are a more representative estimate of drug expenditure. Unit costs from eMIT (presented in Table 36 in Section 5.2.8) are generally considerably lower than those in the BNF, and the use of the BNF costs will overestimate drug expenditure, biasing the analysis in favour of entrectinib. The results of this analysis are presented as Scenario 10 in Table 51.

Drug wastage

Wastage has the potential to significantly impact upon drug expenditure, and the ERG is concerned that the company's model, which excludes drug wastage in the base-case analysis, underestimates the drug costs that would be incurred by the NHS. The ERG explored a scenario that allowed for drug wastage. The results of this analysis are presented as Scenario 11 in Table 51.

Results

The results of the scenarios described above are presented in Table 51. The analysis was not sensitive to the inclusion of eMIT unit costs for the therapies in the established management arm; however, the

inclusion of drug wastage for entrectinib resulted in an increase to the ICER of approximately £2,750, due to the additional drug costs in this arm.

Table 51 Results of the ERG analysis on treatment costs

Scenario	Inc costs	Inc QALYs	ICER
Base case	[redacted]	[redacted]	£52,609
Scenario 10: eMIT costs for comparator therapies	[redacted]	[redacted]	£52,081
Scenario 11: With drug wastage	[redacted]	[redacted]	£55,357

6.4 ERG alternative base-case

Table 52 presents the results of the ERG alternative base-case analysis. These incorporate a number of changes to key model parameters and assumptions, which were previously explored individually in Section 6.3, as well as a number of additional assumptions which were previously explored by the company in scenario analyses (Section 5.2.9.3). Most notably, this includes an alternative extrapolation of available PFS and OS data for entrectinib. As discussed in Section 5.2.6, the ERG considered the Weibull function to be a more preferable model for OS and PFS, as it uses the more reasonable assumption of increasing hazards over time, and resulted in a more plausible estimate of time spent in the post-progression health state.

The ERG alternative base-case analysis includes the following changes to the company base-case analysis:

- Inclusion of children and primary CNS tumours in the population (see Section 5.2.3),
- Weibull distribution for entrectinib OS and PFS (Section 5.2.6),
- Inclusion of marginal testing costs only,
- Confirmatory RNA-based NGS test after WGS test, and removal of NGS testing costs for lung cancer patients,
- WGS test to identify *NTRK* tumours in paediatric patients,
- Testing costs estimated using the number needed to screen based on the whole *NTRK* population,
- Second-line therapy following discontinuation of entrectinib, limited to 6 month duration,
- eMIT costs for therapies in the established management arm,
- Inclusion of drug wastage for entrectinib.

Under the ERG's alternative set of assumptions, the ICER for entrectinib versus established care is £77,120 per QALY.

Table 52 ERG alternative base-case analysis

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Entrectinib *	[redacted]	[redacted]	[redacted]	[redacted]	£77,109
Established management	£19,853	1.03	-	-	-
* Note that these results have changed following the factual accuracy check to include the change made in Scenario 9					

6.5 Exploratory analysis on ERG base-case

6.5.1 Estimation of comparator outcomes based on a response model

Motivation

As discussed in Section 5.2.6.1, the ERG considers that the company should have also considered alternative methods of generating comparator effectiveness estimates. These could include a dual-partitioned response-based model, which distinguishes between responders to treatment and non-responders. This additional model complexity allows for a distinction to be made in the health-related quality of life (HRQoL) of responders and non-responders, as well as allowing for potential differences in the costs of care.[1] An alternative would be based around a surrogate relationship between response and PFS and OS. In the FDA evaluation of larotrectinib, it was considered that these surrogate relationships were reasonably likely to predict meaningful benefit. However, this approach has its own drawbacks: a review of the relationship between the more long-term outcomes of PFS and OS suggested that it varies considerably by cancer type and is not always consistent even within one specific cancer type.[53] In the absence of specific guidance on the surrogate relationship between response and survival, the use of this type of model structure would need to be accompanied by a review of studies in *NTRK* fusion patients to consider the extent to which response-based outcomes can be considered a robust surrogate endpoints for PFS and OS, and to establish how these relationships might be quantified in a modelling approach.

The ERG implemented an exploratory responder-based approach, which uses effectiveness data on non-responder patients as a proxy for patients not receiving an active treatment. The ERG recognises that such an approach is subject to limitations particularly regarding the maturity of the data and the number of patients included in the analysis, but believes that presenting the results of this analysis can provide some degree of reassurance regarding the predicted comparator effect estimates, given the large amount of uncertainty in the approach taken in the company base case. This approach ensures that the population used to model the comparator and the intervention arms are consistent with each other, which is of particular importance given that the prognostic status of *NTRK* fusion tumours is unknown and likely to differ based on each tumour type. However, by ensuring that the efficacy in

both arms is reflective of the trial, it does limit the applicability of the findings to the general eligible population.

As described in Section 5.2.1, a further advantage of the response-based model is that it is easier to generate ICERs specific to each tumour type. This was implemented in the response-based model by altering the rate of response in the model and subsequently changing the survival predictions for the entrectinib arm. As in the model above, survival in the established management arm was modelled assuming a 0% response in the comparator arm. This analysis therefore makes the strong assumption that effectiveness for established management is the same across all tumour types. The assumed response for each tumour type was based on the Bayesian hierarchical analysis presented in Section 4.3.1.

Methods

The ERG constructed a response-based model using the heterogeneous response rates across tumour types, estimated by the BHM in Section 4.3.1, and linked these to OS and PFS. This method facilitated linking response to costs and QALYs and so create histology specific estimates of cost effectiveness. It should be emphasised that this model is for illustrative purposes and given the data made available it has been necessary to make strong assumptions to explore heterogeneity.

At the clarification stage, the ERG requested that the company provide KM plots for PFS and OS for non-responders and responders to entrectinib. The company provided this information for the population in the original analysis, and the population that also includes paediatric patients and patients with primary CNS tumours.

The ERG reconstructed the individual participant data (IPD) for the population including paediatric patients and patients with primary CNS tumours, and fit standard parametric models (exponential, Weibull, log-logistic, lognormal, Gompertz, and generalised gamma) to the dataset for each outcome, using response status as a covariate. Survival in the established management arm was assumed to be equivalent to that of the non-responder patients, and survival in the entrectinib was estimated as a weighted average of survival in the responder and non-responder patients, weighted by the estimated response rate of [redacted] from the BHM described in Section 4.3.1.

To determine the most appropriate model, the ERG referred to fit statistics (AIC and BIC, Table 53 below), visual fit to the observed KM curves, and clinical plausibility of survival estimates. Figure 20 a graphical summary of each curve and their fit to the observed KM data. For PFS, generalised gamma had the best statistical fit; however, it appears to produce long-term projections that were considered overly optimistic and, therefore, implausible. The fit statistics for OS favoured the

lognormal and the exponential. However, the lognormal did not appear to fit well when compared to the responder population.

Figure 20 Survival extrapolations

Figure redacted

Table 53 Fit statistics for survival models fit to the whole population in the integrated analysis

	Overall survival		Progression-free survival	
	AIC	**BIC**	**AIC**	**BIC**
Generalised gamma	209.5179	218.2766	275.3453	284.1039
Weibull	210.2889	216.8579	284.6180	291.1870
Exponential	209.3053	213.6846	283.5421	287.9214
Loglogistic	208.5645	215.1334	281.5299	288.0988
Lognormal	208.2129	214.7819	279.6231	286.1920
Gompertz	210.9277	217.4967	285.5215	292.0904

On the basis of the plausibility of the long-term predictions, comparisons with the KM plots for the responder and non-responder population, and for consistency with the assumptions made in the ERG alternative base-case analysis, the Weibull survival function was selected to model both PFS and OS in this exploratory cost-effectiveness analysis. Figure 21 compares the predicted survival for entrectinib and established management (using the Weibull distribution) compared with the predicted survival for each arm as used in the ERG alternative base-case analysis in Section 6.4. The two methods result in similar survival extrapolations for entrectinib; however the survival curve for established management estimates higher survival in the response-based model.

Figure 21 Comparison of survival functions used in the ERG base case and the responder-based cost-effectiveness analysis

Figure redacted

Results

The analysis was based on the assumptions made under the ERG alternative base-case analysis set out in Section 6.4.

The ICER in this analysis was £95,723 per QALY (Table 54). The costs and QALYs generated for the entrectinib arm were similar to that of the ERG base-case analysis in Table 44. However, this method produces higher QALYs and costs in the established management arm as a result of the higher rate of survival for these patients. Consequently, the ICER that was estimated using this method is higher than that using the same assumptions under the model structure presented by the company.

Table 54 Results of responder-based cost-effectiveness analysis

	Total costs	**Total QALYs**	**Incremental costs**	**Incremental QALYs**	**ICER (£/QALY)**
Entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	£95,709
Established management	£28,507	1.32	-	-	-
* Note that these results have changed following the factual accuracy check to include the change made in Scenario 9					

The results of the analyses by tumour type are presented in Table 55. The ICERs ranged from £57,451 per QALY for sarcoma patients, to £128,663 for thyroid tumours.

Table 55 Results of the responder model by tumour type

Tumour type	**ICER ***
CRC	£98,493
MASC	£111,464
Thyroid	£128,663
NSCLC	£89,668
Pancreatic	£89,770
Sarcoma	£57,451
Neuroendocrine	£108,634
Breast	£86,697
Glioma	£117,456
IFS	£119,787
Melanoma	£114,868
Other	£98,164
All tumours	*£95,709*
* Note that these results have changed following the factual accuracy check to include the change made in Scenario 9	

6.5.2 Value of heterogeneity and net population benefit

The costs and health consequences associated with using entrectinib in colorectal cancer (CRC) with *NTRK* fusions are used to illustrate the importance of taking account of heterogeneity in histology independent assessments. CRC was chosen as it had a low predicted response rate compared to other tumour types, and accounted for a large proportion of patients eligible for treatment with entrectinib. From the responder-based model, entrectinib was associated with [redacted] in extra costs and [redacted] additional QALYs in CRC. The additional QALYs from entrectinib are lower in CRC as the estimated response rate for these patients, as estimated from the hierarchical model, is [redacted], which is below the pooled response rate across all tumour types, which is [redacted]%. This results in an ICER of approximately £95,451 per QALY for CRC. This can be compared to the additional costs and QALYs associated with using entrectinib across all tumour types. For this pooled *NTRK* population (which includes CRC) entrectinib is associated with [redacted] in extra costs and [redacted] additional QALYs. This results in an ICER of approximately £93,532 per QALY.

This simple comparison of a subgroup specific ICER to a pooled population ICER illustrates that the cost-effectiveness of entrectinib could vary significantly between individual tumour types. This also means that the 'average' ICER could be more favourable if the subgroup with a CRC histology were excluded. To understand the implications of this for population health requires that benefits and costs are expressed as net health effects (NHE). The NHE is the difference between any health gained with the intervention and the health forgone elsewhere in the health-care system, all expressed in QALY terms. With an ICER in CRC of approximately £96,451 per QALY, the incremental NHE at a threshold of £50,000 is -0.45 QALYs per patient, that is, the additional health gained with the intervention is more than offset by health forgone elsewhere. This means that for every CRC patients who receives entrectinib, 0.45 QALYs could potentially be lost elsewhere in the health system.

Table 56 Value of heterogeneity – an illustrative example using the CRC population

	Total cost	Total QALYs	ICER	NHE (QALYs)	Incremental NHE (QALYs)
Entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	-0.45
Established management	£32,460	1.43	-	0.785	-
Note, a threshold of £50,000 was used to estimate NHE					

The advantage of NHE is that they can be used to help understand the population level consequences of decisions. The number of CRC patients with *NTRK* fusions in the UK was estimated by the ERG to be approximately 29 per year (Appendix A). This means that an 'non-optimised' recommendation which includes CRC might result in an additional 12.99 QALYs per year to the health system

compared to established management. In other tumour types, entrectinib may provide positive QALYs to the health system but further analysis would be required to identify these tumour types, if they exist. As the cost effectiveness of entrectinib could depend on the tumour type treated, this analysis also illustrates the importance of understanding the distribution of tumour types expected to receive the treatment in practice.

6.6 Conclusions from ERG analyses

The ERG has presented a number of additional analyses carried out in stages. These exploratory analyses were undertaken on a model provided by the company at the clarification stage, which addressed an error identified by the ERG, and included a more recent data cut of the survival data from the integrated efficacy analysis. The impact of these changes was to decrease the ICER from £54,646 to £52,609 per QALY.

Using the corrected and updated model, the ERG then presented a number of analyses considering a range of issues raised in Section 5.2. These scenario analyses addressed the following issues:

- An alternative distribution of tumour types;
- Testing costs to identify *NTRK* fusion+ patients;
- Estimation of treatment-related costs.

The scenarios associated with the greatest impact on cost-effectiveness outcomes related to changes made by the ERG, which involved the removal of testing costs in the comparator arm to more accurately reflect the incremental cost of testing to identify *NTRK* fusions. Testing costs comprise a significant proportion of the total costs, and removing the testing costs for the comparator resulted in the ICER increasing from £52,609 to £63,329. A scenario analysis that explored the impact of an alternative distribution of tumour types demonstrated that the results of the model are sensitive to this assumption. This sensitivity was a consequence of tumour types being associated with different QALYs or costs, highlighting the heterogeneity in this patient population.

The ERG alternative base-case implemented a number of alternative assumptions that were included in the company exploratory analyses. The assumptions that had the largest impact on the ICER was the restriction of duration of second-line therapy to 6 months following discontinuation of entrectinib, and the implementation of a Weibull survival model to estimate overall and progression-free survival of entrectinib. This analysis estimated entrectinib to be more costly (cost difference [redacted]) and more effective ([redacted] QALY gain) compared with established management, and suggests that the ICER for entrectinib compared with established management is £77,109 per QALY.

The final part of this section carried a further series of exploratory analyses that explored the impact of an alternative method to estimate survival. This method used the survival of non-responder patients

to estimate survival predictions in the established management arm. The entrectinib arm was based on a weighted average of responder and non-responder survival predictions, which allowed for the exploration of cost-effectiveness in different tumour types by varying the response rate used to estimate the weighted average. The ICER for the pooled group was £95,705 per QALY. This was higher than the ICER estimated in the ERG analysis, as a result of the higher survival rates predicted by the response-based model for the established management arm. When varied by tumour type, the ICERs ranged from £57,451 per QALY for sarcoma patients, to £128,663 for thyroid tumours.

7 End of life

In the CS and clarification response, the company state that entrectinib meets the end-of-life criteria compared to the current established management across all patients potentially eligible for entrectinib, on the basis of the results of the integrated efficacy analysis.

The conventional application of end-of-life (EoL) criteria to a highly heterogeneous population with no established comparators is challenging in two respects. Firstly, the EoL criteria may apply across some tumour types and not others. Secondly, there is a great deal of uncertainty around estimates of both life expectancy and extension to life which may vary widely by tumour type, this is further exacerbated by the uncertainty around the positioning of entrectinib in the treatment pathway for each tumour type. While there is little precedent for decision optimisation on the basis of the eligibility of sub-populations for EoL, it does not appear appropriate to apply a higher willingness-to-pay threshold to sub-populations that do not otherwise meet the necessary criteria based on the 'unmet need' of other included cancer types. Application of the higher willingness-to-pay threshold in such cases necessarily implies that patients are able to access therapy that otherwise would be considered cost-ineffective based on conventional thresholds and potentially raise issues about equity of access treatment, as QALY generated in *NTRK* positive and *NTRK* negative patients are being valued differently.

Application of the higher EoL threshold across all tumour types regardless of whether they all meet EoL also potentially offers as a commercial advantage to histology independent products as competitor products for tumour types not meeting EOL, which are required to be priced in accordance with conventional £20,000 to £30,000 thresholds, and as a consequence potentially distorts investment incentives towards histology independent therapies.

Criterion 1: The treatment is indicated for patients with a short life expectancy, normally less than 24 months.

In the ERG's base-case analysis, the population who are anticipated to meet the eligibility criteria in the product license have an average mean OS of 20.89 months (median 15.7 months). In the ERG's

response-based model, the mean OS of those patients who did not respond to entrectinib was 24.72 months (median 19.9 months) using a Weibull function. The ERG favours the use of the mean to represent life expectancy, as it better represents the distribution of OS, and measures of health benefit upon which decision about cost-effectiveness are made on the basis of mean values (mean QALYs). While the base-case mean survival falls under the two years stipulated in Criterion 1, the ERG does not consider this figure appropriate for decision-making for a number of reasons.

Firstly, as discussed in Section 3.3, the comparability of the comparator population with patients eligible for entrectinib in clinical practice is highly uncertain, particularly given the company's pooling of often very different life expectancy data from TAs covering multiple lines of therapy within the same indication. As a testing strategy will dictate when entrectinib is made available, it will always be at the same point in the pathway, making one comparator life expectancy estimate more appropriate than the other.

Secondly, the company anticipate NICE's recommendation to cover all tumour types affected by *NTRK* gene fusions. However, the majority of tumour types are not represented in the company's trial or comparator searches, and therefore the life expectancy of these populations is unknown, and may be significantly different to those included in the CS.

The company also stated in their clarification response that the prognostic implications of *NTRK* gene fusions mean these patients are likely to have a lower OS than the population considered in the NICE appraisals. As previously discussed, the ERG does not consider existing literature to support the concept of *NTRK* as being consistently prognostic of a shorter life expectancy.

It is highly uncertain whether the presented average OS estimate represents life expectancy at the time patients would become eligible for entrectinib. Furthermore, it does not reflect the heterogeneity of life expectancy across tumour types when they reach eligibility for treatment. A summary of mean and median OS estimates by tumour type used in the ERG base-case analysis are presented in Table 57.

Table 57 Average SoC OS by tumour type

Tumour type	Median OS (months)	Exp. mean OS (months)
Breast	12.18	17.56
Colorectal	9.07	13.08
MASC	13.80	19.91
Neuroendocrine	39.61	57.14
NSCLC	10.65	15.36
Cholangiocarcinoma	17.23	24.86
Pancreatic	8.80	12.70
Sarcoma	14.30	20.63
Thyroid	30.95	44.65
CNS	7.95	11.46
Infantile fibrosarcoma	17.23	24.86
Melanoma	6.40	9.23
Total (EEA weights)	**15.74**	**22.71**
Total (ERG weights)	**16.39**	**23.64**

These OS estimates suggest that patients with thyroid and neuroendocrine tumours would not meet the first of the EoL criteria, and represent approximately 31% of the incident *NTRK* fusion population (see Section 3.1).

Criterion 2: There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment.

The CS states that the median OS has not yet been reached in the EEA dataset, however, based on the company's extrapolation of the latest data cut, the estimated mean OS is predicted to be [redacted] This suggests a mean OS benefit of [redacted]; therefore, the company conclude that entrectinib offers at least a three-month extension to life. The ERG base-case predicted a mean OS on treatment with entrectinib of 31.1 months, suggesting a mean OS benefit of 10.2 months.

The ERG's response-based model presented in scenario analysis demonstrates that this extension to life may not be consistent between tumour types, which was also the case in company scenario analyses in their clarification response (Appendix E Table 76). Patients who responded to entrectinib had a median OS (Weibull) of 25.2 months, suggesting an OS benefit of 5.3 months (non-responder median OS = 19.9 months). Mean OS benefit by tumour type in the ERG's base-case ranged between 6.40 months (CRC) and 8.80 months (MASC).

While the ERG notes the significant challenges in obtaining a robust estimate of the extension to life generated by entrectinib, these ranges of values suggest that extension to life across the majority of tumour sites is likely to be greater than 3 months. The benefit of entrectinib in unrepresented tumour types is unknown and cannot be assumed to be equal to that seen in the trials. Utilising the predictive distribution estimated in section 4.3.1 in the response based model, extension to life in unobserved tumour sites could potentially range from [redacted] months to [redacted] months. Some tumour sites may therefore not meet the 3-month criteria.

8 Overall conclusions

The clinical evidence for entrectinib is very limited. Most of the efficacy evidence comes from an *NTRK* positive subgroup of patients of a phase 2, uncontrolled basket trial. A total of only 66 *NTRK*-fusion positive patients across 13 tumour types were included in the efficacy evidence, and each of the tumour types was represented by between one and 13 patients. Overall, the trial evidence showed a clinically meaningful overall response rate across tumour types. However, there is considerable uncertainty regarding the extent to which the response observed translates into clinically meaningful survival benefits. OS, PFS and DOR data presented were immature. Despite substantial censoring and the small number of patients at risk in the tails of the Kaplan-Meier curve, the crossing of OS curves between entrectinib responders and non-responders is of some concern.

Due to limited evidence, there is considerable uncertainty about the precision of response and survival benefit estimates and heterogeneity by tumour type and line of therapy. The ERG explored heterogeneity in response rates between 13 tumour types using a Bayesian hierarchical model, which assumes the response probabilities are similar (i.e. exchangeable) across tumour types, rather than identical (the company's preferred assumption). Although the ERG's analyses found that response rates obtained were similar to those presented in the company submission, there was considerable uncertainty in the level of heterogeneity of response rates across tumour types. Therefore, the possibility that some tumour types could have response rates that differ significantly from the pooled [redacted] response rate cannot be excluded. Due to small numbers of patients and subgroups there was insufficient evidence to explore formally whether response and survival may differ by *NTRK* fusion subtype or line of therapy.

The company's updated base-case ICERs for entrectinib compared with established management and presented single ICER of £52,609 per QALY (inclusive of the confidential PAS) to cover all the [redacted] anticipated marketing authorisation.

The ERG's review of the company's presented analysis centred round the challenges associated with assessing cost-effectiveness in a histology independent indication and the uncertainties associated with the limit evidence effectiveness available. The ERG proposed an alternative base-case to address several of the key uncertainties identified. These included uncertainties associated with testing and identify patients with NTRK fusions. This analysis explored alternative estimates of the NNS, as well as testing costs in tumour types not represented in the trial. The ERG also considered a number of plausible modifications to the testing strategy proposed by the company. Uncertainty surrounding the extrapolation of survival data for entrectinib was explored, with the ERG preferring to model PFS and OS using a Weibull function instead of an exponential function proposed by the company. The ERG base-case also included a number of minor alterations to costs as a scenario analysis presented by the

company as part their clarification response which explored alternative assumptions regarding the duration of post progression therapy.

Despite the ERG's attempt to address all the relevant uncertainties, data limitations imply that some key uncertainties could not be fully explored. These unresolved uncertainties could potentially have a profound impact on the cost-effectiveness of entrectinib and would require further data to fully address.

First, the cost-effectiveness estimates are based on an uncontrolled comparison used, which used data from previous NICE TA's as a source of effectiveness data. The ERG, however, found that the population included in the comparator trials is unlikely to match the entrectinib efficacy population, notably due to the unknown prevalence of NTRK fusions in most of the comparator evidence, and the mismatch of previous lines of therapy with the treatment pathway. The ERG therefore has substantive concerns about the validity of the comparator effectiveness data and explored alternative methods of generating a comparator data set by modifying the company model structure so that PFS and OS outcomes were determined based on response to treatment. In this scenario analysis comparator outcomes were generated assuming that all patients were non-responders.

Second the ERG has concerns about the implicit assumption of a homogenous treatment effect across all tumour types and that the presentation of a single ICER conceals the potential for significant variation in tumour specific ICERs. To explore this uncertainty, the ERG, utilised the response base model to integrate the results of the Bayesian hierarchal analysis discussed above, and generate tumour type specific ICERs. This exploratory analysis showed that the tumour type specific ICER's varied significantly from £57, 451 per QALY in sarcoma to £128,663 per QALY in Thyroid cancer (ICER for all tumour types £95,723 per QALY). Methods for further exploring the heterogeneity in the ICER in using population NHE were also present in brief with an illustrative example presented in CRC. This considered the implications of an optimised decision in which CRC was excluded from any NICE recommendation.

Third there are a number of uncertainties relating to the population treated and the positioning of entrectinib in the treatment pathway. This has implications for the modelled comparators that are assumed to represented established management. The company stated that they expect entrectinib to be positioned towards the end of a patient's treatment pathway and this was reflected in the comparator data selected to represent established management. However, the anticipated marketing authorisation for entrectinib is ambiguous in this regard and potentially permits entrectinib to be used as a first-line therapy in several tumour types. Alterative assumptions about the position of entrectinib will necessarily have implications on the model including on comparator costs, effectiveness and HRQoL.

8.1 Implications for research

Exploring outstanding uncertainties differential response rates and survival benefits across tumour sites represents potentially valuable aim of any further research. While evidence from a large RCT would be preferred, it is acknowledged that it is unlikely to be feasible to conduct one in this population. However, a mature and appropriately powered basket trial recruiting patients with a wide range of tumour types in statistically sufficient numbers, and at clinically appropriate and consistent positions in treatment pathways will be necessary to assess heterogeneity of response to entrectinib to inform optimised decision making in future.

9 References

1. Hatswell AJ, Sullivan WG. Creating historical controls using data from a previous line of treatment - Two non-standard approaches. *Stat Methods Med Res* 2019:962280219826609. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30698076
2. Amatu A, Sartore-Bianchi A, Siena S. NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. *ESMO Open* 2016;**1**:e000023. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27843590
3. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A, et al. ALK, ROS1, and NTRK Rearrangements in Metastatic Colorectal Cancer. *J Natl Cancer Inst* 2017;**109**. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29370427
4. Musholt TJ, Musholt PB, Khaladj N, Schulz D, Scheumann GF, Klempnauer J. Prognostic significance of RET and NTRK1 rearrangements in sporadic papillary thyroid carcinoma. *Surgery* 2000;**128**:984-93. Available from: https://www.ncbi.nlm.nih.gov/pubmed/11114633
5. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. *Cancer* 2016;**122**:1097-107. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26784937
6. Vokuhl C, Nourkami-Tutdibi N, Furtwängler R, Gessler M, Graf N, Leuschner I. ETV6–NTRK3 in congenital mesoblastic nephroma: A report of the SIOP/GPOH nephroblastoma study. *Pediatr Blood Cancer* 2018;**65**:e26925.
7. Gatalica Z, Xiu J, Swensen J, Vranic S. Molecular characterization of cancers with NTRK gene fusions. *Mod Pathol* 2019;**32**:147-53. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30171197
8. Okamura R, Boichard A, Kato S, Sicklick JK, Bazhenova L, Kurzrock R. Analysis of NTRK Alterations in Pan-Cancer Adult and Pediatric Malignancies: Implications for NTRK-Targeted Therapeutics. *JCO Precis Oncol* 2018;**2018**. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30637364
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6329466/
9. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* 2017;**23**:703-13. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28481359
10. Chmielecki J, Bailey M, He J, Elvin J, Vergilio JA, Ramkissoon S, et al. Genomic Profiling of a Large Set of Diverse Pediatric Cancers Identifies Known and Novel Mutations across Tumor Spectra. *Cancer Res* 2017;**77**:509-19. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28069802
11. Pavlick D, Schrock AB, Malicki D, Stephens PJ, Kuo DJ, Ahn H, et al. Identification of NTRK fusions in pediatric mesenchymal tumors. *Pediatr Blood Cancer* 2017;**64**. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28097808
12. Solomon JP, Hechtman JF. Detection of NTRK Fusions: Merits and Limitations of Current Diagnostic Platforms. *Cancer Res* 2019.
13. Bozzetti C, Nizzoli R, Tiseo M, Squadrilli A, Lagrasta C, Buti S, et al. ALK and ROS1 rearrangements tested by fluorescence in situ hybridization in cytological smears from advanced non-small cell lung cancer patients. *Diagn Cytopathol* 2015;**43**:941-6.
14. Selinger CI, Rogers T-M, Russell PA, O'toole S, Yip P, Wright GM, et al. Testing for ALK rearrangement in lung adenocarcinoma: a multicenter comparison of immunohistochemistry and fluorescent in situ hybridization. *Mod Pathol* 2013;**26**:1545.
15. England N. *National Genomic Test Directory*. 2019. Available from: https://www.england.nhs.uk/publication/national-genomic-test-directories/ [accessed 15th May 2019].
16. Penault-Llorca F, Rudzinski ER, Sepulveda AR. Testing algorithm for identification of patients with TRK fusion cancer. *J Clin Pathol* 2019. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31072837
17. Marchiò C, Scaltriti M, Ladanyi M, Iafrate AJ, Bibeau F, Dietel M, et al. ESMO recommendations on the standard methods to detect NTRK fusions in daily practice and clinical research. *Ann Oncol* 2019.

18. Liu J, Weng L, Ming Y, Yin B, Liu S. Fusion Genes and Their Detection through Next Generation Sequencing in Malignant Hematological Diseases and Solid Tumors. *Diagnostic Pathology: Open Access* 2015;**1**:108.
19. Williams HL, Walsh K, Diamond A, Oniscu A, Deans ZC. Validation of the Oncomine™ focus panel for next-generation sequencing of clinical tumour samples. *Virchows Archiv: European Journal of Pathology*;**473**:489-503.
20. Bishop JA, Yonescu R, Batista DA, Westra WH, Ali SZ. Cytopathologic features of mammary analogue secretory carcinoma. *Cancer Cytopathol* 2013;**121**:228-33. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23042752
21. Excellence NIfHaC. *Gene expression profiling and expanded immunohistochemistry tests to guide selection of chemotherapy regimes in breast cancer management*: National Institute for Health and Clinical Excellence; 2011.
22. England N. *National Genomic Test Directory for Rare and Inherited Disease*. In: England Ni, editor.; 2019. Available from: https://www.england.nhs.uk/publication/national-genomic-test-directories/
23. Cross P, Chandra A, Giles T, Johnson S, Kocjan G, Poller D, et al. *Guidance on the reporting of thyroid cytology specimens*. London: Royal College of Pathologists; 2016.
24. Abulafi M, Ross P, Gaya A, Riddle P, Monahan K, Burling D, et al. *London Cancer Alliance Colorectal Cancer Clinical Guidlines* London; 2014.
25. Reid R, Bulusu R, Carroll N, Eatock M, Geh I, Judson I, et al. *Guidelines for the management of gastrointestinal stromal tumours (GIST)*. In; 2009. Available from: http://www.augis.org/wp-content/uploads/2014/05/GIST_Management_Guidelines_180809.pdf
26. Ganesan R, Singh N, McCluggage G. *Dataset for histological reporting of endometrial cancer* London: Royal College of Pathologists; 2017.
27. Fotopoulou C, Hall M, Cruickshank D, Gabra H, Ganesan R, Hughes C, et al. British Gynaecological Cancer Society (BGCS) epithelial ovarian/fallopian tube/primary peritoneal cancer guidelines: recommendations for practice. *Eur Addict Res* 2017;**213**:123-39.
28. Hirschowitz LA, Alex; Brown, Laurence; Das, Nagindra; Ganesan Raji; Herrington, Simon; Lopes, Tito; McCluggage, Glen; Meechan, David; Singh, Naveena; Smith, John; Thackeray, Alexandra; Wild, Sue. *Histopathology Reporting in Cervical Screening – an Integrated Approach*. Sheffield; 2012.
29. Oncology DfHaN. *National Head and Neck Cancer Audit 2014 Report 7*. Leeds: NHS Digital; 2014. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/national-head-and-neck-cancer-audit/national-head-and-neck-cancer-audit-2014-dahno-tenth-annual-report#related-links
30. Oxley JV, Murali; Berney, Dan. *Dataset for histopathology reports for prostatic carcinoma*. London; 2016.
31. Alliance LC. *LCA Brain/CNS Cancer Clincial Guidelines*. London; 2014. Available from: http://www.londoncanceralliance.nhs.uk/media/74514/LCA_BCNS_Guidelines_FINAL.pdf
32. UK CR. *Cancer Research UK Policy Statement: Patient Access to Molecular Diagnostics and Targeted Medicines in England*. London; 2018. Available from: https://www.cancerresearchuk.org/sites/default/files/access_to_molecular_diagnostic_tests_and_targeted_medicines_in_england_0.pdf
33. Wright SJ, Newman WG, Payne K. Accounting for Capacity Constraints in Economic Evaluations of Precision Medicine: A Systematic Review. *Pharmacoeconomics* 2019;**37**:1011-27.
34. Bainbridge S CR, Meredith M, Furness P, Gordon B. *Testing Times to Come? An Evaluation of Pathology Capacity Across the UK*: Cancer Research UK; 2016. Available from: https://www.cancerresearchuk.org/sites/default/files/testing_times_to_come_nov_16_cruk.pdf
35. Brada M, Stenning S, Gabe R, Thompson LC, Levy D, Rampling R, et al. Temozolomide versus procarbazine, lomustine, and vincristine in recurrent high-grade glioma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010;**28**:4601-8.
36. Batchelor TT, Mulholland P, Neyns B, Nabors LB, Campone M, Wick A, et al. Phase III randomized trial comparing the efficacy of cediranib as monotherapy, and in combination with lomustine, versus lomustine alone in patients with recurrent glioblastoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013;**31**:3212-8.

37. Middleton MR, Grob JJ, Aaronson N, Fierlbeck G, Tilgen W, Seiter S, et al. Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2000;**18**:158-66.
38. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;**52**:377-84.
39. Thall PF, Wathen JK, Bekele BN, Champlin RE, Baker LH, Benjamin RS. Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. *Stat Med* 2003;**22**:763-80.
40. Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. Robust exchangeability designs for early phase clinical trials with multiple strata. *Pharmaceutical Statistics* 2016;**15**:123-34.
41. Freidlin B, Korn EL. Borrowing information across subgroups in phase II trials: is it useful? *Journal of Clinical Cancer Research* 2013;**19**:1326-34.
42. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay J-Y, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. 2015;**373**:726-36.
43. Heinrich MC, Joensuu H, Demetri GD, Corless CL, Apperley J, Fletcher JA, et al. Phase II, open-label study evaluating the activity of imatinib in treating life-threatening malignancies known to be associated with imatinib-sensitive tyrosine kinases. 2008;**14**:2717-25.
44. Chugh R, Wathen JK, Maki RG, Benjamin RS, Patel SR, Myers PA, et al. Phase II Multicenter Trial of Imatinib in 10 Histologic Subtypes of Sarcoma Using a Bayesian Hierarchical Statistical Model. *J Clin Oncol* 2009;**27**:3148-53. Available from: https://ascopubs.org/doi/abs/10.1200/JCO.2008.20.5054
45. Cunanan KM, Iasonos A, Shen R, Gonen M. Variance prior specification for a basket trial design using Bayesian hierarchical modeling. *Clin Trials* 2019;**16**:142-53.
46. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. *The BUGS book*. Boca Raton, FL: CRC Press; 2013.
47. R Core Team. *R: A language and environment for statistical computing (https://www.R-project.org/)*. [program] Vienna, Austria: R Foundation for Statistical Computing; 2019.
48. Sturtz S, Ligges U, Gelman A. R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software* 2005;**12**:1-16 (http://www.jstatsoft.org/v2/i03). Available from: http://www.jstatsoft.org/v12/i03
49. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013*. London: NICE; 2013. Available from: https://www.nice.org.uk/process/pmg9/
50. National Instiute for Health and Care Excellence. *Eribulin for treating locally advanced or metastatic breast cancer after 1 chemotherapy regimen*. London: NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta515
51. National Instiute for Health and Care Excellence. *Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy*. London: NICE; 2018.
52. National Instiute for Health and Care Excellence. *Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer*. London: NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta476
53. Latimer NR. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making* 2013;**33**:743-54. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23341049
54. Center for Drug Evaluation and Research. *NDA Multidisciplinary Review and Evaluation NDA 210861 and NDA 211710 VITRAKVI (larotrectinib)*: U.S. Food and Drug Administration, ; 2016. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/210861Orig1s000_211710Orig1s000MultidisciplineR.pdf
55. National Institute for Health and Care Excellence. *Lutetium (177Lu) oxodotreotide for treating unresectable or metastatic neuroendocrine tumours*. London: NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta539
56. Excellence NIfHaC. *Eribulin for treating locally advanced or metastatic breast cancer after 1 chemotherapy regimen*; 2018.

57. National Institute for Health and Care Excellence. *Denosumab for the prevention of skeletal-related events in adults with bone metastases from solid tumours. Technology appraisal guidance.* London: NICE; 2012.
58. Phillippo DM, A. E. Ades, S. Dias, S. Palmer, K. R. Abrams and N. J. Welton. . *NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE*. Sheffield, UK, ScHARR, University of Sheffield.; 2016. Available from: https://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.pdf
59. National Institute for Health and Care Excellence. *Trifluridine–tipiracil for previously treated metastatic colorectal cancer*. London: NICE; 2016. Available from: https://www.nice.org.uk/guidance/ta405
60. National Institute for Health and Care Excellence. *Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine*. London: NICE; 2018. Available from: https://www.nice.org.uk/guidance/ta535
61. National Institute for Health and Care Excellence. *Pembrolizumab for treating PDL1-positive non-small-cell lung cancer after chemotherapy*. London: NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta428
62. National Institute for Health and Care Excellence. *Olaratumab in combination with doxorubicin for treating advanced soft tissue sarcoma*. London: NICE; 2017. Available from: https://www.nice.org.uk/guidance/ta465
63. Curtis L, Burns A. *Unit Costs of Health & Social Care 2018*. Canterbury, Kent: Personal Social Services Research Unit, University of Kent; 2018. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/
64. Hodgson R BM, Morgan P, Mebrahtu T, Harden M, Woolacott N. *Crizotinib for untreated anaplastic lymphoma kinase-positive advance non-small-cell lung cancer: A Single Technology Appraisal*: CRD and CHE, University of York; 2016.
65. Council SSA. *Informing the Future of Genomic Medicine in Scotland*; 2019.
66. Department of Health. *Drugs and pharmacuetical electronic market information (eMit)*. 2011. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed 4th July 2019
67. National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management. draft for consultation, October 2018*. London: NICE; 2018. Available from: https://www.nice.org.uk/guidance/ng122/documents/short-version-of-draft-guideline
68. Excellence NIfHaC. *Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy*; 2018.
69. Georghiou T, Bardsley M. *Exploring the cost of end of life care*. London: Nuffield Trust; 2014.
70. Ltd. FH-LR. *Data on file*. In; 2019.
71. Hsiao SJ, Zehir A, Sireci AN, Aisner DL. Detection of tumor NTRK gene fusions to identify patients who may benefit from TRK inhibitor therapy. *J Mol Diagn* 2019. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31075511
72. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* 2017;**23**:703-13. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28481359 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5461196/
73. Sarah Caul JB. *Cancer Registration Statistics, England: 2017*. In: Statistics OfN, editor.; 2019.
74. Nicky Coombes LE-B, Sam Johnson, Jane Lyons, Clara Mackay, Helen Morement, Claudia Oehler and Mick Peake. *Rare and less common cancers: Incidence and Mortality in England, 2010 to 2013*. In: Network NCI, editor. London, UK: Public Health England; 2015.
75. Tracey Genus CB, Kwok Wong, Rajaventhan Srirajaskanthan, Brian Rous, Denis Talbot, Juan Valle, Mohid Khan, Neil Pearce, Mona Elshafie, Nicholas Reed, Tu Vinh Luong, Alia Munir, John Ramage. *Incidence and prevalence of neuroendocrine tumours in England*. In: UKI NETS. London, UK: Endocrine Abstracts; 2017.
76. Cancer Research UK. *Soft tissue sarcoma incidence statistics*. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/soft-tissue-sarcoma/incidence#heading-Four [accessed 27th February 2019].

77. Care Quality Improvement Department RCoP. *National Lung Cancer Audit annual report 2018 (for the audit period 2017)*. In: Physicians RCo, editor. London, UK: Healthcare Quality Improvement Partnership (HQIP) 2019.
78. UK CR. *Uterine cancer incidence by stage at diagnosis* Cancer Research UK; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/uterine-cancer/incidence#heading-Three [accessed Jun 2019].
79. UK CR. *Pancreatic cancer incidence by stage at diagnosis* Cancer Research UK; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/pancreatic-cancer/incidence#heading-Three [accessed Jun 2019].
80. UK CR. *Ovarian cancer incidence by stage at diagnosis* Cancer Research UK; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer/incidence [accessed Jun 2019].
81. UK CR. *Cervical cancer incidence by stage at diagnosis* Cancer Research UK; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/cervical-cancer/incidence#heading-Three [accessed Jun 2019].
82. UK CR. *Bowel cancer incidence by stage at diagnosis* Cancer Research UK; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/soft-tissue-sarcoma/incidence#heading-Zero [accessed Jun 2019].
83. Cancer Research UK. *Breast cancer incidence (invasive) statistics*. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive [accessed 15th May, 2019
84. Board PATE. *PDQ Gastrointestinal Stromal Tumors Treatment*. National Cancer Institute; 2018. Available from: https://www.cancer.gov/types/soft-tissue-sarcoma/hp/gist-treatment-pdq. [accessed June 2019].
85. Austin M, Xing Y, Hayes-Jordan A, Lally K, Cormier JJJops. Melanoma incidence rises for children and adolescents: an epidemiologic review of pediatric melanoma in the United States. 2013;**48**:2207. Available from: https://www.sciencedirect.com/science/article/pii/S0022346813005113?via%3Dihub
86. Berner AM, Morrissey GJ, Murugaesu NJCGMR. Clinical Analysis of Whole Genome Sequencing in Cancer Patients. 2019:1-8.
87. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 1998;**7**:434-55.
88. Gelman A, Rubin DB. Inferences from iterative simulation using multiple sequences. *Statistical Science* 1992;**7**:457-72.

10 Appendices

Appendix A: ERG estimates of eligible population

As the population is defined in the CS as people with *NTRK* fusion-positive [redacted], the following formula was used to estimate the eligible population for each solid tumour type, x:

$$Annual\ eligible\ population = \sum_x FNTRK_x * \ I_x * \ s_x * \ p_x$$

Where FNTRK is the frequency of *NTRK* fusions in a specific tumour type; I is the annual incidence of the tumour type in England; s is the % of that specific tumour type with stage III/IV cancer at diagnosis and p is the position in the treatment pathway. This is done for all of those solid tumours in which an NTRK fusion has been found in the literature. [71]

The frequency of *NTRK* fusions in each specific tumour type are taken from the Foundation Medicine Inc. (FMI) dataset provided to the ERG as a response to clarification questions. This dataset was chosen, as it is a comprehensive set of over [redacted] tumour samples. The *NTRK* gene fusion frequencies for sinonasal adenocarcinoma and renal cell carcinoma were not available in the dataset so it was assumed the frequency was equivalent to that seen in head and neck squamous cell carcinoma and kidney cancer, respectively.

As the FMI data set does not provide sufficient granularity on the frequency of *NTRK* gene fusions within tumours types included in the efficacy evaluable analysis set, further estimates of the frequency of *NTRK* fusions were required. Estimates for MASC, secretory breast carcinoma, papillary thyroid cancer, gastro-oesophageal junction adenocarcinoma, congential mesoblastic nephroma and infantile fibrosarcoma were obtained from the Larotrectinib NDA Multidisciplinary review and evaluation document submitted to the FDA. [54] Estimates of *NTRK* gene fusions for paediatric melanoma and paediatric high and low grade glioma were obtained from Okamura et al. [8]

[72737171](56)The annual incidence for each cancer were primarily obtained from the Office for National Statistics Cancer Registration Statistics [73] and the Rare and Less Common Cancer Statistics. [74757373](58)(48) Annual Incidence data for neuroendocrine tumours, NSCLC and soft-tissue sarcomas were obtained from other sources. [75-77] Stage at diagnosis data were obtained from Cancer Research UK. [78-83] If data were not available for specific subtypes then estimates were obtained from a pragmatic literature search. [75, 77, 84, 85] For tumour types in which a known proportion of the patient population had an unidentified stage at diagnosis, the unidentified proportion was assumed to follow the same distribution as the known proportion.

Finally, to reflect the influence the TRK-inhibitor's proposed/estimated position in the systemic therapy pathway will have on the eligible population, the position was specified for each tumour type and was used to adjust the eligible population. Entrectinib's position in the treatment pathway is proposed for people [redacted]. For those tumour types represented in the efficacy evaluable analysis and with a clear position outlined, the position was assumed to be the same as the one provided in the CS. For those tumour types represented in the efficacy evaluable analysis set but without a clear position, i.e. those with comparator data from multiple lines of therapy, it was conservatively assumed the positioning of the drug was the earliest possible position out of the options provided in the CS.

For those tumours types not represented in the entrectinib CS, it is assumed that entrectinib is positioned as a 3rd line systemic therapy. This decision was made following advice from the ERG's CA and from the identification of the position of chemotherapy, hormone therapy or best supportive care in NICE pathways.

Based on the company's assumption of patients fit enough for treatment, it was assumed for those patients in which entrectinib was first-line, 90% of patients would be eligible. For those using entrectinib as a second-line and third-line therapy, it was assumed 60% and 30% of patients respectively would be eligible.

The annual population eligible for TRK-inhibitors based on tumour types in which an *NTRK* gene fusion has been identified in the literature is 196 patients per year.

Table 58 Summary of data used to estimate annual eligible population

Tumour Type (Low Level)	Frequency of NTRK fusion	Cancer Incidence (England)	% with Stage III/IV Cancer	Position of Entrectinib in line of systemic therapy	Annual TRK-inhibitor eligible population
MASC	100.00%	11	22%	1	2
NSCLC (Adenocarcinoma & squamous cell carcinoma)	[redacted]	32576	57%	2	28
Breast cancer	[redacted]	46102	15%	1	27
Secretory breast carcinoma	91.70%	7	9%	2	0

Papillary thyroid tumour	13.30%	1057	31%	2	26
Thyroid tumour	[redacted]	3254	31%	2	8
Colon/colorectal	[redacted]	34825	55%	2	29
Melanoma	[redacted]	13740	10%	3	1
Neuroendocrine	[redacted]	4363	53%	2	5
Gastrointestinal stromal tumour	[redacted]	734	40%	1	2
Cholangiocarcinoma	[redacted]	556	60%	1	1
Pancreatic	[redacted]	8388	78%	1	12
Appendix	[redacted]	540	74%	3	0
Uterine	[redacted]	7862	18%	1	2
Ovarian	[redacted]	2724	55%	1	4
Cervix	[redacted]	2591	24%	1	1
Soft tissue sarcoma	[redacted]	2740	32%	1	14
Head and neck squamous cell carcinoma	0.24%	9946	63%	3	5
Salivary gland (non MASC)	2.69%	517	63%	3	3
Sinonasal adenocarcinoma	0.24%	4	63%	3	0
Gastro-esophageal junction	0.10%	7569	73%	3	2
Prostate cancer	0.23%	41201	43%	3	12
Renal cell carcinoma	0.07%	7438	43%	3	1
Low-grade glioma	0.42%	929	0%	3	0
High grade glioma (inc. glioblastoma multiforme)	0.42%	2781	100%	3	4
Paediatric high grade glioma	5.30%	67	100%	3	1
Congenital mesoblastic nephroma	60.70%	0	0%	3	0
Paediatric melanoma	11.11%	56	34%	3	1
Infantile fibrosarcoma	90.90%	59	51%	3	8
Paediatric low grade glioma	2.50%	723	0%	3	0
Total:					**196**

*The frequency NTRK fusions in appendix tumours in the FMI data set was reported to be 0%, however it has been reported to be higher than 0% in the literature [2]

Note, totals in the annual TRK-inhibitor eligible population column may add up to greater less than 196 due to rounding.

MASC, mammary analogue secretory carcinoma; NSCLC non-small cell lung cancer

Appendix B: Numbers needed to screen

The number of patients who require screening to identify one individual with an *NTRK* fusion varies depending on the prevalence of gene rearrangement. Table 59 presents the number of patients who need to be screened to identify one individual with a *NTRK* fusion. This is calculated by the following equation:

$$\text{NNS: } \frac{1}{\textit{NTRK} \text{ fusion rate}}$$

According to the company submission, the NNS to identify one patient with an *NTRK* rearrangement varies between 1 (MASC, *NTRK* prevalence = 100%) and 1250 (Pancreatic Cancer, *NTRK* prevalence = 0.08%). ERG estimates of the NNS to identify one *NTRK*+ patient for all tumour types mentioned in the CS range from 1 (MASC, Secretory Breast Carcinoma and Infantile Fibrosarcoma) to 2000 (High Grade Glioma, *NTRK* prevalence = 0.05%). Due to discrepancies in the recorded prevalence of NTRK fusions, the NNS reported in the CS differ to the NNS calculated by the ERG.

Table 59 Number needed to screen: company and ERG estimates

Tumour Type	Prevalence of *NTRK* fusion (ERG)	Number Needed to Screen (ERG)	Number Needed to Screen (Company)
Salivary gland (MASC)	[redacted]	[redacted]	[redacted]
NSCLC	[redacted]	[redacted]	[redacted]
Breast cancer (not specified)	[redacted]	[redacted]	[redacted]
Secretory breast carcinoma	[redacted]	[redacted]	[redacted]
Papillary thyroid tumour	[redacted]	[redacted]	[redacted]
Thyroid Tumour (NOS)	[redacted]	[redacted]	[redacted]
Colon/colorectal	[redacted]	[redacted]	[redacted]
Neuroendocrine (NOS)	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]
Pancreatic	[redacted]	[redacted]	[redacted]
Uterine	[redacted]	[redacted]	[redacted]
Ovarian	[redacted]	[redacted]	[redacted]
Cervix	[redacted]	[redacted]	[redacted]
Soft tissue sarcoma	[redacted]	[redacted]	[redacted]
High grade glioma	[redacted]	[redacted]	[redacted]
Paediatric high grade glioma	[redacted]	[redacted]	[redacted]
Congenital mesoblastic nephroma	[redacted]	[redacted]	[redacted]
Paediatric melanoma	[redacted]	[redacted]	[redacted]
Infantile fibrosarcoma	[redacted]	[redacted]	[redacted]

Paediatric low grade glioma	[redacted]	[redacted]	[redacted]

Using the calculated estimates of the annual population eligible for a TRK inihibitor, the ERG estimated the total patient population that would require IHC and NGS screening to identify individuals eligible for entrectinib. This is equal to the annual population of individuals in England with [redacted].

$$\text{Population Requiring IHC screening}_x = I_x \times s_x \times p_x$$

where *x* is the tumour type in which an *NTRK* fusion has been identified; I_x is the annual incidence of the tumour type in England; s_x is the % of patients with that specific tumour type who have stage III/IV cancer at diagnosis and p_x is the position of the therapy in the treatment pathway.

ERG calculation of the annual population who require IHC screening, based on the tumours where an *NTRK* fusion has been identified, indicate that approximately 51,958 patients would need IHC screening to identify potential individuals eligible for entrectinib.

According to the CS, IHC will identify 89% of *NTRK* fusion negative individuals, hence 11% of the population screened with IHC will require confirmatory NGS screening. Whole genome sequencing also requires confirmatory RNA-based NGS; as the diagnostic accuracy of WGS is unclear, it was assumed that 11% of individuals screened with WGS would require confirmatory RNA-based NGS. [86]

$$\text{Population Requiring NGS screening}_x = (I_x \times s_x \times p_x) \times 0.11$$

The ERG calculated annual population requiring NGS screening, based on the tumours where an *NTRK* fusion has been identified, indicates that approximately 5,806 patients would need confirmatory NGS to identify the patients eligible for entrectinib.

Table 60: Annual population requiring IHC or NGS screening in order to identify patients with an NTRK fusion.

Tumour Type (Low Level)	Annual TRK-inhibitor eligible population	Population requiring IHC screening	Population Requiring NGS Screening
Salivary gland (MASC)	2	-	-
NSCLC Lung (Adenocarcinoma & squamous cell carcinoma)	9	11141	1226
Breast cancer (not specified)	4	6224	685
Secretory breast carcinoma	1	0	0

Papillary thyroid tumour	24	197	22
Thyroid tumour (NOS)	5	605	67
Colon/colorectal	12	11492	1264
Melanoma (NOS)	2	412	45
Neuroendocrine (NOS)	4	1387	153
Gastrointestinal stromal tumour	3	264	29
Cholangiocarcinoma	0	300	33
Pancreatic	15	5888	648
Appendix	9	120	13
Uterine	1	1274	140
Ovarian	3	1348	148
Cervix	2	560	62
Soft tissue sarcoma	4	-	87
Head and neck squamous cell carcinoma (NOS)	21	1880	207
Salivary gland (non MASC)	3	98	11
Sinonasal adenocarcinoma	0	1	0
Gastro-oesophageal junction	5	1658	182
Prostate cancer	24	5315	585
Renal cell carcinoma	4	960	106
High grade glioma (inc. glioblastoma multiforme)	1	834	92
Paediatric high grade glioma	3	-	2
Congenital mesoblastic nephroma	0	-	0
Paediatric melanoma	0	-	1
Infantile fibrosarcoma	25	-	1
Paediatric low grade glioma	0	-	0
Total	**188**	**51958**	**5806**

Appendix C: Comparator evidence

See Excel file.

Appendix D: STARTRK-2 Quality assessment checklist

Checklist		STARTRK-2 (Interim CSR)
Reporting	Is the hypothesis/aim/objective of the study clearly described?	yes
	Are the main outcomes to be measured clearly described in the Introduction or methods section?	yes
	Are the characteristics of the patients included in the study clearly described?	yes
	Are the interventions of interest clearly described?	yes
	Are the distributions of principal confounders in each group of subjects to be compared clearly described?	no
	Are the main findings of the study clearly described?	yes
	Does the study provide estimates of the random variability in the data for the main outcomes?	yes
	Have all important adverse events that may be a consequence of the intervention been reported?	yes
	Have the characteristics of patients lost to follow-up been described?	yes
	Have actual probability values been reported for the main outcomes except when the probability is less than 0.001?	no
External validity	Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	no
	Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	no
	Were the staff, places, and facilities where the patients were treated representative of the treatment the majority of patients receive?	yes
Internal validity - bias	Was an attempt made to blind study subjects to the intervention they have received?	no
	Was an attempt made to blind those measuring the main outcomes of the intervention?	no
	If any of the results of the study were based on "data dredging", was this made clear?	yes
	In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	yes
	Were the statistical tests used to assess the main outcomes appropriate?	yes
	Was compliance with the interventions reliable?	yes
	Were the main outcome measures used accurate?	yes
	Were the patients in different intervention groups recruited from the same population?	no

Internal validity - confounding (selection bias)	Were study subjects in different intervention groups recruited over the same period of time?	yes
	Were study subjects randomised to intervention groups?	no
	Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	no
	Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	no
	Were losses of patients to follow-up taken into account?	yes
Power	Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	no

Appendix E: Analysis of response heterogeneity - methods and additional results

For all analyses 55,000 iterations were run on 2 parallel chains and the first 5,000 iterations discarded as "burn-in". Convergence was assessed by visual inspection of the Brooks-Gelman-Rubin plots and assessment of the $\hat{R}$ statistic.[87, 88]

Table 61 shows the posterior probabilities estimated by the base-case BHM, using BIRC-assessed data with imputation and the prior distribution is equation (1).

Table 61 Posterior probabilities of response for all tumour types (BIRC-assessed data with imputation)

	Tumour type	mean	median	95%CrI
1	Sarcoma	[redacted]	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]	[redacted]
9	Breast	[redacted]	[redacted]	[redacted]
10	Thyroid	[redacted]	[redacted]	[redacted]
11	CNS Primary	[redacted]	[redacted]	[redacted]
12	Paediatric CNS Primary	[redacted]	[redacted]	[redacted]
13	Paediatric (non-CNS)	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3;
Uniform prior distribution for the between-tumour standard deviation

Table 62 has the model fit statistics for the base-case and all sensitivity analyses, which show that all models fit the data well. Inspection of box-plots of individual groups' contributions to the residual deviance (not shown) support this.

Table 62 Model fit statistics for the base-case and sensitivity analyses.

	Posterior mean of the residual deviance	DIC
Base-case (all patients, prior for response rate centred on 0.3, Uniform prior for heterogeneity)	11.8*	40.1
Sensitivity analysis 1 (all patients, prior for response rate centred on 0.3, inverse-gamma prior for heterogeneity)	10.9*	44.3
Sensitivity analysis 2 (all patients, prior for response rate centred on 0.5, Uniform prior for heterogeneity)	11.9*	40.0
Sensitivity analysis 3 (no primary CNS or paediatric patients, prior for response rate centred on 0.3, Uniform prior for heterogeneity)	9.1†	31.3

* compare to 13 groups; † compare to 10 groups

Sensitivity analysis 1

An Inverse Gamma(2, 20) prior distribution for the between-tumour variance was used, instead of the Uniform prior on the between-tumour standard deviation. This means the between-tumour precision has prior mean 0.10 and variance 0.005, which implies that the between-tumour standard deviation has a prior mean ≈ 3.97 and variance ≈ 4.33. Note that both the prior mean and variance are higher than those implied by the Uniform(0,5) prior distribution, which are 2.5 and 2.08, respectively. This results in a much larger estimate of the between-tumour heterogeneity with median [redacted] and 95% CrI ([redacted]). Figure 22 shows the prior and posterior distributions for the between-tumour heterogeneity. The prior and posterior distributions in the base-case are also included for comparison. We can see that the inverse-gamma prior distribution places much more weight on large values of heterogeneity and does not allow for values close to zero, which is then reflected in the posterior distribution, which also excludes small values.

Figure 22 Sensitivity analysis 1: [redacted]

[redacted]*Figure redacted*

The very large estimated heterogeneity means that the 95% CrI for the probability of response is much wider than in the base-case, and the predictive interval for the response rate for an unrepresented tumour type covers nearly the whole range of probabilities from zero to 1 (Table 63).

Table 63 Sensitivity analysis 1: Probabilities of response according to the BHM.

	Overall posterior probability of response		
	mean	median	95% CrI

Posterior probability of response	[redacted]	[redacted]	[redacted]
Predictive probability of response	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3;
Inverse-gamma prior distribution for the between-tumour variance.

Given that the inverse-gamma prior is not derived from genuine prior beliefs that low levels of heterogeneity are not plausible, the ERG caution against the results from this analysis. However, for completeness, the distribution of the response rates for each tumour type are shown in Table 64 and Figure 23.

Table 64 Sensitivity analysis 1: Probabilities of response for all tumour types (IRC-assessed data with imputation)

	Tumour type	observed response (%)	Estimated mean response based on BHM (%)	Prob of response rate at least 30%	Prob of response rate at least 10%
1	Sarcoma	[redacted]	[redacted]	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]	[redacted]	[redacted]
9	Breast	[redacted]	[redacted]	[redacted]	[redacted]
10	Thyroid	[redacted]	[redacted]	[redacted]	[redacted]
11	CNS Primary	[redacted]	[redacted]	[redacted]	[redacted]
12	Paediatric CNS Primary	[redacted]	[redacted]	[redacted]	[redacted]
13	Paediatric (non-CNS)	[redacted]	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3; Inverse-gamma prior distribution for the between-tumour variance

Figure 23 Sensitivity analysis 1: [redacted]

Figure redacted

Sensitivity analysis 2

Results using a more favourable prior distribution for the log-odds of response, centred on an *a priori* probability of response of 50%, are presented in this section. The prior distributions used in this sensitivity analysis are

$$\mu \sim \text{Normal}(0,10)$$
$$\sigma \sim \text{Uniform}(0,5)$$

The BHM estimates moderate between-group heterogeneity, similar to the base-case (posterior median [redacted] 95% CrI ([redacted])).

The estimated mean response rate across all tumour types and the predictive probabilities are similar to the base-case (Table 65) and the estimated probabilities of response for each tumour type in Table 66 are almost identical to the results obtained in the base-case (Table 61). We therefore conclude that the prior distribution for the mean probability of response does not have any meaningful impact on the results.

Table 65 Sensitivity analysis 2: Probabilities of response according to the BHM.

	Overall posterior probability of response		
	mean	median	95% CrI
Posterior probability of response	[redacted]	[redacted]	[redacted]
Predictive probability of response	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.5;
Uniform prior distribution for the between-tumour standard deviation

Table 66 Sensitivity analysis 2: Posterior probabilities of response for all tumour types

	Tumour type	**mean**	**median**	**95%CrI**
1	Sarcoma	[redacted]	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]	[redacted]
9	Breast	[redacted]	[redacted]	[redacted]
10	Thyroid	[redacted]	[redacted]	[redacted]
11	CNS Primary	[redacted]	[redacted]	[redacted]
12	Paediatric CNS Primary	[redacted]	[redacted]	[redacted]
13	Paediatric (non-CNS)	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3;
Uniform prior distribution for the between-tumour standard deviation

Sensitivity analysis 3

Results excluding primary CNS and paediatric patients are presented in this section. Note that this analysis includes only BICR-assessed data. The prior distributions used in this sensitivity analysis are given in equation (1).

The BHM estimates moderate between-group heterogeneity, similar to the base-case although with a wider 95%CrI due to less tumour types being included (posterior median [redacted] 95% CrI ([redacted])).

The estimated mean response rate across all tumour types and the predictive probabilities are similar to the base-case (Table 67) and the estimated probabilities of response for each tumour type in Table 68 are only slightly larger than the results obtained in the base-case (Table 61, including all tumour types). We therefore conclude that there are similar amounts heterogeneity across adult non-primary CNS tumours as across all tumour types, and a similar amount of uncertainty in the response rate expected in an unrepresented adult solid tumour.

Table 67 Sensitivity analysis 3: Probabilities of response according to the BHM.

	Overall posterior probability of response		
	mean	median	95% CrI
Posterior probability of response	[redacted]	[redacted]	[redacted]
Predictive probability of response	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.5;
Uniform prior distribution for the between-tumour standard deviation

Table 68 Sensitivity analysis 3: Posterior probabilities of response for all tumour types

	Tumour type	**mean**	**median**	**95%CrI**
1	Sarcoma	[redacted]	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]	[redacted]
9	Breast	[redacted]	[redacted]	[redacted]
10	Thyroid	[redacted]	[redacted]	[redacted]

Prior distribution for log-odds of response centred on a probability of 0.3;
Uniform prior distribution for the between-tumour standard deviation

OpenBUGS code

Bayesian Hierarchical Model: Uniform(0,5) prior distribution for the between-tumour standard deviation

```
# CODE ADAPTED FROM: Thall et al (2003)
# Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes.
# Statist. Med., 22: 763-780. doi:10.1002/sim.1399
#
# Uniform prior distribution for between-group SD, as recommended by Cunanan et al. (Clinical
Trials, 2019)
#
model{
for (i in 1:numGroups){  # numGroups is k, the number of different probabilities
  x[i] ~ dbin(p[i],n[i]) # In each group, x is the number of responses and n is the number of
patients
  # set up deviance code with correction for zero cells
  x1[i] <- max(x[i],0.1) # zero cell correction
  xhat[i] <- p[i] * n[i] # expected value of the numerators
  xhat1[i] <- max(xhat[i], 0.1) # zero cell correction
  # Deviance contribution with zero cell correction
  dev1[i] <- 2 * (x1[i] * (log(x1[i])-log(xhat1[i]))
            +  (n[i]-x1[i]) * (log(n[i]-x1[i]) - log(n[i]-xhat1[i])))
  # deviance contribution for for zero cells
  dev0[i] <- 2 * n[i] * log(n[i]/(n[i]-xhat[i]))
  # deviance contribution
  dev[i] <- dev1[i] * (1-equals(x[i],0)) + dev0[i] * equals(x[i],0)
  # logit model for p
  logit(p[i]) <- rho[i]
  rho[i] ~ dnorm(mu,tau) # RE for log-odds
  # Probability that the response rate for each group is > than targetResp (given as data)
  pg[i] <- step(p[i] - targetResp)
  pg2[i] <- step(p[i] - targetResp2)
 }
totresdev <- sum(dev[])            # total residual deviance
# Priors
mu ~ dnorm(mean.Mu, perc.Mu)       # pooled mean of log-odds
#tau ~ dgamma(tau.alpha, tau.beta) # used in Thall (2003)
#sd <- 1/sqrt(tau)                 # between-group sd (log-odds scale)
sd ~ dunif(0,5)                    # recommended by Cunanan (2019)
tau <- pow(sd,-2)
# predictive distribution
rho.new ~ dnorm(mu,tau)            # log-odds response across groups
# convert to probabilities
logit(p.pooled) <- mu    # mean probability of response across groups
logit(p.new) <- rho.new  # probability response across groups
# predictive probabilities of response rates > targetResp (given as data)
pg.new <- step(p.new - targetResp)
pg2.new <- step(p.new - targetResp2)
}
```

Data

list(x=c(), n=c(), numGroups=13, mean.Mu=-0.847298, perc.Mu=0.1, targetResp=0.3, targetResp2=0.1)

Appendix F: Drummond Checklist

Table 69 presents the quality assessment.

Table 69 Quality assessment of included CEA study using Drummond et al. checklist completed by the ERG

	CEA quality assessment questions	Answer (Yes/No/Unclear)	Notes/Explanation for No or Unclear
1	Was the research question stated?	Yes	-
2	Was the economic importance of the research question stated?	Yes	-
3	Was/were the viewpoint(s) of the analysis clearly stated and justified?	Yes	-
4	Was a rationale reported for the choice of the alternative programmes or interventions compared?	Yes	-
5	Were the alternatives being compared clearly described?	No	Many of the interventions used to construct the weighted average comparator were described but no comparators were provided in the 'other TBC' resulting in a lack of clarity as to what the alternatives are.
6	Was the form of economic evaluation stated?	Yes	-
7	Was the choice of form of economic evaluation justified in relation to the questions addressed?	Yes	-
8	Was/were the source(s) of effectiveness estimates used stated?	Yes	-
9	Were details of the design and results of the effectiveness study given (if based on a single study)?	Yes	-
10	Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of a number of effectiveness studies)?	Yes	-
11	Were the primary outcome measure(s) for the economic evaluation clearly stated?	Yes	-

12	Were the methods used to value health states and other benefits stated?	Yes	-
13	Were the details of the subjects from whom valuations were obtained given?	Yes	-
14	Were productivity changes (if included) reported separately?	No	Excluded from base-case analysis. Model did provide option to include productivity losses.
15	Was the relevance of productivity changes to the study question discussed?	No	Not mentioned
16	Were quantities of resources reported separately from their unit cost?	Yes	-
17	Were the methods for the estimation of quantities and unit costs described?	Yes	-
18	Were currency and price data recorded?	Yes	-
19	Were details of price adjustments for inflation or currency conversion given?	Unclear	End of life costs were adjusted for inflation, however no other parameters seem to have been adjusted for inflation/currency conversion.
20	Were details of any model used given?	No	-
21	Was there a justification for the choice of model used and the key parameters on which it was based?	Yes	-
22	Was the time horizon of cost and benefits stated?	Yes	-
23	Was the discount rate stated?	Yes	-
24	Was the choice of rate justified?	Yes	-
25	Was an explanation given if cost or benefits were not discounted?	No	-

26	Were the details of statistical test(s) and confidence intervals given for stochastic data?	No	CI for stochastic data were provided however, the CIs for the comparator data were all assumed and not taken from the original sources. Formal significance tests were not performed.
27	Was the approach to sensitivity analysis described?	Yes	-
28	Was the choice of variables for sensitivity analysis justified?	No	Justified for some of the variables, but testing costs were not explored thoroughly and this was a big driver of the cost-effectiveness.
29	Were the ranges over which the parameters were varied stated?	Yes	-
30	Were relevant alternatives compared? (That is, were appropriate comparisons made when conducting the incremental analysis?)	No	It is unclear what the relevant comparators are for the population considered in the marketing authorisation as not only are there unrepresented tumour types but it's unclear where entrectinib will be used in the treatment pathway.
31	Was an incremental analysis reported?	Yes	-
32	Were major outcomes presented in a disaggregated as well as aggregated form?	Yes	-
33	Was the answer to the study question given?	Yes	-
34	Did conclusions follow from the data reported?	Yes	-
35	Were conclusions accompanied by the appropriate caveats?	No	-
36	Were generalisability issues addressed?	No	Issues of generalisability were not fully addressed. Unclear whether the entrectinib population and the comparator population are generalisable due to unrepresented tumour types; proportion of tumour types used to weight comparator costs and utilities; prior therapies; underrepresentation of *NTRK2* fusions in the CS and issues of where in the treatment pathway entrectinib will be used.

wan

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

ERG report – factual accuracy check

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Tuesday 6 August 2019** using the below comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 The company's proposal to proactively enter the Cancer Drugs Fund (CDF) with a comprehensive data collection plan, a key element of the appraisal, is largely omitted in the report

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 1, pages 14-24: The summary makes no reference to Roche's efforts to address clinical and cost-effectiveness uncertainties through proactive entry into the CDF and a comprehensive associated data collection proposal.	Add a concluding paragraph to section 1.7 to cover this topic, such as "It is important to note that the company acknowledges many of the uncertainties raised and is therefore proposing proactive entry into the Cancer Drugs Fund with a substantial data collection plan in addition to routine trial follow-up and SACT data collection."	This is an important omission in this section, and this information should be provided in the interest of transparency and balance so that readers are informed of this key consideration up front. The CDF proposal, which is pivotal for this appraisal, is only referenced once in section 3.5.	Not a factual error. The ERG appreciates the particular challenges associated with this appraisal and the company's proposed entry the CDF to address key uncertainities . However, the efforts made by the company and the proposed entry into the CDF lie beyond the scope of the ERG report. The proposed changes are therefore not appropriate.
Section 8, pages 161-163. As highlighted in the summary the ERG report conclusion section makes no reference to Roche's efforts to address clinical and cost-effectiveness uncertainties through proactive entry into the Cancer Drugs Fund and a comprehensive associated data collection proposal.	As detailed for the ERG summary the company would suggest the addition of a paragraph within section 8 to cover this topic, such as "It is important to note that the company acknowledges many of the uncertainties raised and is therefore proposing proactive entry into the Cancer Drugs Fund with a substantial data collection plan in addition to routine trial follow-up and SACT data collection."	As highlighted for the ERG summary the company feels this is an important omission in this key section. This information should be provided in the interest of transparency and balance so that readers are informed of this key consideration.	See above.

Issue 2 ERG assertion that no attempt was made to adjust for prognostic factors in the naïve weighted comparison

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 1.2, page 15, last paragraph: it is stated that "No attempts were made to adjust for differences in population	Amend sentence to: "Important prognostic factors such as the impact of NTRK fusions and brain metastases on outcomes were explored in scenario analyses".	Factual inaccuracy.	Text edited to clarify that these concerns relate to the base-case analysis (Section 1.2 page 15)

characteristics between the entrectinib and comparator trial populations; comparisons were naïve and did not account for any potentially important prognostic factors". This is factually inaccurate as the key prognostic factors of presence/absence of brain metastases and NTRK fusions were explored in scenario analyses.			
Section 3.3, page 43, second paragraph: It is stated that "These comparisons do not account for any potentially important patient characteristics, such as age, performance status, NTRK fusion status or the prevalence of CNS metastases". This is factually inaccurate as the key prognostic factors of presence/absence of brain metastases and NTRK fusions were explored in scenario analyses.	Amend sentence to "These comparisons do not account for potentially important patient characteristics, such as age, performance status, though NTRK fusion status and the prevalence of CNS metastases were explored in scenario analyses."	Factual inaccuracy.	Not a factual error. It is clear from the context that we are referring to the base-case comparison.
Section 3.3, page 49, third paragraph: It is stated that "Comparisons were naïve and do not account for any potentially important prognostic factors, such as age, performance status, NTRK fusion status, prevalence of CNS metastases, or specific tumour mutations within each tumour type. No attempts were made to adjust for differences in population	Amend sentences to "Comparisons were naïve and do not account for all potentially important prognostic factors, such as age, performance status, or specific tumour mutations within each tumour type. NTRK fusion status and prevalence of CNS metastases were explored in scenario analyses."	Factual inaccuracy.	Text edited to clarify that these concerns relate to the base-case analysis (Section 3.3, page 49).

characteristics between the entrectinib and comparator trial populations." This is factually inaccurate as the key prognostic factors of presence/absence of brain metastases and NTRK fusions were explored in scenario analyses.			
Section 4.4, page 82, second paragraph: It is stated that "No attempts were made to adjust for differences in population characteristics between the entrectinib and comparator trial populations; comparisons were naïve and do not account for any potentially important prognostic factors." This is factually inaccurate as the key prognostic factors of presence/absence of brain metastases and NTRK fusions were explored in scenario analyses.	Amend sentence to: "Important prognostic factors such as the impact of NTRK fusions and brain metastases on outcomes were explored in scenario analyses".	Factual inaccuracy.	Text edited to clarify that these concerns relate to the base-case analysis (Section 4.4, page 82).

Issue 3 Misleading over-interpretation of post-hoc exploratory analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 1.3, page 16, first paragraph: It is stated that "Despite substantial censoring and known limitations of responder analyses, the crossing of OS curves between entrectinib responders and non-responders	Remove sentence: "~~Despite substantial censoring and known limitations of responder analyses, the crossing of OS curves between entrectinib responders and non-responders is concerning.~~"	Misleading over-interpretation of highly uncertain data with substantial limitations.	Text edited to emphasise the limitations of the available data (Section 1.3, page 16).

is concerning." Given the extensive and significant limitations of this data and the responder analysis approach, this is an extremely unreliable and misleading over-interpretation of the KM data. Also applies to: Section 4.2.6.1, page 58/59 Section 4.4, page 82, third paragraph Section 8, page 161			

Issue 4 Incorrect number of tumour types specified for the integrated analysis data set

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 1.4, page 17, paragraph 3: "The integrated analysis set included 54 patients across 13 different tumour types"	Amend sentence to: "The integrated analysis set presented within the CS included 54 patients across 10 different tumour types"	Factual inaccuracy.	The ERG acknowledges that the CS states that there are 10 tumour types. This classification, however, considers it important to differentiate between the gynaecological, thyroid and breast cancers included in the integrated efficacy analysis. An additional sentence has been added to Section 3.1 (pg. 35) to acknowledge the different positions.
Section 3.1, page 35, paragraph 1: "The clinical efficacy evidence submitted by the company included Trk inhibitor-naïve patients with NTRK fusion-	Please amend to: "The primary clinical efficacy population (n=54) presented within the CS included Trk inhibitor-naïve patients with NTRK fusion-positive solid tumours that is limited to 10 tumour types included in the entrectinib	Factual inaccuracy.	See above.

positive solid tumours that is limited to 13 tumour types included in the entrectinib clinical trials." The original primary efficacy population presented by the company (n=54) covered 10 tumour types.	clinical trials."		
Section 5.2.6, page 101, paragraph 4: "The integrated analysis set included 54 patients across 13 different tumour types"	Amend sentence to: "The integrated analysis set presented within the CS included 54 patients across 10 different tumour types"	Factual inaccuracy.	See above.

Issue 5 Updated clinical cut-off date (CCOD) not marked as AIC

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 1.4, page 17, paragraph 3: Updated CCOD is not highlighted as AIC.	Mark date as AIC	CCOD is currently confidential.	Data cut-off date now marked as AIC (Section 1.4, page 17).
Section 4.2.6.1, page 64, figure 4. Updated CCOD in figure legend and heading not marked as AIC.	Please mark data cut-off date in Figure 4 legend as AIC.	CCOD is currently confidential.	Data cut-off date now marked as AIC (Section 4.2.6.1, page 64).
Section 4.2.6.1, page 65, figure 5. Updated CCOD in figure heading not marked as AIC.	Please mark data cut-off date in Figure 5 heading as AIC.	CCOD is currently confidential.	Data cut-off date now marked as AIC (Section 4.2.6.1, page 65).
Section 4.2.6.1, page 68, figure 6 / page 69 figure 7. Updated CCOD in figure heading not marked as AIC.	Please mark data cut-off date in Figure 6 heading as AIC.	CCOD is currently confidential.	Data cut-off date now marked as AIC (Section 4.2.6.1, page 69).

Issue 6 Inconsistent deterministic versus probabilistic ICERs presented

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 1.4, page 18, last paragraph: The original	Amend deterministic ICER to £52,609 per the updated base case.	Factual inaccuracy.	Text changed as proposed (Section 1.4, page 18).

submission base case deterministic ICER of £54,609 is given while the updated base case probabilistic ICER of £52,052 is given.			

Issue 7 Inaccuracies relating to NTRK fusion testing

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 2.2.2.1, page 29, third paragraph: It is stated that "Immunohistochemistry detects NTRK fusions by proxy by identifying overexpression of Trk proteins." IHC does not detect fusions specifically, it only detects the presence of Trk proteins.	Amend sentence "Immunohistochemistry detects NTRK fusions by proxy by identifying overexpression of Trk proteins" to "Immunohistochemistry detects overexpression of Trk proteins, a subset of which may be the result of NTRK gene fusions."	Factual inaccuracy.	Text changed as proposed (Section 2.2.2.1, page 29).
Section 2.2.2.2, page 30, paragraph 1: It is stated that "There is currently no established strategy for detecting for NTRK fusions across tumour types in the NHS. With the exception of MASC, where patients receive the ETV6-NTRK3 FISH test, NTRK fusions are not routinely tested for." NTRK targets are specified within the Testing Directory against infantile fibrosarcoma, spindle cell soft tissue tumour, myeloproliferative neoplasm, histiocytosis and congenital mesoblastic nephroma in addition to MASC.	Amend sentence to: "There is currently no established strategy for detecting for NTRK fusions across tumour types in the NHS. With the exception of MASC, infantile fibrosarcoma, spindle cell soft tissue tumour, myeloproliferative neoplasm, histiocytosis and congenital mesoblastic nephroma, NTRK fusions are not routinely tested for."	Factual inaccuracy.	Revised sentence to clarify that except for where WGS is available and for some haematological tumours, NTRK fusion testing is only routinely implemented for MASC (Section 2.2.2.2, page 30).
Section 2.2.2.2, page 30, paragraph 3: It is stated that "The company	Amend sentence to: "The company suggest the use of the Roche Ventana pan-TRK assay for	Misleading representation of performance of NTRK IHC assay,	Text revised to clarify that the poorer sensitivity in NTRK3

suggest the use the Roche Ventana pan-TRK assay for IHC testing, which detects Trk proteins A, B, and C, identifying any NTRK gene rearrangement in one test. The CS states that Ventana assay eliminates 89% of NTRK fusion-negative samples. However, the sensitivity of pan-TRK IHC has been estimated to be as low as 50% for NTRK3 fusions.7, 19. Thus, up to half of the individuals with an NTRK3 rearrangement could incorrectly test negative.". This is a potentially misleading statement, as both references refer to overall sensitivity and specificity figures in excess of 95%.	IHC testing, which detects Trk proteins A, B, and C, identifying any NTRK gene rearrangement in one test. The CS states that Ventana assay eliminates 89% of NTRK fusion-negative samples. The sensitivity and specificity of the assay have been shown to be 95.2% and 100%, respectively, though based on limited case numbers it could drop to 50% for NTRK3 fusions.7, 19".	context of overall assay performance needed.	rearrangements was found in other pan-TRK assays and not necessarily the Ventana pan-TRK assay (Section 2.2.2.2, page 30). The ERG do not consider that it is necessary to report overall assay performance and make it clear that the poorer sensitivity is only seen in NTRK3 rearrangements. Reference 19 removed due to lack of clarity regarding sensitivity of IHC for NTRK3. Text amended to reflect this change. (Section 2.2.2.2, page 30).
Section 2.2.2.2, page 30 paragraph 4. It is stated that "NTRK-fusion panels cannot be added to DNA-based NGS panels,12 so for the majority of cancer types where NGS is currently offered, additional RNA fusion-based or RNA/DNA hybrid NGS will be required to identify NTRK rearrangements." This sentence is unclear and incorrect. The ability to add new target content to a DNA based NGS panel is dependent on the panel type; ie custom seq. capture, amplicon based versus commercially available CE-IVD panels.	Amend sentence to reflect complexity of technical aspects highlighted, or delete.	Clarification to reflect complexity of NGS testing methods.	Sentence revised to better reflect the ERG's concerns are about the ability of DNA-based NGS to detect NTRK rearrangements (Section 2.2.2.2, page 30-31).

Section 5.2.8.7, page 128, paragraph 4: "The ERGs clinical advisers stated NGS would be needed after WGS to confirm an NTRK-fusion positive tumour." This sentence is technically unclear. NGS, Next-Generation Sequencing, technically refers to sequencing technologies which superceded Sanger sequencing. WGS is carried out using NGS. It is therefore unclear what is being referred to in the use of "NGS" in this sentence. Is this referring to DNA panels, RNA panels, sequence capture etc? Throughout the ERG report the use of "NGS" is ill defined in terms of methodology being referred to.	Please amend to clearly define which type of sequencing is being run on NGS in confirmation of WGS. Is this referring to a specific targeted methodology? If so please define.	Request for clarification of testing methods.	Sentence amended to define type of NGS sequencing (Section 1.8 pages 22-23; Section 2.2.2.2, page 30; Section 5.2.8.7, page 128; Section 6.3.2, page 146-149; Appendix B, page 174).
Section 5.2.8.7, page 127, first paragraph: "which suggested that the IHC testing approach will remove 89% of NTRK-fusion negative samples". IHC testing is intended to screen out patients with tumours that are negative for any NTRK mutation, not just NTRK fusions.	Amend sentence to "which suggested that the IHC testing approach will remove 89% of patients with tumours negative for any type of NTRK mutation."	Factual inaccuracy.	Not a factual inaccuracy. The ERG considers that the wording is clear and note that the exact same wording is used in the CS (Section B.3.5.4, page 127).
Section 2.2.2.2, page 32, first paragraph: It is stated that "RNA-based NGS fusion panels are available on the NHS for NSCLC, targeting a range of genes including NTRK1, ALK, and ROS1". Upon confirmation of the availability of the stated fusion panel within the	Please correct or clarify statements regarding the RNA-based NGS fusion panels available for NSCLC.	Factual inaccuracy.	Text amended to clarify that the use of RNA-based NGS fusion panels is not routinely used on the NHS and currently does not detect NTRK fusions (Section 2.2.2.2, page 32; Section 5.2.8.7, page 128; Section

Testing Directory we can find no reference to a panel that includes NTRK1 fusions against Non-Small Cell Lung (CI Code M4). In addition, Roche is not aware of this putative panel being routinely used in clinical practice. Also applies to: Section 5.2.8.7, page 128, fifth paragraph. Section 6.3.2, page 146, second paragraph.			6.3.2, page 146).
Section 5.2.8.7, page 130, first paragraph: It is stated that "The ERG considers the inclusion of screening costs included in the comparator arm, as seen in Table 41, to be inappropriate. The focus of testing costs should be on the incremental testing costs associated with identifying NTRK fusion positive patients and therefore the ERG suggests the comparator testing costs should be removed unless current testing is able to identify NTRK fusions." Testing costs for other biomarkers besides NTRK should be included in the comparator arm, since the NGS testing applied to find NTRK+ patients is also likely to identify patients with other driver mutations such as EGFR, ALK, BRAF etc. The testing pathway implemented to identify NTRK+ patients would	Acknowledge rationale for inclusion of testing costs in comparator arm and overestimation of testing costs for primary brain tumours due to inclusion of paediatric patients in this group.	Error in interpretation of application of testing costs in model, clarification of issue relating to primary CNS tumour testing costs.	The ERG have amended the error made in the company's revised base-case model and updated the ICERs reflect the alternative assumptions (Section 1.8, pages 23-24; Section 6.3.2, page 148; Section 6.3.3, page 148-149; Section 6.4, page 149 -150, Section 6.5.1, page 153-154; Section 6.5.2, page154-155, Section 6.6, page 156; Section 8, page 162). Additional text has been added to the report to reflect these changes (Section 6.3.2, page 147). Note some Table number have also changed as a result of this new analysis.

therefore duplicate or supplant tests already in place to identify other driver mutations. In addition, it should be noted that screening costs for glioma are overestimated since the costs represent a mixture of adult (five) and paediatric (four) primary brain tumours; screening costs for paediatric gliomas are significantly lower to due to inclusion in the genomic test directory. Also applies to : Section 1.5, page 20, fourth paragraph. Section 6.3.2, page 146, first paragraph.			
Section 5.2.8.7, page 130, first paragraph: It is stated that "The ERG also considers the scenarios presented by the company in which testing costs are shared across 2 and 4 NTRK targeting therapies as the base-case analysis should represent the incremental costs of implementing entrectinib in the NHS, not the implementation of a range of hypothetical agents that may or not be available in the future." The sentence does not make sense as written, but appears to reject scenarios whereby testing costs are shared across multiple NTRK inhibitors. These scenarios	Correct sentence, and amend paragraph to acknowledge that scenarios have been discussed and should be taken into consideration.	Contradiction of prior discussion regarding valid scenarios.	Text amended for sense (Section 5.2.8.7, page 130). The ERG, however, stand by the point made. Further, while this issue was discussed at the checkpoint meeting, the ERG and NICE advised against this scenario.

have been discussed with NICE, the ERG and NHS England, notably at the Checkpoint teleconference on 10th April 2019 where it was suggested that these scenarios should in fact be considered, especially in context of the concurrent larotrectinib appraisal. In addition, since a scenario is included whereby testing costs of excluded altogether to account for a future situation, it is logically inconsistent to disregard test cost-sharing scenarios.			
Section 5.2.8.7, page 130, third paragraph: the company-assumed cost for NGS testing is not marked as CIC	Mark cost as CIC	NGS cost is an internal assumption based on specialist advice from genomic testing hubs	NGS cost now marked as CIC (Section 5.8.2.7, page 130).
Section 6.3.2, page 146, paragraph 3: "It is assumed WGS will remove 89% of NTRK fusion negative samples, reducing the requirement for NGS confirmatory testing to 11% of the NNS population". This is a highly questionable assumption, since WGS is likely to be significantly more robust than IHC in identifying NTRK fusions. The true figure for positive WGS tests requiring confirmation with another type of NGS test (if it is needed at all) is likely to be considerably lower.	Acknowledge the severe limitations of this assumption and that the true figure for which confirmatory NGS is required is likely to be significantly smaller.	Further transparency needed to address weakness of assumption used for modelling purposes.	Not a factual error. The ERG accepts there are significant uncertainties regarding the specificity of WGS, but considered this a reasonable assumption given the absence of any evidence.

Issue 8 Inaccuracy regarding purported discrepancy in entrectinib positioning

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 3.3, page 49, first paragraph: It is stated that "The ERG found discrepancies between this source and the proposed positioning presented in CS table 6 for three tumour types, including MASC, soft tissue sarcoma, pancreatic cancer, which are positioned as first line in CS table 6, and second or subsequent line in the company appendix table." This is incorrect - the only discrepancy is the inclusion of trabectedin as a 2nd line therapy for sarcoma; positioning for doxorubicin and MASC and pancreatic cancer are consistent between the tables.	Amend sentence to "The ERG found a discrepancy between this source and the proposed positioning presented in CS table 6 for soft tissue sarcoma; the appendix table includes a second line comparator (trabectedin) while CS table 5 positions entrectinib as first line for soft tissue sarcoma."	Factual inaccuracy.	Text amended as proposed (Section 3.3, page 49).

Issue 9 Inaccuracies relating to rationales for exclusion of patient subgroups from the base case analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 4.2.1, page 54, third paragraph: It is stated that "The company's exclusion of patients with primary CNS tumours from the integrated efficacy dataset due to poor response (1 of 5 included patients responded) suggests that this assumption may not hold across all tumour types given	Remove sentence "~~The company's exclusion of patients with primary CNS tumours from the integrated efficacy dataset due to poor response (1 of 5 included patients responded) suggests that this assumption may not hold across all tumour types given larger samples~~."	Factual inaccuracy.	Text removed as proposed (Section 4.2.1, page 54).

larger samples." This is incorrect; patients with primary CNS tumours were excluded due to the limited follow-up at the time of the primary analysis and because of the different method used to measure response (RANO versus RECIST in systemic solid tumours), not due to any perceived differential response versus other tumour types.			
Section 5.2.1, page 96, third paragraph: It is stated that "The ERG is concerned that the company is seeking a recommendation in patients with primary CNS tumours and paediatric patients, yet despite data being available for these patient groups, the company has decided to omit information from the base case provided in the CS due to differing response measurements when in fact response outcomes are not used in the economic model." This statement is inaccurate as the reason for exclusion of the paediatric subgroup was due to data being unavailable at the time the submission was made. It was also agreed with the ERG at the Checkpoint teleconference on 10th April that incorporation of paediatric data may be uninformative due to the lack of a	Remove sentence: "~~The ERG is concerned that the company is seeking a recommendation in patients with primary CNS tumours and paediatric patients, yet despite data being available for these patient groups, the company has decided to omit information from the base case provided in the CS due to differing response measurements when in fact response outcomes are not used in the economic model.~~" Or amend to: "The company is seeking a recommendation in patients with primary CNS tumours and paediatric patients but due to data availability at the time of submission was not able to include these patients in the base case analysis. The company therefore provided a scenario analysis including these patients at the earliest opportunity."	Factual inaccuracy.	ERG have amended sentence to clarify that exclusion criteria for primary CNS patients for the original CCOD was due to different methods used to measure response. The ERG acknowledges that in the most recent CCOD, primary CNS and paediatric patients were excluded due to limited follow-up time (Section 5.2.1, page 96).

counter-factual. The issue of differing response measures for primary CNS tumours was raised with regard to interpretation of response data which is directly linked to measurement of progression, and so combining PFS for the two response measures also presents problems. Roche appreciates the concern regarding the lack of the paediatric and primary CNS data in the base case, in context of the scope population, and provided the additional scenario in response as soon as was feasible.			

Issue 10 Error regarding summary of impact of reducing subsequent therapy treatment duration on entrectinib costs

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Section 5.2.4, page 100, fifth paragraph: It is stated that "The impact of this assumption is to increase costs in the entrectinib arm of the model and therefore to substantially decrease the ICER." This appears to be incorrect as the impact of reducing the duration of subsequent therapy is to reduce, not increase, costs in the entrectinib arm, and also decrease the ICER.	Amend sentence to: "The impact of this assumption is to decrease costs in the entrectinib arm of the model and therefore to substantially decrease the ICER."	Factual inaccuracy.	Text amended as proposed (Section 5.2.4, page 100).

Issue 11 Incorrect assertion that company number needed to screen (NNS) is much lower than ERG NNS

Description of problem	Description of proposed amendment	Justification for amendment	ERG's response
Appendix B, page 173, second paragraph: It is stated that "Due to discrepancies in the recorded prevalence of NTRK fusions, the NNS reported in the CS are, in general, much lower than the NNS calculated by the ERG." This is incorrect as, in eight tumour types the company NNS is either similar or higher than the ERG's estimate, and in general where the company NNS is lower, it is not substantially so.	Remove sentence: "~~Due to discrepancies in the recorded prevalence of NTRK fusions, the NNS reported in the CS are, in general, much lower than the NNS calculated by the ERG.~~"	Factual inaccuracy	Sentence amended to only acknowledge differences in NNS (Appendix B, page 173).

Evidence Review Group's Report
Entrectinib for treating NTRK fusion-positive solid tumours (addendum)

Probabilistic results of the ERG base-case scenario

.

1 Probabilistic results of the ERG alternative base-case scenario

This addendum presents the probabilistic results of the ERG alternative base-case analysis, as presented in Section 6.4 of the main ERG report.

1.1 Probabilistic results of the ERG base-case analysis

Addendum Table 1 presents the probabilistic results of the ERG base-case model.

Note, a number of comparator treatments have a confidential patient access schemes (PAS) discount, including eribulin, everolimus, nab-paclitaxel, nintedanib, trabectedin and trifluridine-tipiracil. For probabilistic results with the confidential PAS applied for these therapies, please see the accompanying confidential appendix to the addendum.

Table 1 ERG probabilistic base-case

	Total costs	**Total QALYs**	**Incremental costs**	**Incremental QALYs**	**ICER (£/QALY)**
Entrectinib	[redacted]	[redacted]	[redacted]	[redacted]	£74,809
Established management	£19,897	1.04			

NICE National Institute for Health and Care Excellence

Technical engagement response form

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **Wednesday 2 October 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Roche Products Ltd**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Introductory statement – revised base-case: While uncertainty is increased due to more limited follow-up and differing methods of response assessment, Roche is in agreement with NICE and the ERG regarding the inclusion of paediatric patients and adult patients with primary CNS tumours into the base-case population, in light of the anticipated label and proposed reimbursement population (see Issue 8). It is therefore important to note that, where new scenario analyses have been requested, these are based on the model that includes these additional patients. This model was provided in response to clarification question B2. A summary of this base-case is provided in Table 1. The new base-case ICER including these patients is plausibly cost-effective at the end-of-life threshold.

Table 1: Revised base-case ICER

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£61,228	1.61	1.03	[redacted]	[redacted]	[redacted]	£49,358
Entrectinib	[redacted]	[redacted]	[redacted]				

Questions for engagement

Issue 1: Prevalence of NTRK gene fusion	
1.Which dataset more accurately reflects the prevalence of NTRK fusion for each tumour type?	We are in agreement with the ERG's preference for using the Foundation Medicine (FMI) data set as the source of NTRK prevalence data. While this source uses a predominantly United States-based population, which presents some generalisability issues, it represents the single largest and therefore most robust study of NTRK gene fusion prevalence. The overall prevalence figure from the FMI data set of [redacted] is also backed up by recent literature reporting similar, but smaller scale studies (1-3). As requested by the NICE technical team, we have therefore conducted a scenario analysis using the ERG-preferred NTRK prevalence estimates. This results in an incremental cost-effectiveness ratio (ICER) change of -£3,417 (Table 2). This reduction is driven by an overall reduction in the number-needed-to-screen (NNS).

Table 2: Scenario analysis exploring the impact ERG's preferred NTRK prevalence/NNS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case

Established management	£49,247	1.61	1.03	[redacted]	[redacted]	[redacted]	£45,941	-£3,417
Entrectinib	[redacted]	[redacted]	[redacted]					

We are proactively proposing entry of entrectinib into the Cancer Drugs Fund (CDF), so it is important to note that this area of uncertainty is one which may be addressed by collection of data during a CDF funding period.

Issue 2: Treatment pathway and positioning

2.How will 'acceptability' in the context of this appraisal be defined in clinical practice?	Roche recognises the uncertainty associated with our anticipated label and took steps to rationalise this by applying the decision criteria and resultant comparator therapies in our initial company submission. We stand by this approach, though we understand there is a degree of subjectivity to the definition of 'acceptable' therapy. Upon further feedback from clinical experts, there are further considerations that may be taken into account. For example, the lack of a target biomarker for the therapy could dictate choice, while an ORR of less than 20% and a PFS of less than four months could indicate an 'unacceptable' outcome, though these criteria are clearly dependent on context and a number of variables such as tumour type. In recognition of this uncertainty, we have provided a scenario analysis using the NHSE and NHSI's alternative recommended positioning for entrectinib (see Issue 11, question 25).
3.For each tumour type, at what points in the respective treatment pathways will entrectinib be used in clinical practice in England?	Our position on the points at which entrectinib may be used in respective tumour treatment pathways is described in the initial company submission and response to clarification questions. However, we recognise the uncertainty around the positioning of entrectinib in clinical practice and in response to Issue 11, question 25, have therefore conducted a scenario analysis using the NHSE and NHSI's alternative recommended positioning for entrectinib. We also agree with the NICE technical team's view that this issue may only be resolved through the use of entrectinib in clinical practice, and propose proactive entry of entrectinib into the CDF with a view to resolving this and other uncertainties identified within the appraisal.

4.Is the evidence for entrectinib generalisable to clinical practice in England?	Roche's view is that the clinical evidence available for entrectinib is generalisable to clinical practice in England. Entrectinib has demonstrated efficacy and tolerability across all tumour types represented within the clinical trials, regardless of line of therapy. We note the concern from the ERG and the NICE technical team that the proportion of treatment-naive patients in the entrectinib patient population may confound comparative results in favour of entrectinib, due to line of therapy being a determinant of prognosis. However, we dispute this assumption for the following reasons: • The assumption that patients in second or later lines of therapy have a poorer prognosis in general is flawed. Patients surviving and remaining fit enough for subsequent lines of therapy may in fact have a similar or even improved prognosis when compared with an unselected treatment-naive patient. • The assumption does not apply to a targeted therapy such as entrectinib, the efficacy of which is almost exclusively governed by the presence or absence of the given sensitising mutation, in this case NTRK gene fusions. • The main determinant of prognosis for entrectinib is therefore the patient's fitness, as defined by their Eastern Cooperative Oncology Group Performance Status (ECOG PS), not their line of therapy. To further address this concern, [redacted] 1 [redacted] 2 [redacted]

	[redacted] . **Figure 1: PFS for treatment-naive vs pre-treated patients in the new base-case population** [redacted] **Figure 2: OS for treatment-naive vs pre-treated patients in the new base-case population** [redacted] However, we acknowledge the uncertainty regarding the potential patient population for entrectinib in England and have proactively proposed entry into the CDF. This will allow the make-up of the real-world population in England to be established, in addition to ascertaining the activity of entrectinib in tumour types for which there are currently no data.
Issue 3: NTRK gene fusion screening pathway	
5. What is the likely screening pathway to identify NTRK positive solid tumours?	The utilisation of the IHC pre-screen in circumstances where next generation sequencing (NGS) availability is limited is in line with the most recent ESMO guidance (4) on the identification of NTRK fusion-positive patients, and is therefore rooted in evidence-based recommendations. As acknowledged by the NICE Technical Report we have taken this pragmatic approach to testing in light of the current fragmented implementation of genomic testing within the NHS. The introduction of the IHC pre-screen is seen as an interim step to ensure patient access to entrectinib and avoid the implementation of NGS-based testing within the NHS, across all tumour types, becoming a rate-limiting step. This approach could also reduce inequity of access to testing and hence to tumour agnostic therapy. At present, funding for whole-genome sequencing (WGS) is in place for paediatric cancers and sarcomas. However, NGS-based testing implementation across the remaining tumour types is still underway and is not anticipated to be available to all mestastatic cancer patients until 2021 at the earliest. We view the wide-scale implementation of appropriate NGS-based testing capable of the sensitive and specific detection of NTRK fusion-positive tumours as early in the patient pathway as possible as the most optimal testing route. This is in alignment with the vision for genomic testing within the NHS. The IHC

	pre-screen is an interim step to support access to entrectinib in tumour types where appropriate NGS-based testing has not been fully implemented. The Clinical Lead for the Cancer Drug Fund states that the greatest benefit to patient and clinician is derived from using NGS testing as early as possible in the treatment pathway, in particular by testing prior to the administration of systemic therapy. We are in agreement with this opinion. An additional consideration is the proposed entry of entrectinib into the Cancer Drug Fund. Should this be successful, this presents a window in which to optimise the proposed screening pathway around early use of appropriate NGS testing. This again could be supported by the interim use of the IHC pre-screen in order to avoid patient drug access as a result of NGS test implementation time lag.
6. At what point in the treatment pathway for each tumour type is/will NTRK gene fusion testing carried out?	As detailed in our answer to question 5 we feel that screening for NTRK fusion-positive solid tumours is best achieved in the long term through the use of appropriate NGS technology. Due to the potential to identify multiple markers in parallel, the greatest benefits are derived from using this approach as early as possible in the patient treatment pathway across all tumour types. We are in agreement with the Clinical Lead for the CDF, who states that the greatest benefit to the patient and clinician is derived from using NGS testing as early as possible in the treatment path, in particular by testing prior to the administration of systemic therapy. An additional technical consideration should be given to the availability of biopsy tissue in certain tumour types. In many cases the material for testing is extremely limited. Minimising the need for serial testing on limited biopsy tissue is therefore beneficial. By deploying parallel marker identification via NGS on limited tissue as early as possible in the treatment pathway further maximises any benefits to be gained.
Issue 4: NTRK gene fusion testing costs	
7. Should testing costs for molecular testing that is already done in the NHS	Roche's position is that testing costs should be excluded from the appraisal of entrectinib for reasons described in response to Issue 4, question 11. However, in scenarios where testing is included, Roche's position is that costs for molecular testing for any biomarker or oncogenic driver mutation should be included in the comparator arm, whether or not it is for NTRK gene fusions. The primary reason for this is that testing for NTRK fusions requires either up-front detection or confirmation by an NGS panel, which is likely to be capable of detecting a wide range of actionable mutations (such as EGFR in lung cancer) and markers (such as PD-L1 gene expression across a range of tumour types). Excluding testing costs from the comparator

be included in the economic model, even if it is not to detect NTRK gene fusions?	arm for those tumour types where testing is currently performed for other targets therefore double-counts testing costs, as NTRK testing via NGS-based testing is likely to replace existing testing by other methods.
8. Are the costs of adding a panel to an RNA-based Next Generation Sequencing (NGS) test negligible?	Theoretically the costs associated with the addition of a new biomarker to an established and validated RNA-based NGS panel could be negligible. However, it should be noted that such costings are subject to a number of technical parameters which could have a significant impact. In addition, it should be noted that this figure would, at present, potentially be subject to lab-to-lab variation. We understand that currently RNA sequencing is not being offered routinely as part of the NHS Genomic Medicine Services implementation. As a result, any RNA sequencing within the NHS is subject to local variation in tests offered by local labs. It is therefore very difficult to provide a precise estimate of potential costs.
9. Is it appropriate to include the costs of confirmatory NGS for people who have Whole Genome Sequencing (WGS)?	Roche's position is that it is appropriate to exclude all testing costs from the cost-effectiveness analysis of entrectinib. A full summary of the reasoning behind this position is given in our answer to Issue 4, question 11. We wish to see the implementation of appropriate NGS-based testing which is capable of the sensitive and specific detection of NTRK 1/2/3 fusions across all tumour types. As discussed this can technically be approached in a number of ways. However, it should be noted that the use of NGS hybrid panels capable of the parallel detection of both DNA and RNA targets minimises the need for confirmatory follow-up testing. Additional consideration should be given to the availability of test material which may be limited in certain tumour types. Biopsy material may be insufficient or low in tumour content. In such cases, the ability to do any follow up confirmatory testing may be significantly limited. Where confirmatory testing is based on RNA sequencing it should also be noted that the performance of such testing is highly dependent on the quality of extracted RNA.
10. Should the economic model	Roche agrees with the NICE technical team's view that it is inappropriate to include testing costs for unrepresented tumour types in the absence of outcome data in these tumour types.

include testing costs for unrepresented tumour types?	
11. What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?	Roche's position is that it is appropriate to exclude all testing costs from the cost-effectiveness analysis of entrectinib for the following reasons: • Due to the anticipated tumour agnostic license, the testing for NTRK gene fusions required to identify patients appropriate for entrectinib is unlike any previous new biomarker testing in oncology. The associated costs of this testing technology, and the extreme rarity of the gene fusion results in a much higher testing cost-per-patient-identified than has been seen previously. • Applying any testing costs to entrectinib penalises the innovation of the technology; the first therapies utilising this testing methodology will be forced to bear significant testing costs, while therapies of the same class entering the market at a later time point are likely to fit into existing pathways without requiring any additional testing costs to be factored into their cost-effectiveness assessments. • As well as the concurrent appraisal of the NTRK larotrectinib, numerous other tumour agnostic therapies are in development; for NTRK gene fusions alone two further NTRK inhibitors, selitrectinib and repotrectinib, are anticipated in the near future. Tumour agnostic products targeting RET fusions and KIT mutations are also in development alongside molecules targeting other markers. It is therefore inappropriate to attempt to apply a proportion of testing costs in the face of future pan- or tumour-agnostic developments. • Testing for NTRK gene fusions is conducted by NGS-based technology either up-front or through confirmation testing post e.g. immunohistochemistry (IHC) pre-screening. NGS-based testing provides a wealth of additional genetic information, well beyond the value of identifying an NTRK gene fusion positive patient. At a minimum, this information may be used to inform a patient's prognosis, eligibility for a clinical trial, or an alternative treatment choice for the patient. The value of this information is high, and is not currently factored into the cost-effectiveness analysis for entrectinib. This therefore leads to considerable uncertainty in assessing the costs and benefits of testing for NTRK gene fusions.

- Health technology appraisal of tumour agnostic therapies is currently subject to considerable uncertainty on a number of fronts, a key one being NTRK testing, particularly given the current state of transition within the NHS regarding genomic testing in cancer. Removing the testing component from the process therefore significantly reduces the uncertainty burden.

- The NHS's drive towards pan-tumour, up-front NGS-based testing for all cancer patients is taking place regardless of the advent of tumour agnostic therapies such as entrectinib. The cost of implementing this initiative will therefore ultimately be borne by NHS budget, regardless of the emergence of new therapies that could play a role in this area.

We note the NHSE and NHSI's view of how testing should be implemented, i.e. that our proposed pragmatic interim approach of IHC pre-screen followed by NGS-based confirmation should be disregarded in favour of up-front NGS testing. We are in agreement that this is an ideal scenario for testing, as long as an appropriate NGS technology is used that detects all three NTRK gene fusions to a high degree of sensitivity and specificity. However, this view does not reflect NICE's standard perspective of the present day, i.e. the current state of affairs as of the time of this appraisal. Our approach represents a pragmatic step that could be implemented at the point of reimbursement of entrectinib, and could therefore be considered to be reflective of the present day. The NHSE and NHSI's proposal for up-front NGS testing also has a significant impact on the costs incurred by entrectinib, as demonstrated by the NICE technical team's requested scenario analyses shown in Table 3 (including additional 10% and 25% proportions). This approach further exacerbates the issue of testing cost burden.

This analysis was conducted by removing the IHC pre-screening from the model and applying the [redacted] test cost to all patients, while keeping comparator testing costs as per the company base-case. The incremental testing cost for MASC, secretory breast cancer and soft tissue carcinoma was held at 0 for this analysis.

Table 3: Scenario analyses exploring the impact of up-front NGS-based testing at varying proportions

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case
100% up-front NGS-based testing								
Established management	£61,258	1.61	1.03	[redacted]	[redacted]	[redacted]	£96,641	£47,283
Entrectinib	[redacted]	[redacted]	[redacted]					
50% up-front NGS-based testing								

Established management	£41,233	1.61	1.03	[redacted]	[redacted]	[redacted]	£66,205	£16,847
Entrectinib	[redacted]	[redacted]	[redacted]					
33% up-front NGS-based testing								
Established management	£34,588	1.61	1.03	[redacted]	[redacted]	[redacted]	£56,060	£6,702
Entrectinib	[redacted]	[redacted]	[redacted]					
25% up-front NGS-based testing								
Established management	£31,221	1.61	1.03	[redacted]	[redacted]	[redacted]	£50,988	£1,630
Entrectinib	[redacted]	[redacted]	[redacted]					
10% up-front NGS-based testing								
Established management	25,213	1.61	1.03	[redacted]	[redacted]	[redacted]	£41,857	-£7,501
Entrectinib	[redacted]	[redacted]	[redacted]					
Testing costs excluded								
Established management	£21,208	1.61	1.03	[redacted]	[redacted]	[redacted]	£35,770	-£13,588
Entrectinib	[redacted]	[redacted]	[redacted]					

Issue 5: Identification of NTRK gene fusions - diagnostic accuracy	
12. What is the expected diagnostic accuracy of Next Generation Sequencing	As acknowledged within our submission the NHS is currently undertaking a significant implementation of genomic testing at scale. The Genomic Laboratory Hubs are still in the process of becoming operational and as a result our experience has been that there is still significant variation in the NGS-based offering between sites. As such, there is no single representative NGS-based offering demonstrating a given diagnostic accuracy. In addition, it should be noted that next-generation sequencing is a platform for multiple different assay types ranging from DNA panels, through whole exome into amplicon sequencing and hybrid capture. Each assay type has numerous technical factors which have the potential to

(NGS) testing?	impact on sensitivity and specificity. It is therefore difficult to present a single set of figures which are representative of the wide variation in potential next-generation based assays. Based on recent peer-reviewed publications aiming to examine the advantages and limitations of various potential technologies for the detection of NTRK fusions, Marchio et al (4) states that the sensitivity of DNA based sequencing is "moderate" and specificity is "high". While RNA based NGS is stated as having "high" sensitivity and specificity. It should however be noted that RNA sequencing performance is highly dependent on RNA quality. Solomon et al (3) quantified assay performance relative to RNA-based sequencing in a large cohort of NTRK fusion-positive tumours. Here, the authors show an overall sensitivity and specificity of 81.1% and 99.9% respectively for a DNA-based sequencing assay when compared to RNA-based sequencing. It should however be noted that the DNA assay quoted (MSK-IMPACT), is a hybridisation capture assay designed to target all exons and selected introns across NTRK1-3. As such, this is an example of one such NGS approach and cannot be taken as representative of the performance of all NGS assays. By contrast the Thermofisher Oncomine Focus assay has a stated sensitivity and specificity of 100% for the detection of fusions (ThermoFisher Scientific). In summary, this demonstrates the potential variation in potential diagnostic accuracy across NGS-based assays and is an area of potential further evidence generation, particularly in the specific context of NTRK fusion detection in clinically derived samples.
13. Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types?	The minimisation of any potential false positive or false negative results is a function of the sensitivity and specificity of the technology employed in diagnosis. As discussed in our answer to question 12 it is difficult to provide a precise diagnostic accuracy for any single NGS-based approach due to the variation in methods being employed within the Genomic Laboratory Hubs. However recent literature is in agreement that the sensitivity and specificity of NGS-based approaches is generally considered to be high. We feel that the utilisation of a RNA/DNA hybrid panel NGS approach has the potential to maximise sensitivity and specificity, therefore minimising the risk of false negatives/positives. As presented in our submission the utilisation of an IHC pre-screen prior to NGS confirmation would also minimise false-negatives and is alignment with the most recent ESMO guidance.

14. Is it appropriate to limit testing to avoid false positive results and the associated costs?	In light of the potential tumour agnostic indication being sought for entrectinib, it would not be appropriate to limit testing and exclude certain tumour types. As described in our answer to question 13 we feel that the implementation of an appropriate NGS based technology has the potential to achieve high sensitivity and specificity.
Issue 6: Distribution of tumour types	
15. Are the ERG's estimates of the distribution of tumour types reflective of what would be seen in clinical practice in England? If no, what is the likely distribution of NTRK positive	Roche acknowledges that distribution of tumour types is an area of uncertainty in this appraisal. The entrectinib trial populations were recruited though an extensive testing programme in more than 10,000 patients taking place across 150 sites in 15 countries, including the UK. It is therefore reasonable to expect that the tumour proportions seen in this population may reflect that seen in clinical practice. Over-representation of rare tumour types such as MASC and secretory breast cancer can be explained as these tumour types are characterised by a very high prevalence of NTRK fusions (1-3). While we recognise that the ERG's estimated distribution may be plausible, the likely distribution of NTRK gene fusion-positive tumour types in England may only be definitively answered once comprehensive NGS-based testing is implemented in an unbiased fashion across all cancer patients. For this reason, among others, Roche proposes proactive entry of entrectinib into the CDF; during a CDF period, this uncertainty may be addressed through a data collection plan. The ERG's scenario analysis in which they re-weighted comparator costs and outcomes according to their estimated tumour distribution in England is inappropriate since entrectinib costs and outcomes were not also adjusted due to unavailability of patient data to the ERG. In order to address this, and in response to the NICE technical team's request, we re-weighted the entrectinib data according to the number of patients with each tumour type calculated by the ERG (see Table 4). [redacted] 3 [redacted] 4 [redacted]. The ERG also state that the increase to the ICER was due to a reduction in the proportion of sarcoma patients, a tumour type associated with high comparator costs; it is important to note that the reason

tumour types in clinical practice?

for these higher costs, is the existing WGS conducted for sarcoma as specified in the Genomic Testing Directory. The higher ICER is therefore artificially inflated due to high testing costs.

Table 4: Tumour type re-weighting according to ERG calculations

Tumour Type	Observed no.	Observed %	ERG %	Tumour Weight	Weighted no.
Sarcoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
NSCLC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
MASC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Breast	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Thyroid	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CRC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pancreatic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neuroendocrine tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gynaecological	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CNS Primary	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Skin Cancer	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 3: Re-weighted PFS data

[redacted]

	Figure 4: Re-weighted OS data
Issue 7: Unrepresented NTRK gene fusion positive solid tumour types	
16. Are the results of the entrectinib studies that includes 13 different NTRK positive tumours sites generalisable to all NTRK-fusion positive tumour types?	Roche acknowledges the uncertainty associated with the existing entrectinib clinical data in terms of the potentially unrepresented tumour types. However, the existing data for entrectinib demonstrates compelling efficacy across tumour types, as demonstrated by the waterfall plots provided in the initial submission and responses to clarification questions. In combination with the fact that NTRK fusions are thought to play a similar role across tumour types (5), this leads clinical experts to believe that there is compelling reason for entrectinib to have notable activity and response rates across tumour types. In addition, targeted therapies aimed at actionable mutations such as NTRK gene fusions typically demonstrate high response rates; for example, EGFR inhibitors in lung cancer have shown response rates of 55–75% (6) , while a pooled analysis of ALK inhibitors has shown an aggregated response rate of 65% (7). It is also important to note that entrectinib has a tumour agnostic indication in the United States and Japan, and is anticipated to have a similar indication in Europe; this implies that regulatory bodies have judged the benefit:risk profile in this population to be acceptable, based on the available data and the scientific rationale behind the therapeutic approach. We note the NICE technical team's request to investigate tumour type heterogeneity in time-to-event outcomes using the ERG's BHM framework. However, our view is aligned with the ERG in that the current data is not robust enough to provide any meaningful estimates using this approach. We have proactively proposed entry of entrectinib into the CDF alongside a comprehensive data collection plan, one of the aims of which will be to collect data in patients with tumour types that are not currently represented in the clinical data available to date. Our view is that the only appropriate way to resolve this uncertainty is through this approach.
Issue 8: Primary CNS tumours and paediatric tumours	
17. Should people with primary CNS tumours and	As described in our introductory statement to our responses, Roche agrees that it is appropriate to include people with primary CNS tumours and paediatric tumours in the analysis, since this is in line with the anticipated marketing authorisation and the proposed reimbursement population. As stated in our response to clarification questions, these patients were not originally included in the analysis population in the initial company submission due to the very limited follow-up data available

paediatric tumours be included in the analysis?	at that time. There are also concerns with combining data sets that use differing methods to assess response, for example investigator assessed versus independent review committee, and Response Assessment in Neuro-Oncology criteria (RANO) versus Response Evaluation Criteria In Solid Tumours (RECIST v1.1), which remains an issue.
Issue 9: Heterogeneity of response across different solid tumour types	
18. Is a uniform level of response to entrectinib across the different tumour types with a NTRK fusion reasonable?	As discussed in the response to Issue 7, question 16, the currently available evidence for entrectinib and the literature regarding the role of NTRK gene fusions across tumour types (5) suggests that entrectinib is active across the studied range of tumour types, and will be active in as yet unknown tumour types. The assumption of a uniform level of response across a heterogeneous tumour type population is therefore reasonable. However, the numbers of patients with each given tumour type are too small for any definitive conclusion regarding a consistent response between tumour types, so we acknowledge the uncertainty in this assumption. We have therefore proactively proposed entry of entrectinib into the CDF alongside a comprehensive data collection plan. During data collection, we propose to collect further data in tumour types that are currently represented in the available data, as well as new data in patients with tumour types that are not currently represented. This data collection will help to establish the variability or otherwise of response rates across tumour types.
19. Is the Bayesian Heirarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types	Roche's perspective is that the existing subgroup data for entrectinib is not robust enough at the specific tumour type level for reliable modelling to assess tumour or response heterogeneity. Even the largest subgroup, soft tissue sarcoma (13 patients), actually represents a potentially very heterogeneous population in itself with a range of underlying histologies and variable prognoses. As discussed in our response to Issue 9, question 18, we believe the most appropriate way to address the issue of potential response heterogeneity is through continued data collection during a potential interim CDF funding period.

for this appraisal?	
20. Would it be appropriate to apply the Bayesian Heirarchical Modelling (BHM) framework to explore the heterogeneity in the time to event outcomes?	Roche agrees with the ERG that the currently available time-to-event data is not yet robust enough to use BHM to explore heterogeneity in these outcomes across tumour types. In addition, as discussed in our response to Issue 9, question 19, the individual tumour type subgroups are too small to model outcomes using the BHM framework. As requested by the NICE technical team, we have provided descriptive PFS and OS KM curves split according to tumour type (Figure 5 and Figure 6). These demonstrate that any attempt to model time-to-event outcomes using a BHM framework is unlikely to result in useful information. Please note that for the purposes of this descriptive data, paediatric data have been combined with adult in the sarcoma and primary CNS tumour groups. **Figure 5: PFS split by tumour type** [redacted] **Figure 6: OS split by tumour type** [redacted]
21. Are the response rates in tumour sites not represented in the trial data suitable for consideration?	As discussed in previous responses to questions under Issues 7 and 9, Roche acknowledges the uncertainty regarding activity of entrectinib in unknown tumour types. However, the existing clinical evidence and rationale regarding the consistent role of NTRK gene fusions across tumour types provides reason to believe that entrectinib will be active in them. The ERG's BHM-predicted response rates in unknown tumour types of [redacted]. While we acknowledge that it is plausible that response rates are likely to fall within this range, the range is so broad that in our view, its usefulness for decision making is limited. In addition, as discussed in our response to Issue 7, question 16, entrectinib has shown compelling responses across a wide range of tumour types. Furthermore, targeted therapies aimed at actionable oncogene driver mutations such as NTRK gene fusions typically demonstrate high response rates; for example, EGFR inhibitors in lung cancer have shown response rates of 55–75% (6), while a pooled analysis of ALK inhibitors has shown an aggregated response rate of 65% (7). We consider that the most appropriate method to address this uncertainty is through continued data collection.
Issue 10: Constructing a comparator arm	

22. Is the company's comparator arm suitable for decision making?	Roche believes the comparator arm modelled in our initial submission, and the updated arm provided in response to clarification questions to account for paediatric and primary CNS tumours, provides a reasonable and pragmatic basis on which to assess the comparative effectiveness of entrectinib. The decision criteria used to define comparators were discussed with clinical experts, and individual comparators alongside their outcomes were also subsequently validated with clinical experts in many cases. However, we acknowledge that there are some key uncertainties in the comparison, in particular the potential impact of NTRK gene fusions on prognosis, and the relatively high proportion of patients with brain metastases in the entrectinib clinical trial population. As outlined in our initial submission, we believe these issues are likely to worsen the outcomes of a comparator population when compared to patients with no oncogenic driver mutations or brain metastases. Roche has recently conducted a literature review to identify evidence of the prognostic implications of NTRK gene fusions, the results of which are summarised in Issue 12, question 27. A scenario analysis using an amended comparator arm with changes recommended by the NHSE and NHSI has been conducted; please refer to Issue 11, question 24 for the outcomes of this analysis. [redacted] 7 [redacted] 5

[redacted]

Figure 7: Intra-patient analysis TTNT curve alongside modelled comparator and entrectinib PFS curves

[redacted]

Table 5: Therapies administered immediately prior to entrectinib

Tumour type	Prior therapies
Breast	[redacted]
Cholangiocarcinoma	[redacted]
CRC	[redacted]
Gynaecological	[redacted]
NETs	[redacted]
NSCLC	[redacted]
Pancreatic	[redacted]
MASC	[redacted]
Sarcoma	[redacted]
Thyroid	[redacted]

While we believe the modelled comparator arm is suitable to assist decision making, and is supported by the intra-patient analysis, we also acknowledge that comparative analysis is an uncertainty in the assessment of entrectinib, and is a key

	reason why we have proposed proactive entry into the CDF. During the CDF period, Roche intends to propose a substantial range of data collection activities aimed at addressing this and other uncertainties.
23. Is it appropriate to use non-responders as a proxy for patients not having an active treatment to generate a comparator arm for this appraisal?	Roche believes the use of non-responding patients as a proxy comparator arm is a deeply flawed approach and is inappropriate for decision-making for the following reasons, some of which are acknowledged by the ERG and the NICE technical team: • The response data is derived from a pool of 66 patients in total, comprising [redacted]. The sample sizes in each arm are therefore too small to provide a meaningful or robust comparison. • Response has been measured differently for groups of patients within responder analysis; primary CNS tumours were measured by Response Assessment in Neuro-Oncology Criteria (RANO) while other solid tumours were measured by Response Evaluation Criteria In Solid Tumours (RECIST v1.1). • Non-responding patients may still benefit from entrectinib while by halting or slowing tumour growth while not meeting the RECIST v1.1/RANO threshold for response; for example, patients may experience stable disease or a reduction in tumour size of under 30%. Clinical expert feedback suggests that it may be more appropriate to use patients with a best response of progressive disease as a comparator arm for this reason. • Results may be biased by a large proportion of patients who have a reduction in tumour size while not meeting the RECIST v1.1/RANO definition of response. Approximately [redacted] patients in the data set may fit these criteria. This risks underestimating the magnitude of benefit of entrectinib through an over-performing comparator arm. • The heterogeneity of the patient population means that key prognostic factors such as tumour type, presence of brain metastases and ECOG PS are unbalanced, resulting in a highly unreliable comparison. • There is no adjustment for covariates such as those described above, which in such a heterogeneous population is a fundamental issue. • The "usual method" used by the ERG is the least robust of the methods by which responders may be compared with non-responders due to inherent bias, and its use is recommended against in the literature (8). • Overall survival data for entrectinib is still immature leading to uncertainty in longer term outcomes.

	• For primary CNS and paediatric tumours, there is even more limited follow-up, leading to greater uncertainty in this analysis. • The intra-patient analysis conducted by Roche is supportive of the modelled weighted naive comparator data, lending further weight to its validity for decision making. Response-based analyses are rarely used in HTAs conducted by NICE, with manufacturers, ERGs and NICE generally preferring comparison to historical control where no direct comparative data are available, as per NICE TSD 17. In the rare instances that these analyses have been used over the last 10 years, it has generally been supportive to a primary comparative analysis method (NICE TA347, TA380, TA421, and TA536). Landmark response-based analyses have been the primary comparative method in only two previous appraisals (TA486 and TA530), and in both cases the ERG was heavily critical of the approach, leading to the Committees involved to question their relevance for decision making. In conclusion, we believe that the weighted naive comparison supported by the outcome of the intra-patient analysis is the more appropriate comparative methodology.
Issue 11: Comparator treatments	
24. Is the company's modelled comparator population representative of the eligible population in England?	Roche's comparator population was designed according to the likely therapies that patients would receive in the absence of entrectinib in accordance with the anticipated licence. The initial scope and decision criteria for comparators were discussed with clinical experts. Individual comparators were derived from NICE's own published clinical guidelines and technology appraisals, thus ensuring that comparators consisted only of those routinely available on the NHS in England. We acknowledge the uncertainty regarding appropriate comparators. Our approach, whereby comparators from multiple lines were used in some tumour types (for example 2nd and 3rd line in colorectal cancer), was designed to address this by acknowledging that there could be multiple points in a treatment pathway where entrectinib could be used, depending on an individual clinician's interpretation of "acceptable" therapy. However, as requested by the NICE technical team we have performed a scenario analysis using the NHSE and NHSI's preferred comparator positioning at single and in some cases later lines of therapy. This scenario analysis was conducted by removing the costs and outcomes of comparators that the NHSE and NHSI deemed inappropriate; these were capecitabine in second-line breast cancer, FOLFIRI and irinotecan in second-line colorectal cancer, and everolimus in neuroendocrine tumours (NETs). Where there was uncertainty in suggested positioning, no changes were made. These adjustments resulted in change to the ICER of -£1,464 (see Table 6). However, it should be noted that the best supportive care outcomes used for

NETs are heavily confounded by cross-over to active therapy, as is acknowledged in NICE TA449 and TA539. Arbitrary reduction of overall survival for BSC in NETs to 24 months reduced the ICER by a further £1,903.

Table 6: Scenario analysis exploring NHSE and NHSI's preferred comparators

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case
Established management	£59,119	1.57	1.01	[redacted]	[redacted]	[redacted]	£47,894	-£1,464
Entrectinib	[redacted]	[redacted]	[redacted]					

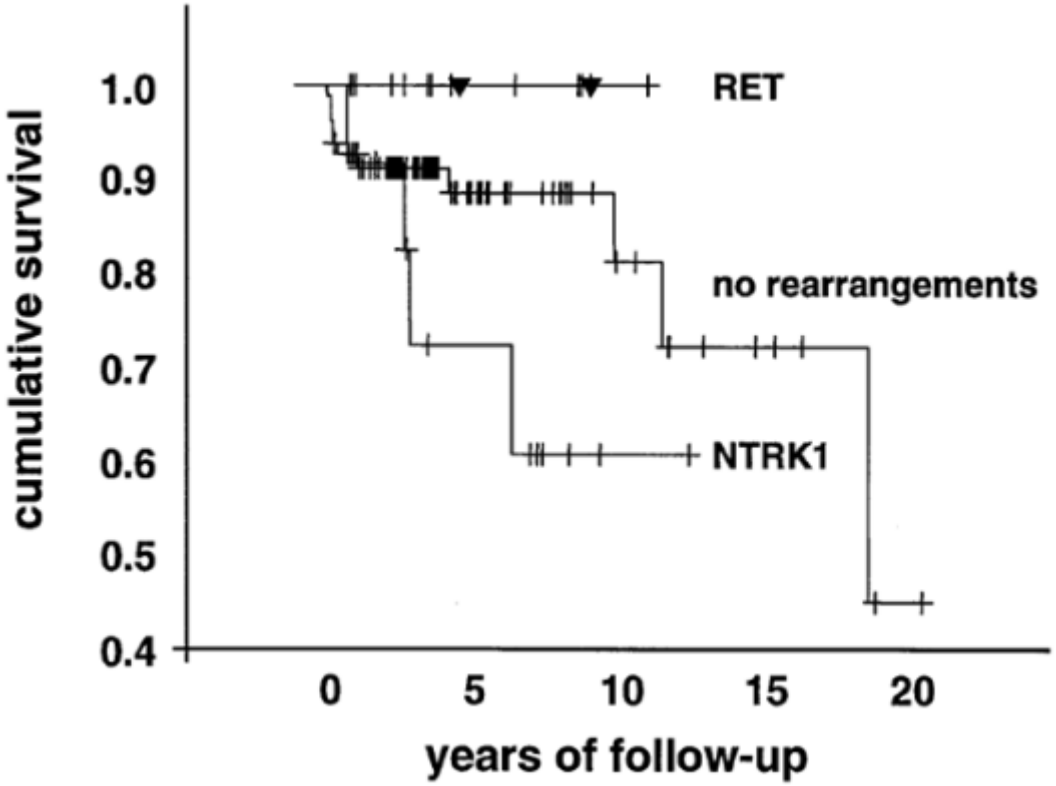
Regarding the proportion of treatment-naive patients in the entrectinib trial population, this issue is addressed in our response to Issue 2, question 4. As stated in that response, this issue is not anticipated to have a significant impact on the overall efficacy of entrectinib.

In light of the uncertainty regarding comparators, Roche proposes proactive entry of entrectinib into the CDF. In addition to the routine data collection performed via systemic anti-cancer therapy (SACT) during a CDF period, Roche plans to perform a range of data collection activities which may help to address this uncertainty.

25. Should the comparator data be matched to the position that entrectinib is given in the efficacy evaluable dataset?

Roche's perspective is that it is not appropriate to match comparator data to the position that entrectinib was given in the trial populations. The primary reason for this is that Roche agrees with the ERG's and the NICE technical team's view that the trial populations do not fully match the anticipated licence in terms of entrectinib's anticipated positioning in treatment pathways. As demonstrated in our response to Issue 2 question 4, the line of therapy at which entrectinib is administered is not a prognostic factor, meaning the entrectinib outcome data can be considered representative of outcomes seen at lines of therapy matched to the modelled comparators.

Issue 12: Prognostic factors	
26. Is it appropriate to adjust the comparator data to account for worsened prognosis of CNS metastases?	Roche's position is that it is appropriate to adjust comparator data to account for the worse prognosis associated with the presence of CNS metastases, given the relatively high prevalence of CNS metastases in the original base-case population (20.4%). This proportion is significantly higher than the highest proportion reported in comparator TAs (14% in TA426). In addition, although the inclusion/exclusion and reporting of CNS metastases prevalence in comparator TAs outside of lung cancer are not reported, in some cases they were excluded altogether. It is therefore reasonable to assume that the efficacy of entrectinib is more greatly impacted by the higher proportion of CNS metastases than the weighted comparator arm. In our initial submission, we provided a scenario analysis in which the comparator arm was adjusted for the possible impact of brain metastases. This was done by applying a proportionately worse PFS, OS and utility values based on the literature (9, 10) to an equivalent number of equivalent tumour type patients in the weighted comparator arm. Applying this adjustment to the new base-case population reduces the ICER by £4,337.
27. Is it appropriate to adjust the comparator data to adjust for NTRK fusions? If yes, are there any additional evidence sources which could be used to inform whether presence of an NTRK	Roche acknowledges the uncertainty regarding the prognostic implications of the presence of NTRK gene fusion and consequently undertook a recent literature review in order to identify any new literature on this topic, in an attempt to address this. The search was conducted on 25th June 2019, and the full report is provided as an appendix to this response. In summary, the search did not identify any further literature beyond that previously cited in our initial submission and identified in the ERG's subsequent report. However, the majority of the limited evidence available to date are suggestive of a worse prognosis for NTRK gene fusion positive tumours (11-13). The studies by Musholt *et al* and Prasad *et al* in particular, which specifically look at NTRK gene fusions in thyroid cancer, suggest a trend to worse OS (see Figure 8). While it is possible that the prognostic effect of an NTRK gene fusion may vary between tumour types, we believe it is appropriate to consider a scenario where the comparator population prognosis is worsened. In particular, given two of the three studies show a worse prognosis of NTRK fusions in thyroid cancer, this should be taken into account for the comparator arm and end-of-life criteria (see response to Issue 21).

fusion is prognostic?	**Figure 8: Kaplan-Meier estimates obtained with the use of tumour-related death calculations of patients with papillary thyroid cancer harbouring RET (n = 15 patients) or NTRK1-derived chimeric oncogenes (n = 13 patients), respectively, matched up to patients with papillary thyroid cancer without TK receptor rearrangements (n = 89 patients) (12)**  We anticipate that this uncertainty may be addressed during a potential CDF data collection period, and we have data collection proposals aimed at addressing this question.
Issue 13: Subsequent therapies	

28. What percentage of people, who receive entrectinib, would be expected to receive subsequent therapy following disease progression?	Roche's base-case assumption is a proportion of [redacted] based on data from the STARTRK-2 trial. While there is uncertainty associated with this figure, we believe it is the most reasonable basis to derive rates of subsequent therapy at this time. In clinical practice this figure is likely to be governed by a number of factors such as where in the pathway entrectinib is used, the availability of a subsequent therapy and a patient's fitness, and willingness, for further therapy. According to the anticipated licence [redacted]. In combination with a potential drop in fitness upon progression on entrectinib, and hence eligibility for further treatment, this means it is unlikely that a high proportion of patients will receive active therapy after entrectinib. Where therapy is administered, it is likely to be with one of the comparator standard of care options identified in this appraisal, which are essentially displaced by entrectinib. If these points are taken into account, Roche's base-case assumption is reasonable and potentially higher than is likely in clinical practice. We note the NICE technical team's request to adjust the entrectinib trial data to remove the effect of subsequent therapies from the entrectinib arm where appropriate. Unfortunately, this is not possible as standard methods such as rank-preserving structural failure time (RPSFT) and inverse probability-of-censoring weights (IPCW) require either a randomised trial or large patient numbers, respectively. In addition, censoring at the point of subsequent therapy is known to introduce significant bias. We also question the appropriateness of adjusting for subsequent therapy since, while the subsequent therapies used after entrectinib may not match what would be expected in clinical practice in all cases, there is likely to be a subsequent therapy effect in reality. In addition, the impact of subsequent therapies after entrectinib is likely to be small given the limited follow-up available.
29. What percentage of people, who receive established management, would be expected to receive subsequent therapy	Roche's base-case assumption is a proportion of 0%, because of the [redacted]. However, we recognise that in reality the figure is likely to be higher than 0%, since it is possible that some patients will be fit enough for the limited options still available to them, or eligible for a clinical trial. The assumption of 0% is therefore conservative in terms of the cost-effectiveness of entrectinib.

following disease progression?	
30. Which subsequent therapies would be used and in what proportions? Please answer for following entrectinib and following established management.	In clinical practice, Roche anticipates that certain active comparators identified in the model would be used after entectinib, since entrectinib may displace them in the patient pathway. For example, docetaxel +/- nintedanib may be used in lung cancer patients progressing on entrectinib, provided they remain fit and willing enough for further therapy. In the model, a subsequent therapy cost was applied via the application of a mean monthly comparator drug cost until death. This approach assumes an equivalent distribution of comparator therapies and subsequent therapies. We are unable to provide comment at this time regarding the proportions in which different subsequent therapies are likely to be used either after entrectinib or after established management. It should be noted that, given the rarity of NTRK gene fusion-positive patients, the potential role of clinical trials in subsequent therapy (for example, for a new NTRK inhibitor) may be more significant than for it is for other conditions, potentially reducing the cost impact of subsequent therapies.
31. How long would these people be expected to be treated with subsequent therapy? Please answer for following entrectinib and following established	Roche acknowledges that our initial assumption of subsequent therapy until death is considerably less plausible than the shorter three and six month durations discussed by the ERG. We agree that six months' duration is a plausible alternative, though even this may overly pessimistic given the typical 4–6 cycle duration of many standard chemotherapies; for example, docetaxel therapy may only be administered for 2–4 months. This duration is likely to apply to therapy subsequent to both entrectinib and established management. Applying a 6-month subsequent therapy duration reduces the base-case ICER by £9,468, from £49,358 to £39,890.

management .	
Issue 14: Model structure	
32. What is the most appropriate model structure for this appraisal?	Roche's position is that the partitioned survival analysis remains the most appropriate model structure for decision making. The limitations of the response-based model are numerous and are listed in our response to Issue 10, question 23. In addition, we have concerns regarding the use of the BHM framework to explore tumour heterogeneity, particularly with regard to time-to-event outcomes, and do not believe this is appropriate for decision making. These concerns are described in responses to questions under issues 7 and 9. Our view is that continued recruitment of new patients and follow-up of existing patients, alongside other data collection proposals, will provide further confidence in the base-case model by reducing the uncertainties identified during the appraisal process.
Issue 15: Extrapolation of overall and progression-free survival	
33. Is the exponential or Weibull distribution most appropriate for extrapolating overall and progression-free survival?	Roche's base-case uses the exponential distribution to extrapolate PFS and OS for the following reasons: • The exponential distribution was consistently the best statistical fit for both PFS and OS by both Akaike information criterion (AIC) and Bayesian information criterion (BIC) across two separate data cuts. • The exponential distribution was deemed to be one of the most clinically plausible based on consultation with clinical experts. In addition, the exponential distribution provided the most stable predictions of OS across both data cuts (see Table 7), with mean OS only increasing by 2.13 months (6%), and mean PPS only increasing by 0.98 months (5%). By contrast, for the Gompertz distribution these values increased by 6.3 (29.6%) and 4.39 months (66.4%), respectively. Of note, it is apparent that as follow-up increases, all distributions predict more favourable survival outcomes, with the most pessimistic (gamma and Gompertz) showing the most marked improvement. Please note that these analyses were done using the original integrated analysis population, i.e. excluding primary CNS and paediatric tumours, for consistency with the original base-case.

Table 7: Stability of OS predictions across two data cuts by distribution

Distribution	31st May 2018 data cut		[redacted]	
	Mean OS (m)	Mean PPS (m)	Mean OS (m)	Mean PPS (m)
Exponential	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]

We also question the appropriateness of only changing the distribution for the entrectinib arm while not exploring the impact of changes to the comparator arm. The ERG acknowledge in their report that comparator OS may be overestimated and that consequently the ICER may also be overestimated. We have attempted to address this by exploring a proxy method for applying a Weibull distribution to the comparator. This was done by assuming a shared gamma parameter between entrectinib and the comparator while individual comparator Lambda parameters were estimated based on the reported median survival. The results of this analysis demonstrate that applying the proxy Weibull method to the comparator as well as entrectinib (Table 8) significantly reduces the impact of an entrectinib Weibull distribution on the ICER from £13,604 to £5,886.

Table 8: Scenario analysis exploring impact of proxy Weibull method on comparator arm

	Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case
Entrectinib Weibull distribution	Established management	£61,228	1.61	1.03	[redacted]	[redacted]	[redacted]	£62,962	£13,604
	Entrectinib	[redacted]	[redacted]	[redacted]					
Entrectinib and comparator Weibull distribution	Established management	£60,485	1.42	0.92	[redacted]	[redacted]	[redacted]	£55,244	£5,886
	Entrectinib	[redacted]	[redacted]	[redacted]					

	To conclude, our position is that the exponential distribution is most appropriate for extrapolating overall survival. However, we acknowledge the uncertainty in predicting long-term OS outcome with entrectinib, and have proactively proposed that entrectinib be considered for the CDF. During an interim CDF period, this uncertainty is anticipated to be resolved with continued trial follow-up.
Issue 16: Drug wastage and source of treatment costs	
34. Is it appropriate to assume no drug wastage of entrectinib?	Roche acknowledges that there may be a small amount of drug wastage in clinical practice. Implementing drug wastage in the model increases the ICER by £2,745 (Table 9). However, in the context of this issue it is important to note that, as the ERG acknowledges, we have taken the conservative assumption of 100% dosing intensity for entrectinib, when the mean observed dose in the original 31st May 2018 analysis was [redacted]. Applying this dose intensity to the new base-case reduces the ICER by £4,461. Consequently, it could be considered that the drug wastage is negated by reduced dosing intensity.

Table 9: Scenario analysis exploring impact of drug wastage

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case
Established management	£61,228	1.61	1.03	[redacted]	[redacted]	[redacted]	£52,103	£2,745
Entrectinib	[redacted]	[redacted]	[redacted]					

35. Should drug acquisition costs for the comparator therapies be sourced from the British National Formulary (BNF) or the electronic	Roche acknowledges that the drugs and pharmaceutical electronic market information tool (eMIT) may provide a more real-world approximation of certain drug prices than the British National Formulary (BNF) in certain cases. As the ERG identified, using eMIT rather than BNF costs only has a small effect on the ICER, however we have conducted a new analysis using eMIT prices the new base-case in order to eliminate this uncertainty (see Table 10). This results in a small reduction in the ICER of £255.

Table 10: Scenario analysis exploring the impact of using eMIT costs instead of BNF, where appropriate

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case
Established management	£59,873	1.61	1.03	[redacted]	[redacted]	[redacted]	£49,103	-£255

market information tool (eMIT)?

Entrectinib	[redacted]	[redacted]	[redacted]					

Issue 17: Administration costs and resource use

36. Have administration costs been adequately captured in the company's model?

Roche implemented a simplified clustering approach to implementing comparator treatment administration costs in the model due to the number of possible comparator treatments. Due to time constraints and the anticipated small impact of individualised treatment administration costs, we have not been able to conduct the requested scenario analysis. However, we have implemented the oral chemotherapy tariff cost to both entrectinib and the comparator oral chemotherapies. The results of this are shown in Table 11. While comparator administration costs are slightly increased, there is a more notable increase in entrectinib total costs, resulting in a small increase to the ICER of £2,118.

Table 11: Revised administration costs

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)	Change from base-case
Established management	£61,401	1.61	1.03	[redacted]	[redacted]	[redacted]	£51,476	£2,118
Entrectinib	[redacted]	[redacted]	[redacted]					

37. Has resource use been adequately captured in the company's model?

Roche's assessment of resource use in the progressed disease state was based on a recent appraisal for eribulin in previously-treated metastatic breast cancer (NICE TA515). The components and costs were accepted by the ERG and committee. This was used as a recent proxy for the progressed disease state in tumour types covered in this appraisal. The costs and resource use therefore match this appraisal, with updated reference costs.

However, we note that in other recent oncology appraisals for previously treated cancers (for example, TA428, TA483, TA484, and TA520), monitoring costs are also included in the progressed disease state. We have therefore provided a scenario whereby these costs are added. The revised progressed state costs are shown in Table 12; it should be noted that these represent potentially conservative estimates, since clinician feedback suggest there would be minimal imaging after progression on a treatment late in a patient's pathway. Incorporating these costs raises the monthly progressed disease state healthcare resource use costs to £402.93 from £331.59, resulting in a small increase to the ICER of £289.

Table 12: Additional progressed disease state costs

Item	Number used	% of patients	Unit cost	Monthly cost	Reference*	Source
Medical oncology, outpatient visit	1	100	£162.05	£162.05	Total Outpatient Attendances -row 370(Outpatient, consultant-led)	TA515
Medical oncology, outpatient, nurse-led	1	100	£104.00	£104.00	Total Outpatient Attendances -row 370(Outpatient, non consultant-led)	TA515
GP home visit	1	100	£37.40	£37.40	PSSRU 2018(143): 10.3b GP unit costs (9.22 minutes patient time)	TA515
Nurse community visit	1	67	£42.00	£30.15	PSSRU 2018(143): 10.2 GP practice nurse unit costs (assumes 1 hour patient contact)	TA515
Full blood count	1	100	£2.51	£2.51	DAPS05 (haematology)	TA520
Liver function test	0.67	100	£1.11	£1.11	DAPS04 (clinical biochemistry)	TA520
Renal function test (with electrolytes)	0.67	100	£1.11	£1.11	DAPS04 (clinical biochemistry)	TA520
CT scan	0.41	100	£132.75	£54.43	RD22Z (CT scan of one are, pre-and post-contrast, outpatient)	TA520
X-ray	0.41	100	£31.49	£12.91	DAPF	TA484
Average monthly cost				£402.93		

Issue 18: Implementation and training costs	
38. What additional infrastructure and training requirements	Whole-genome sequencing in sarcomas and paediatrics via the genomic laboratory hubs (GLH) is nearing readiness. Over the next 12 months to 2 years the implementation of NGS-based testing across all tumour types will continue to ramp up. During this time there is an ideal opportunity to engage with the wider clinical community to educate and increase awareness around genomics in medicine. We are in agreement that the primary additional training requirement would be for the

could be considered for this appraisal?	education of oncologists and pathologists in the new concepts associated with tumour-agnostic therapies. Roche intends to collaborate closely with the NHS to support training in these areas. Dependent on the chosen optimal testing approach there may also be additional training requirements associated with sample handling of biopsy material. Particular consideration would be required around appropriate material handling where RNA extraction is required for NGS-based testing should a DNA/RNA hybrid panel approach be implemented. These training requirements are however very much in line with the existing training requirements associated with wide scale implementation of NGS-based testing within the NHS. Again, Roche intends to collaborate closely with the NHS to support training in these areas.
Issue 19: Utility values	
39. Is the utility value estimate for the entrectinib progression-free health state collected in the STARTRK-2 trial appropriate?	Roche agrees with the ERG that the utility estimate derived from patients in the STARTRK-2 trial for entrectinib is appropriate for reasons outlined in our initial submission and at clarification questions. Though there is potential variation between the line of therapy at which entrectinib was used in the trial versus what may happen in clinical practice, the data and rationale provided in response to Issue 2, question 4 suggests that in practice this may not have an impact on outcomes, including quality of life.
40. Is the utility value for the established management progression-free health state	Roche agrees with the ERG that the utility estimate derived for the comparator population is appropriate for reasons outlined in our initial submission and at clarification questions. The utility estimates used were robust in that they were collected from patients with relevant tumour types at an advanced/metastatic stage of disease, were aligned with the model structure (PFS/post-progression) and were consistent with the NICE reference case.

appropriate (0.73)?	
41. Should the utility value for the progression-free state be equal between the entrectinib and established management arms? If yes, what should the utility value be?	Roche is in agreement with the ERG that the base-case utility values are plausible and that the utility values should therefore not be equal between entrectinib and established management in the PFS state. Entrectinib is an oral TKI therapy with a more convenient administration and relatively tolerable safety profile when compared with traditional cytotoxic chemotherapies, which form the majority of comparator products.
42. Is it reasonable to assume that the utility value for the progressed disease health state be equal between the entrectinib and established management arms? If yes, is the value	Roche's position is that an equal utility value of 0.59 is appropriate across the entrectinib and established management arms, the value being derived in the same manner as the established management PFS utility value. However, we acknowledge that there is some uncertainty in the entrectinib progressed state utility value, since the data collected from STARTRK-2 were not appropriate for use given the limited number of observations.

of 0.59 appropriate?	

Issue 20: Uncertainty around the cost-effectiveness results

43. Is the uncertainty around the cost-effectiveness results captured appropriately in the company's model?

Roche's view is that the probabilistic sensitivity analysis (PSA) adequately captures the uncertainty associated with the appraisal. However, in order to try to address this issue, we have amended parameter uncertainties in the following ways:

- Standard errors for comparator PFS and OS have been increased from 0.15 to 0.25.
- Confidence intervals for comparator life-years gained, quality adjusted life-years gained and costs have been widened.

The new comparator confidence intervals and the results of the new PSA (2000 iterations) are presented in Table 13.

Due to inconsistent reporting, time constraints and anticipated minimal impact, standard errors were not extracted from original comparator source guidance.

Table 13: Updated PSA with broader comparator standard error and confidence intervals

	Deterministic				Probabilistic			
Technologies	**Total costs (£)**	**Total LYG**	**Total QALYs**	**ICER incremental (£/QALY)**	**Total costs (£)**	**Total LYG**	**Total QALYs**	**ICER incremental (£/QALY)**
Established management	£61,228	1.61	1.04	£49,358	[redacted]	[redacted]	[redacted]	£49,090
Entrectinib	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	

Issue 21: End of life

44. What is the life expectancy of the patient

In Roche's revised base-case, the mean discounted OS for entrectinib and the weighted comparator are 35.8 and 19.3 months, respectively, suggesting both that the current standard of care results in mean OS of less than two years, and that the mean OS gain conferred by entrectinib is 16.5 months. Using the NHSE and NHSI's preferred comparators results in a discounted mean OS of 18.9 months, resulting in a mean OS gain of 16.9 months. As such, Roche is in agreement with

group receiving established management ?	NHSE and NHSI that entrectinib meets both end-of-life criteria. In further support of criterion 1, Roche anticipates use of entrectinib in accordance with its anticipated label [redacted].
45. What is the extension to life of the patient group receiving entrectinib?	With regard to the ERG's four concerns as to why criterion 1 is not met, Roche disputes these for the following reasons: • The reason we considered use at more than one line of treatment for a given tumour type was to account for ambiguity in interpretation of the anticipated licence; in reality, testing may occur at diagnosis with a view to using entrectinib at a later line. Roche has also taken steps to mitigate against this criticism by conducting a scenario analysis using NHSE and NHSI's preferred pathway positioning at later lines, which validates to a degree the assumption that standard of care outcomes is slightly overestimated in our base-case. This gives further confidence that criterion 1 is met. • We acknowledge that the life expectancy of unrepresented tumour types is not known, which leads to a degree of uncertainty regarding whether criterion 1 is met. However, as previously stated, Roche anticipates use of entrectinib in accordance with its anticipated label [redacted]. In addition, this uncertainty may be addressed during a potential CDF data collection period, which Roche has proactively proposed for entrectinib. • We acknowledge that the existing evidence supporting NTRK gene fusions being a poor prognostic indicator lacks robustness. However, the majority of the limited evidence available to date are suggestive of this scenario (11-13). We anticipate that this uncertainty may be addressed during a potential CDF data collection period. • We agree that such a heterogeneous patient population is likely to encompass heterogeneous life expectancies dictated in part by the nature of the tumour type. However, as previously stated, Roche anticipates use of entrectinib in accordance with its anticipated label [redacted]. The ERG also quote high mean overall survival figures for thyroid and neuroendocrine tumours of 44.7 and 57.1 months, respectively. While these tumours may have a relatively good

	prognosis, these figures are heavily confounded by cross-over to active therapy in the case of BSC. By way of illustration, the median OS on placebo in the lenvatinib SELECT trial was 34.5 months (NICE TA535), though when adjusted for cross-over it was just 19.1 months (14). These data also do not account for the very poor prognosis of anaplastic thyroid cancer, which has a median OS of three months (15). In addition, of the three prognostic studies currently published, two are in thyroid cancers, and both show a worsened prognosis compared with wild-type or other oncogene-driven disease (12, 13), suggesting that comparator survival outcomes in this tumour type may be overestimated. Finally, it should also be noted that, even in a cohort of patients within single, poor-prognosis tumour type such as previously-treated non-small cell lung cancer, there is likely to be a cohort of long term survivors even on chemotherapy such as docetaxel; this is demonstrated by recently published results showing that 2.6% of docetaxel-treated patients remained alive after five years (16). With respect to the ERG's uncertainty regarding criterion 2, by the ERG's own estimation, life extension in unknown tumour types could range from [redacted] months, suggesting that life extension greater than three months is by far the more likely outcome. Finally, uncertainty surrounding efficacy in unknown tumour types is likely to be addressed during a potential CDF data collection period, for which Roche proposes a comprehensive data collection plan.
Issue 22: Innovation	
46. Is entrectinib an innovative treatment?	Roche believes entrectinib is an innovative treatment for reasons outlined in the initial company submission.
Issue 23: Cancer Drugs Fund	
47. Does entrectinib meet the criteria for inclusion in the Cancer Drugs Fund?	Roche agrees with the NICE technical team's view that entrectinib meets the criteria for inclusion into the CDF. As discussed in our response to Issue 21, entrectinib comprehensively meets both end-of-life criteria. In addition, our revised base-case ICER of £49,358, which includes 100% of testing costs, falls under the maximum end-of-life threshold. Also of note, the ICER with testing costs excluded, which we believe is more appropriate, is considerably lower (£35,770) and closely mirrors the ERG's base-case ICER excluding testing (£38,030).
48. What data	Roche has proactively proposed entry of entrectinib into the CDF as we acknowledge that there are a number of clinical- and cost-effectiveness uncertainties due to the novel, tumour agnostic nature of the product and its resulting atypical evidence base.

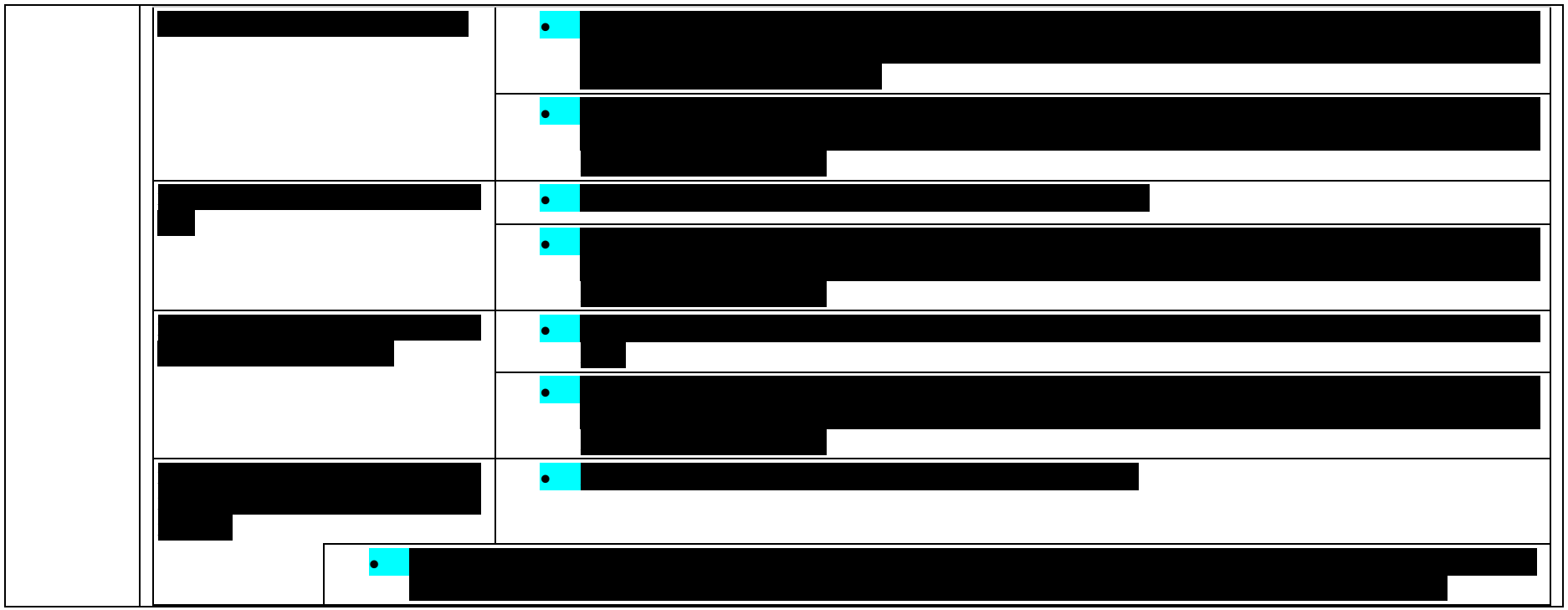
would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.

Roche's proposed data collection concepts are summarised in Table 14, categorised by uncertainties that they address. [redacted].

Data collected by Public Health England (PHE) during a CDF period has the potential to help address a broad range of uncertainties; however, we acknowledge that the nature, scope and quality of data collected by PHE is yet to be determined and so we have not designated particular uncertainties to this method. Roche will collaborate closely with NICE, NHS England and PHE to help determine this, and to further elaborate on our currently existing proposals through the drafting of a data collection agreement.

Table 14: Roche's CDF data collection proposals

Uncertainty/data gap	Data collection proposals
[redacted]	• [redacted] • [redacted] • [redacted] • [redacted]
[redacted]	• [redacted] • [redacted] • [redacted] • [redacted]



References

1. Okamura R, Boichard A, Kato S, Sicklick JK, Bazhenova L, Kurzrock R. Analysis of NTRK Alterations in Pan-Cancer Adult and Pediatric Malignancies: Implications for NTRK-Targeted Therapeutics. JCO Precis Oncol. 2018;2018.
2. Gatalica Z, Xiu J, Swensen J, Vranic S. Molecular characterization of cancers with NTRK gene fusions. Modern Pathology. 2019;32(1):147-53.
3. Solomon JP, Linkov I, Rosado A, Mullaney K, Rosen EY, Frosina D, et al. NTRK fusion detection across multiple assays and 33,997 cases: diagnostic implications and pitfalls. Modern Pathology. 2019.

4. Marchiò C, Scaltriti M, Ladanyi M, Iafrate AJ, Bibeau F, Dietel M, et al. ESMO recommendations on the standard methods to detect NTRK fusions in daily practice and clinical research. Annals of Oncology. 2019.
5. Lange AM, Lo HW. Inhibiting TRK Proteins in Clinical Cancer Therapy. Cancers. 2018;10(4).
6. Krajewski KM, Braschi-Amirfarzan M, DiPiro PJ, Jagannathan JP, Shinagare AB. Molecular Targeted Therapy in Modern Oncology: Imaging Assessment of Treatment Response and Toxicities. Korean J Radiol. 2017;18(1):28-41.
7. Fan J, Fong T, Xia Z, Zhang J, Luo P. The efficacy and safety of ALK inhibitors in the treatment of ALK-positive non-small cell lung cancer: A network meta-analysis. Cancer medicine. 2018;7(10):4993-5005.
8. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1983;1(11):710-9.
9. Ali A, Goffin JR, Arnold A, Ellis PM. Survival of patients with non-small-cell lung cancer after a diagnosis of brain metastases. Current oncology (Toronto, Ont). 2013;20(4):e300-e6.
10. Roughley A, Damonte E, Taylor-Stokes G, Rider A, Munk VC. Impact of Brain Metastases on Quality of Life and Estimated Life Expectancy in Patients with Advanced Non-Small Cell Lung Cancer. Value Health. 2014;17(7):A650.
11. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A, et al. ALK, ROS1, and NTRK Rearrangements in Metastatic Colorectal Cancer. J Natl Cancer Inst. 2017;109(12).
12. Musholt TJ, Musholt PB, Khaladj N, Schulz D, Scheumann GF, Klempnauer J. Prognostic significance of RET and NTRK1 rearrangements in sporadic papillary thyroid carcinoma. Surgery. 2000;128(6):984-93.
13. Prasad ML, Vyas M, Horne MJ, Virk RK, Morotti R, Liu Z, et al. NTRK fusion oncogenes in pediatric papillary thyroid carcinoma in northeast United States. Cancer. 2016;122(7):1097-107.
14. Guo M, Sherman S, Wirth L, Schlumberger M, Dutcus C, Robinson B, et al. 2805 Overall survival gain with lenvatinib vs. placebo in radioactive iodine refractory differentiated thyroid cancer (RR-DTC): An updated analysis. European Journal of Cancer. 2015;51:S559.
15. Kebebew E, Greenspan FS, Clark OH, Woeber KA, McMillan A. Anaplastic thyroid carcinoma. Cancer. 2005;103(7):1330-5.
16. Bristol Myers Squibb. Bristol-Myers Squibb Announces Pooled Five-Year Survival Results for Opdivo (nivolumab) in Previously-Treated Advanced Non-Small Cell Lung Cancer Patients 2019 [Date Accessed 27th September 2019]. Available from: https://news.bms.com/press-release/corporatefinancial-news/bristol-myers-squibb-announces-pooled-five-year-survival-resul.

Additional information request: Solid tumours (NTRK positive) - entrectinib [ID1512]

In response to NICE's request for additional information on 7th October 2019, we are providing six and 12 months Kaplan Meier (KM) data by tumour type for progression-free (PFS) and overall survival (OS) (see Table 1). We note that NICE wishes to use this data to apply Bayesian Hierarchical Modelling (BHM) to these time-to-event (TTE) outcomes in a similar manner to the approach taken with response data.

In their report, the ERG acknowledge that it is unclear whether this type of model would provide useful results with TTE outcomes, given the immaturity of the survival data and the small number of patients in most tumour types. Our perspective, as stated in our responses to Issues 7 and 9 of the technical report, is that this approach will not provide any useful information for decision making for the reasons that the ERG highlight. This is demonstrated by the KM curves provided in our response Issue 9, question 20 and the data provided in Table 1. In almost all cases,

[redacted]. As with response rates, any estimates derived from a BHM framework applied to TTE outcomes are therefore likely to result in extremely wide confidence intervals, rendering them inappropriate for decision making.

We reiterate the point that tumour heterogeneity is most reliably explored through continued data collection from the clinical trial programme and other related data collection activities during a potential Cancer Drugs Fund period.

Table 1: Entrectinib PFS and OS events at 6 and 12 months

Tumour ID	Tumour type	Number of patients (n)	Number of responders (x)	Number of PFS events at 6 months	Number of PFS events at 12 months	Number of deaths at 6 months	Number of deaths at 12 months
1	Sarcoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	NSCLC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	CRC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Neuroendocrine tumours	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	Pancreatic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	Gynaecological	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7	Cholangiocarcinoma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	MASC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

9	**Breast**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	**Thyroid**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11	**CNS Primary**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	**Paediatric CNS Primary**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13	**Paediatric (non-CNS)**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

As requested, we have also provided Excel model scenarios and PDFs of references used in the technical engagement responses.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **Wednesday 2 October 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow. If confidential information is submitted, please also send a second version of

your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	**Jayne Bressington**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**GIST Support UK - respondent**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Prevalence of NTRK gene fusion	
Which dataset more accurately reflects the prevalence of NTRK fusion for each tumour type?	
Issue 2: Treatment pathway and positioning	
How will 'acceptability' in the context of this appraisal be defined in clinical practice?	
For each tumour type, at what points in the respective treatment pathways will entrectinib be used in clinical practice in England?	We believe that if a GIST patient is found to carry an NTRK fusion that they should be given entrectinib as a first line therapy. (If however, because it is a new test the NTRK fusion is only discovered after they have been prescribed the other standard GIST therapies then it would in that case be classified as a 2nd, 3rd or 4th line therapy).
Is the evidence for entrectinib generalisable to clinical practice in England?	
Issue 3: NTRK gene fusion screening pathway	
What is the likely screening pathway to identify NTRK positive solid tumours?	The standard screening pathway for GIST is as follows: 1. **Molecular testing to review KIT** (exons 8, 9, 11, 13, 17) and **PDGFRA** (exons 12, 14 and 18) mutation analysis. 2. **Immunohistochemistry for SDHB** **3. SDHA immunohistochemistry** 4. **BRAF** is often performed as part of a NGS multi gene panel. IHC 5. **NF1**, skeinoid fibers, clinical input. 6. **NTRK - IHC/fusion**

At what point in the treatment pathway for each tumour type is/will NTRK gene fusion testing carried out?	The point at which NTRK fusion gene testing is carried out will vary for each patient. The testing will only happen where surgery or a biopsy has provided a sample of the tissue that can be tested. From a GIST perspective NTRK gene fusion testing will only be carried out when it has been identified that the patient has tested negative for all of the other known GIST mutations. This group of patients are currently classified as **"quadruple negative GIST".**
Issue 4: NTRK gene fusion testing costs	
Should testing costs for molecular testing that is already done in the NHS be included in the economic model, even if it is not to detect NTRK gene fusions?	
Are the costs of adding a panel to an RNA-based Next Generation Sequencing (NGS) test negligible?	
Is it appropriate to include the costs of confirmatory NGS for people who have Whole Genome Sequencing (WGS)?	
Should the economic model include testing costs for unrepresented tumour types?	
What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?	The screening for NTRK fusions is already naturally part of the whole genome sequencing panel in GIST and sarcoma so we think that this indicates that the testing costs should be excluded.
Issue 5: Identification of NTRK gene fusions - diagnostic accuracy	
What is the expected diagnostic accuracy of Next Generation Sequencing (NGS) testing?	Very accurate
Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types?	

Is it appropriate to limit testing to avoid false positive results and the associated costs?	
Issue 6: Distribution of tumour types	
Are the ERG's estimates of the distribution of tumour types reflective of what would be seen in clinical practice in England? If no, what is the likely distribution of NTRK positive tumour types in clinical practice?	It is estimated that c14% of wild-type GIST patients are "quadruple negative". It is estimated that there is c.125 WT GIST patients p.a. so 14% represents 17 patients. Of these 50% are likely to be clear of disease after surgery and will not require therapy. The other 50% c. 8 patients will be screened for NTRK fusions and rare ones will be found to have NTRK fusions.
Issue 7: Unrepresented NTRK gene fusion positive solid tumour types	
Are the results of the entrectinib studies that includes 13 different NTRK positive tumours sites generalisable to all NTRK-fusion positive tumour types?	
Issue 8: Primary CNS tumours and paediatric tumours	
Should people with primary CNS tumours and paediatric tumours be included in the analysis?	Paediatric patients with NTRK fusions should be allowed access to entrectinib.
Issue 9: Heterogeneity of response across different solid tumour types	
Is a uniform level of response to entrectinib across the different tumour types with a NTRK fusion reasonable?	
Is the Bayesian Heirarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal?	

Would it be appropriate to apply the Bayesian Heirarchical Modelling (BHM) framework to explore the heterogeneity in the time to event outcomes?	
Are the response rates in tumour sites not represented in the trial data suitable for consideration?	
Issue 10: Constructing a comparator arm	
Is the company's comparator arm suitable for decision making?	
Is it appropriate to use non-responders as a proxy for patients not having an active treatment to generate a comparator arm for this appraisal?	
Issue 11: Comparator treatments	
Is the company's modelled comparator population representative of the eligible population in England?	
Should the comparator data be matched to the position that entrectinib is given in the efficacy evaluable dataset?	
Issue 12: Prognostic factors	
Is it appropriate to adjust the comparator data to account for worsened prognosis of CNS metastases?	
Is it appropriate to adjust the comparator data to adjust for NTRK fusions? If yes, are there any additional evidence sources which could be used to inform whether presence of an NTRK fusion is prognostic?	

Issue 13: Subsequent therapies	
What percentage of people, who receive entrectinib, would be expected to receive subsequent therapy following disease progression?	100% of the NTRK fusion GIST patients who progress on entrectinib would expect to receive subsequent therapy. Hopefully this would be a drug such as Larotrectinib which we understand could be used in sequence to overcome recurrent resistance mutations should they arise.
What percentage of people, who receive established management, would be expected to receive subsequent therapy following disease progression?	100% of GIST patients who receive established management would expect to receive subsequent therapy following disease progression.
Which subsequent therapies would be used and in what proportions? Please answer for following entrectinib and following established management.	Entrectinib – another NTRK fusion inhibitor?, Imatinib, Sunitinib, Regorafenib, Clinical trial Established – Imatinib, Sunitinib, Regorafenib, Clinical trial
How long would these people be expected to be treated with subsequent therapy? Please answer for following entrectinib and following established management.	
Issue 14: Model structure	
What is the most appropriate model structure for this appraisal?	
Issue 15: Extrapolation of overall and progression-free survival	
Is the exponential or Weibull distribution most appropriate for extrapolating overall and progression-free survival?	

Issue 16: Drug wastage and source of treatment costs	
Is it appropriate to assume no drug wastage of entrectinib?	
Should drug acquisition costs for the comparator therapies be sourced from the British National Formulary (BNF) or the electronic market information tool (eMIT)?	
Issue 17: Administration costs and resource use	
Have administration costs been adequately captured in the company's model?	
Has resource use been adequately captured in the company's model?	
Issue 18: Implementation and training costs	
What additional infrastructure and training requirements could be considered for this appraisal?	
Issue 19: Utility values	
Is the utility value estimate for the entrectinib progression-free health state collected in the STARTRK-2 trial appropriate?	
Is the utility value for the established management progression-free health state appropriate (0.73)?	
Should the utility value for the progression-free state be equal between the entrectinib and established management arms? If yes, what should the utility value be?	

Is it reasonable to assume that the utility value for the progressed disease health state be equal between the entrectinib and established management arms? If yes, is the value of 0.59 appropriate?	
Issue 20: Uncertainty around the cost-effectiveness results	
Is the uncertainty around the cost-effectiveness results captured appropriately in the company's model?	
Issue 21: End of life	
What is the life expectancy of the patient group receiving established management?	
What is the extension to life of the patient group receiving entrectinib?	
Issue 22: Innovation	
Is entrectinib an innovative treatment?	**Yes**
Issue 23: Cancer Drugs Fund	
Does entrectinib meet the criteria for inclusion in the Cancer Drugs Fund?	
What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.	

NICE National Institute for Health and Care Excellence

Technical engagement response form

Entrectinib for treating NTRK fusion-positive solid tumours [ID1512]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments: **Wednesday 2 October 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under [redacted], all information submitted under [redacted]. If confidential information is submitted, please also send a second version of your

comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	**Professor Donal O' Donoghue, submitting on behalf of:**
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**NCRI-ACP-RCP-RCR**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Questions for engagement

Issue 1: Prevalence of NTRK gene fusion	
Which dataset more accurately reflects the prevalence of NTRK fusion for each tumour type?	The ERG dataset appears more accurate for NTRK fusion prevalence based on the large Foundation Medicine (FM) dataset. The only caveat about the true prevalence of the NTRK fusion is how robust the FM assay is at detecting NTRK fusion as a DNA-based only test (no RNA component). There is a possibility the true prevalence may be underestimated. However, the FM dataset would seem to be the most robust data available at the current time. Prevalence is well described in secretory carcinoma and fibrosarcoma. In other groups the prevalence described will be determined by the populations screened in each tumour type.
Issue 2: Treatment pathway and positioning	
How will 'acceptability' in the context of this appraisal be defined in clinical practice?	This will vary between tumour types primarily determined by the alternative standard of care options. We agree the term 'acceptable alternatives' is ambiguous and open to interpretation. In general, post first-line treatment would be appropriate, pending further data collection although there may be circumstances where the risk-benefit ratio could favour entrectinib over first-line SOC treatment. Suggested indications are specified below.

For each tumour type, at what points in the respective treatment pathways will entrectinib be used in clinical practice in England?	As above, this will vary between tumour types primarily determined by the level of evidence supporting alternative standard of care options. Overall we would agree with NHS England's suggested positioning (Table 4) with some minor edits detailed below. We would also encourage adding in some of the paediatric indications such as infantile fibrosarcoma given the high TRK fusion positivity in this setting. Specific disease types: MASC –first-line Soft tissue sarcoma –first-line for chemo-resistant sarcomas and second-line and beyond for chemo-sensitive sarcomas Pancreatic cancer –first or second line. This may depend on the clinical situation/urgency for treatment and turnaround time /success of genetic testing on small biopsy/cytology samples Cholangiocarcinoma – likely second-line and beyond as only n=1 patient treated with cholangiocarcinoma on the trial and not clear which line of treatment this was. NSCLC – we would propose the option of first-line and beyond for this indication. Activity of tyrosine kinase inhibitors in genomically driven sub-types of NSCLC is well established and the community has preference to treat with targeted therapies in early lines of therapy. Breast – we would propose second-line and beyond Thyroid cancer – second-line and beyond would seem reasonable at this stage although as further data are collected, we would anticipate moving into a first-line indication at a later stage. Colorectal cancer – third line and beyond Neuroendocrine carcinomas – consider second line and beyond

	Paediatric tumours with high TRK fusion prevalence (e.g. infantile fibrosarcoma) – first line and beyond
Is the evidence for entrectinib generalisable to clinical practice in England?	Data are available within the ALKA and STARTRK studies for overall a relatively small number of patients (n=54). However, it is recognised that TRK fusion occurs at low frequency across many disease types and the data within the trials are highly encouraging for responses across different diseases. On balance, we feel the data are generalizable in the context of a TRK fusion positive patient (rather than considering traditionally by disease type and line of therapy). However, we would agree that there is discrepancy between the trial data and proposed positioning and feel the indications proposed above are a reasonable approach. Further data collection in context of the CDF, if approved, would be vital in assessing outcomes and providing supportive data to bring into earlier lines of therapy.
Issue 3: NTRK gene fusion screening pathway	
What is the likely screening pathway to identify NTRK positive solid tumours?	IHC followed by confirmatory NGS or up front NGS – determined by resource. We would defer to NHS England on this matter in respect of plans to roll out NGS DNA/RNA panel testing. Issues to consider are 1) An NGS test for all patients potentially covered by this appraisal would be an ideal approach and would permit for parallel screening of other molecular alterations that may be actionable for other treatment (thus more cost-effective) 2) The NGS approach must be robust to detect NTRK fusions. DNA panel sequencing is good but may still result in false negatives and restricts the ability to identify patients with unusual/unknown fusion partners. An RNA-based approach may be preferable if sufficiently validated on FFPE tissue 3) The roll out of NGS testing may not be until 2021. An interim measure may therefore need to be in place to support CDF access (if approved). A 2-step approach with IHC could be considered but this may take 1-2 years to establish and embed in path labs by which point

	NGS may be available. Would the company consider supporting access to NGS testing as an interim measure? 4) We note WGS will be available for paediatric and sarcoma patients. These patients would still likely require a separate RNA fusion panel to identify a TRK fusion as WGS is limited in ability to detect these alterations. Thus testing for TRK still needs to be considered for these patients.
At what point in the treatment pathway for each tumour type is/will NTRK gene fusion testing carried out?	Ideally this should be performed early in the patient's diagnostic pathway for patients with advanced, unresectable cancers to best inform treatment choice. If not feasible testing could be undertaken during standard first-line treatment to best inform second-line treatment options. Panel testing is ideal here to identify other potentially actionable alterations and could ideally be included in the economic model. For those disease types with known high prevalence of TRK fusion such as MASC, secretory carcinoma of breast, upfront testing (pre first-line treatment) is essential and are already included in the test directory.
Issue 4: NTRK gene fusion testing costs	
Should testing costs for molecular testing that is already done in the NHS be included in the economic model, even if it is not to detect NTRK gene fusions?	No
Are the costs of adding a panel to an RNA-based Next Generation Sequencing (NGS) test negligible?	We believe this should read the cost of adding a 'single gene' to an RNA-based NGS test. If so, indeed the cost of this should be negligible.
Is it appropriate to include the costs of confirmatory NGS for people who have Whole Genome Sequencing (WGS)?	Inclusion of confirmatory NGS costs seem reasonable. A confirmatory test may be required but of equal concern it should be noted that WGS may not efficiently detect fusions (unless the assays/bioinformatics are further optimised by the time of roll-

	out). Thus an additional DNA/RNA based test may be needed to identify fusions in the patient groups that have will have access to WGS. In point 4 referred to by the ERG in Issue 4, it is unclear how a figure of 11% has been reached.
Should the economic model include testing costs for unrepresented tumour types?	Yes this seems reasonable accepting the requirement for prospective/phase 4 data collection in these cohorts.
What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?	NTRK fusion is a diagnostic test in secretory carcinoma so should not be included in the model in this group however it is not performed routinely in the other tumour types with low NTRK fusion prevalence so it seems reasonable to include the costs of these other patients in the model.
Issue 5: Identification of NTRK gene fusions - diagnostic accuracy	
What is the expected diagnostic accuracy of Next Generation Sequencing (NGS) testing?	Utility of NGS primarily determined by the quality of the sample input. Our experts note that analysis failure is seen in approximately 20% of cases that screened by NGS. Sensitivity and specificity is however very high if the analysis is successful. The specificity of NGS is expected to be very high (close to 100%). The sensitivity is uncertain and from clinical judgement for DNA NGS may be in the region of 70% (unsubstantiated). RNA testing would provide higher sensitivity (if assays are optimised for FFPE tissue). The likelihood of false positives is in our view very low in the case of fusions. There are stringent QA criteria applied to NGS testing in the Genomic Laboratory Hubs and in general, DNA and RNA based NGS

	assays have very low false positivity rates, especially considering the controls that would be run with each sample. We disagree with the suggestion that the differential responses in the clinical trial may be due to false positive NTRK results. Differential responses may be the result of a host of different factors including, tumour heterogeneity, co-existing mutations, pharmacokinetic variabiltiy etc.
Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types?	As per the European Recommendation https://www.ncbi.nlm.nih.gov/pubmed/31268127 'In tumours where NTRK fusions are highly recurrent, FISH, RT-PCR or RNA-based sequencing panels can be used as confirmatory techniques, whereas in the scenario of testing an unselected population where NTRK1/2/3 fusions are uncommon, either front-line sequencing (preferentially RNA-sequencing) or screening by immunohistochemistry followed by sequencing of positive cases should be pursued.' In our opinion, the issue will not be about identifying false positives, rather avoiding false negatives. A RNA-based NGS approach would be optimal for identification of fusions but would depend on good quality FFPE tissue. A combined DNA/RNA NGS approach could be used to reduce the risk of false negatives (and would also maximise the chance of identifying other molecular alterations that could be actionable [treated with a standard-of-care treatment or guide clinical trial selection]). An alternative would be a 2-step diagnostic approach such as IHC followed by confirmatory DNA/RNA testing but this would not be in line with NHS England ambitions to roll out NGS. In practice, the individual GLHs will select and fully validate their preferred assays for identifying TRK fusion and we would rely on their expertise to minimise false negative rates. .

Is it appropriate to limit testing to avoid false positive results and the associated costs?	Not applicable
Issue 6: Distribution of tumour types	
Are the ERG's estimates of the distribution of tumour types reflective of what would be seen in clinical practice in England? If no, what is the likely distribution of NTRK positive tumour types in clinical practice?	The distributions are hard to determine but we recognise the FM dataset may be the best guide as to the relative prevalence of NTRK fusion. Presumably the percentages presented by the ERG then reflect the relative incidence of the various cancer types. Mostly the figures look reasonable but would highlight the following exceptions: Breast cancer([redacted]) – whilst this is a common tumour type, it is the rare secretory breast cancer sub-type that has the highest prevalence of TRK fusion. We would question whether the histological sub-types of breast cancer in the Foundation Medicine cohort are known to assess whether the prevalence may be skewed. We would anticipate breast cancer patients as a whole to represent a smaller proportion of patents in the model , perhaps around [redacted] Thyroid cancer: ([redacted]) – intuitively this would seem a high proportion of patients and is based on the prevalence seen in the Foundation Medicine cohort. Again, we would question whether the histological sub-type of patients is known in the FM cohort and whether these are representative of the UK population. If so, we would accept this estimation.

	Colorectal cancer ([redacted] – the prevalence of TRK fusion is considered [redacted] (same as NSCLC). As lung cancer is more common than colorectal cancer we would anticipate this cohort to be proportionately lower (unless NSCLC sub-type (excluding SCLC) is similar to CRC incidence). Cholangiocarcinoma – this is estimated to be [redacted] but this seems unlikely given the [redacted] prevalence presented in Table.6. The figures in the table add up to 99%; should these be 100% or does the 1% account for other tumour types not covered here that may harbour a TRK fusion? In particular, should paediatric and primary CNS tumours also be included here.
Issue 7: Unrepresented NTRK gene fusion positive solid tumour types	
Are the results of the entrectinib studies that includes 13 different NTRK positive tumours sites generalisable to all NTRK-fusion positive tumour types?	There is likely to be clinically significant variability in response rates between tumour sites which will have widely varied genomic alterations in addition to NTRK fusion. Data to not exist to support this but it would be reasonable to assume a similar efficacy rate would be seen given the results to date across 13 different tumour types. One could look to publically-available data within the larotrectinib studies (competing TRK inhibitor) to determine if any other tumour types have been recruited and associated response. It is reasonable to consider the drugs have a similar level of efficacy.
Issue 8: Primary CNS tumours and paediatric tumours	

Should people with primary CNS tumours and paediatric tumours be included in the analysis?	Yes. This group may have a different outcome due to their disease biology (as per issue 7 above) and the clinical symptoms of CNS disease, however ideally (with enough data) all sites will be analysed as sub-groups.
Issue 9: Heterogeneity of response across different solid tumour types	
Is a uniform level of response to entrectinib across the different tumour types with a NTRK fusion reasonable?	We would agree with the comments and issues raised by the technical team in relation to this question. It is currently uncertain whether responses would be heterogeneous and the proposed additional data and modelling are appropriate. One point for consideration is that generally fusions are a strong oncogenic driver of disease (and often in the absence of other genomic alterations) compared with a somatic mutation. With the latter there can be considerable heterogeneity between disease types in terms of contribution to disease progression (and often with co-mutations contributing) thus differential responses are frequently seen with tyrosine kinase inhibitors. One may extrapolate, therefore, that inhibition of a signalling pathway (TRK), and thus response to treatment, being driven by a genomic fusion may be more homogeneous. However, this is based on expert clinical opinion rather than robust supporting data.
Is the Bayesian Heirarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal?	Our experts were unsure on how to best capture this heterogeneity in the absence of large numbers of patients with each tumour type.
Would it be appropriate to apply the Bayesian Heirarchical Modelling (BHM) framework to explore the heterogeneity in the time to event outcomes?	

Are the response rates in tumour sites not represented in the trial data suitable for consideration?	This depends how the data was/is collected.
Issue 10: Constructing a comparator arm	
Is the company's comparator arm suitable for decision making?	The comparator survival data seems reasonable for the 'beyond standard of care' cohort. We agree with the ERGs assessment of the flaws associated with the company proposed comparator arms. The use of real-world data would be preferable but availability of these data are likely to be highly limited.
Is it appropriate to use non-responders as a proxy for patients not having an active treatment to generate a comparator arm for this appraisal?	This would have some value but may introduce bias as there will be biological reasons patients do not respond which may in themselves be prognostic (likely associated with adverse prognosis). This approach is also potentially flawed as even in the absence of tumour response, there is still the possibility/likelihood that disease stabilisation may occur and thus impact on PFS and OS. Assessing time to PFS on previous treatment would be a valuable approach but acknowledge these data may not be available if not specifically collected in the supporting trials.
Issue 11: Comparator treatments	
Is the company's modelled comparator population representative of the eligible population in England?	It seems reasonable although, the lines of therapy used for the comparator arms by the company do not fully match the line of therapy/reflect current practice compared with the line of treatment entrectinib was administered in the clinical trials.

Should the comparator data be matched to the position that entrectinib is given in the efficacy evaluable dataset?	Yes, we agree, although acknowledge that the intended positioning of entrectinib may be different from that used in the clinical trials.
Issue 12: Prognostic factors	
Is it appropriate to adjust the comparator data to account for worsened prognosis of CNS metastases?	In our opinion there are insufficient data to determine the impact of the presence of CNS metastases on prognosis in the NTRK fusion population. However, we would point out that usually patients participating in clinical trials would not be eligible to participate with active brain metastases, thus the comparator data may not include such patients. Alternatively patients with brain metastases can be included in trials if local treatment is administered first to control the CNS disease. This in itself (localised brain treatment) could impact on patient outcome (favourably) and could therefore impact survival data in this cohort.
Is it appropriate to adjust the comparator data to adjust for NTRK fusions? If yes, are there any additional evidence sources which could be used to inform whether presence of an NTRK fusion is prognostic?	In our opinion there are limited data to determine the impact of the NTRK fusions on prognosis thus would be appropriate not to adjust for this factor.
Issue 13: Subsequent therapies	
What percentage of people, who receive entrectinib, would be expected to receive subsequent therapy following disease progression?	This will depend in which line of therapy entrectinib is positioned for each of the disease types.
What percentage of people, who receive established management, would be expected to receive subsequent therapy following disease progression?	This will depend on line of therapy under consideration and disease type.

Which subsequent therapies would be used and in what proportions? Please answer for following entrectinib and following established management.	As above
How long would these people be expected to be treated with subsequent therapy? Please answer for following entrectinib and following established management.	As above
Issue 14: Model structure	
What is the most appropriate model structure for this appraisal?	
Issue 15: Extrapolation of overall and progression-free survival	
Is the exponential or Weibull distribution most appropriate for extrapolating overall and progression-free survival?	
Issue 16: Drug wastage and source of treatment costs	
Is it appropriate to assume no drug wastage of entrectinib?	
Should drug acquisition costs for the comparator therapies be sourced from the British National Formulary (BNF) or the electronic market information tool (eMIT)?	
Issue 17: Administration costs and resource use	
Have administration costs been adequately captured in the company's model?	This seems reasonable as it is an oral therapy

Has resource use been adequately captured in the company's model?	Seems reasonable
Issue 18: Implementation and training costs	
What additional infrastructure and training requirements could be considered for this appraisal?	Nil for drug delivery however depending on the strategy for identification of NTRK fusions there will be infrastructure and training considerations to meet this.
Issue 19: Utility values	
Is the utility value estimate for the entrectinib progression-free health state collected in the STARTRK-2 trial appropriate?	
Is the utility value for the established management progression-free health state appropriate (0.73)?	
Should the utility value for the progression-free state be equal between the entrectinib and established management arms? If yes, what should the utility value be?	
Is it reasonable to assume that the utility value for the progressed disease health state be equal between the entrectinib and established management arms? If yes, is the value of 0.59 appropriate?	
Issue 20: Uncertainty around the cost-effectiveness results	
Is the uncertainty around the cost-effectiveness results captured appropriately in the company's model?	

Issue 21: End of life	
What is the life expectancy of the patient group receiving established management?	This will vary considerable depending on tumour type and the line of therapy
What is the extension to life of the patient group receiving entrectinib?	This cannot be said with certainty and the extension gained will be significantly impacted on by the underlying biology/rate of progression and the availability of subsequent therapies
Issue 22: Innovation	
Is entrectinib an innovative treatment?	Yes
Issue 23: Cancer Drugs Fund	
Does entrectinib meet the criteria for inclusion in the Cancer Drugs Fund?	Yes
What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.	Imperative to collect phase 4 data prospectively on both unrepresented tumour types and those that are represented but with only small numbers Number of patients screened to identify one TRK fusion in each disease type Prevalence of NTRK fusion across *all* disease types in the UK population. Number of cycles of entrectinib adminsitered Outcome of patients – RR, real world PFS and OS Registry for all patients with NTRK fusion (regardless of whether patients receive entrectinib) and analysis of outcomes on each treatment received

	Presence or absence of brain metastases

Single Technology Appraisal (STA)

Entrectinib for treating *NTRK* fusion-positive solid tumours

ERG addendum: review of company's response to technical engagement

Produced by **CRD and CHE Technology Assessment Group, University of York, Heslington, York YO10 5DD**

Date **14/10/2019**

Source of funding

This report was commissioned by the NIHR HTA Programme as project number NIHR128763

Declared competing interests of the authors

None

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in [redacted], all academic-in-confidence (AIC) data are highlighted in [redacted].

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

ALK	Anaplastic lymphoma kinase
BHM	Bayesian hierarchical model
CDF	Cancer Drugs Fund
CNS	Central nervous system
CrI	Credible interval
DNA	Deoxyribonucleic Acid
ERG	Evidence review group
ESMO	European Society for Medical Oncology
FISH	Fluorescence in situ hybridisation
FMI	Foundation Medicine Inc
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
IHC	Immunohistochemistry
IV	Intravenous
KM	Kaplan Meier
LYG	Life years gained
MASC	Mammary-analogue secretory cancer
NGS	Next generation sequencing
NMB	Net monetary benefit
NICE	National Institute for Health and Care Excellence
NSCLC	Non-small cell lung cancer
NTRK	Neurotrophic tyrosine receptor kinase
ORR	Objective response rate
OS	Overall survival
PFS	Progression free survival
QALY	Quality adjusted life-year
RNA	Ribonucleic acid
ROS1	Proto-oncogene tyrosine-protein kinase ROS
TA	Technology Appraisal
Trk	Tropomyosin receptor kinase
TTNT	Time to next treatment
WGS	Whole-genome sequencing
WTP	Willingness-to-pay

1 Overview

This addendum to the Evidence Review Group (ERG) report provides the ERG critique of the additional evidence provided by Roche in their response to the draft Technical Report for the appraisal of Entrectinib for treating *NTRK* fusion-positive solid tumours.

The draft Technical Report, outlined 23 key issues for consideration and provides the technical team's preliminary scientific judgement on each issue. The company's response to the draft Technical Report indicated that they accepted the technical team's preliminary judgement on **Issues 1** (prevalence of *NTRK* fusion), **8** (primary CNS and paediatric tumours), **13** (subsequent therapies), **16** (drug wastage and source of treatment costs), **19** (utility values) and **22** (innovation), so these are not discussed in this addendum. The company's responses to all the other issues are discussed Section 2.

The company also provided an appendix describing a review to identify new literature on the prognostic implications of *NTRK* gene fusions (**Issue 12**: Prognostic factors). As discussed in the ERG report, there is insufficient evidence to suggest that *NTRK* fusions are associated with a worse prognosis across all solid tumours compared with a non-*NTRK* population. As the literature review of *NTRK* fusion prognosis did not identify any additional evidence to that presented in the original company submission, this appendix is not looked at further. ERG comments on the company's assumptions can be found in Sections 2.1, 5.2.6.1 and 7 of the ERG report.

2 Description and critique of additional evidence

2.1 Treatment pathway and positioning (Issues 2 and 11)

The technical engagement report outlines a number of questions (**Issues 2** and **11**) concerning the positioning of entrectinib in the treatment pathway. These relate to the wording of the anticipated marketing authorisation, as well as related uncertainties regarding the generalisability of the integrated efficacy analysis to NHS practice. The anticipated marketing authorisation for entrectinib outlines that entrectinib may be used where [redacted] In the ERG's original critique, we highlighted that the term [redacted] is ambiguous and likely to be open to interpretation, leading to uncertainty in the positioning of entrectinib in the treatment pathway.

The company's response accepts that there is some uncertainty in the anticipated label and a degree of subjectivity in the definition of what constitutes [redacted]. The company, however, highlights the decision criteria detailed in the company submission which set out their expectation that comparator therapies will typically come under one of three therapeutic classes: chemotherapy, hormone therapy and best supportive care, and would exclude targeted therapies, immunotherapy, biologic therapy, non-palliative radiotherapy, and surgery with curative intent. The company

additionally provides further guidance on when they would anticipate that entrectinib would be used in practice based on further clinical expert feedback. This guidance suggests that the lack of a target biomarker is a likely factor in the decision to give entrectinib and that an alternative therapy with an objective response rate (ORR) of less than 20%, combined with progression free survival (PFS) of less than 4 months, could be an indicator of an unacceptable outcome.

In response to **Issue 11** regarding the appropriateness of the modelled comparators, the company further acknowledges the uncertainty in the positioning of entrectinib and highlights that their approach, based upon using multiple lines of therapy (2nd and 3rd), was designed to address the fact that entrectinib could potentially be used at multiple points in the pathway and may be dependent upon an individual clinician's interpretation of [redacted] therapy. The company then presents a scenario analysis based on NHS England's (NHSE) and NHS Improvement's (NHSI) response to the technical engagement report where comparator efficacy and costs were modified so as to exclude capecitabine for 2nd line breast cancer, FOLFIRI and irinotecan for 2nd line colorectal cancer, and everolimus for the treatment of neuroendocrine tumours. The results of these adjustments to the comparator arm resulted in an incremental cost-effectiveness ratio (ICER) of £47,894 per quality adjusted life-year (QALY), a decrease of £1,464 per QALY.

However, the ERG notes that for this scenario, the company removed all of the comparator drug costs and drug administration costs for breast cancer, colorectal cancer and neuroendocrine tumours. As NHSE and NHSI requested a scenario with the removal of the costs and outcomes of the four drugs outlined above and not the removal of all of the drug costs for the three tumour types identified, the ERG has provided result of the requested scenario analysis (shown in Table 1).

Table 1 Scenario analysis exploring NHSE and NHSI's preferred comparators

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£60,341	1.59	1.02	[redacted]	[redacted]	[redacted]	£49,294
Entrectinib	[redacted]	[redacted]	[redacted]				

The ERG notes NHSE and NHSI's responses to the technical engagement process and considers the company's new scenario a reasonable reflection of the position of entrectinib outlined therein. The ERG also notes that NHSE and NHSI reasoning regarding the positioning of entrectinib is typically grounded in the availability of NICE approved alternatives and the strength of the evidence base supporting entrectinib which is often weak. This leads NHSE and NHSI to conclude that, for a number of tumour types, entrectinib will be used later in the treatment pathway where alternative

treatments are fewer and tend to be less effective. With respect to this line of reasoning, the ERG notes that the continuing data collection programme outlined by the company and the potential placement of entrectinib within the Cancer Drugs Fund (CDF) are likely to mean that the evidence supporting the efficacy of entrectinib will evolve significantly. The availability of such improved efficacy data may therefore change the decision calculus regarding where entrectinib should be used in the treatment pathway.

Further to the above, as outlined in the ERG's original critique, the ERG considers that the positioning of entrectinib will be influenced significantly by the timing of *NTRK* testing, because clinician preference will be to use a targeted therapy where available. In this regard, the ERG notes NHSE's intention to move towards pan cancer genomic testing at the point of diagnosis. This is likely to have consequences for the placement of entrectinib in the treatment pathway. Such an expansion of molecular testing would likely mean clinicians are aware of patients' *NTRK* status earlier than the hierarchical testing approach proposed by the company. Therefore, contrary to the presented scenario based on NHSE and NHSI recommendations, entrectinib may be used earlier, rather than later, in the treatment pathway reflecting the preference for targeted therapies.

The ERG therefore maintains that there are substantive uncertainties regarding the positioning of entrectinib and while the ERG accepts that the revised scenario may represent the short-term positioning of entrectinib, the ERG highlights that the positioning of entrectinib is likely to change with time and will likely reflect the evolving clinical data as well as the availability of *NTRK* testing.

2.2 Generalisability of integrated efficacy analysis and comparator treatments (Issues 2 and 11)

The ERG critique of the company submission highlighted a number of important uncertainties relating to the generalisability of the population recruited to the studies and included in the integrated efficacy analysis, as well as concerns regarding the comparability of the data used to model comparator therapies. These uncertainties led the ERG to conclude that the modelled population is unlikely to be representative of the population who will receive entrectinib in NHS practice. Further, this also led the ERG to conclude that the population included in the comparator trials is unlikely to match the entrectinib efficacy population, notably due to the unknown prevalence of *NTRK* fusions in most of the comparator evidence, and mismatches in the lines of previous therapy.

As part of their response, the company provides new evidence to support the generalisability of the integrated efficacy analysis and its relevance to NHS practice. This new evidence is used to justify the company's decision not to present further economic analysis where the comparator data is matched to the position that entrectinib is given in the integrated efficacy analysis. This new evidence is derived from a

[redacted]

The ERG does not consider the evidence provided by the company to support such a conclusion and maintains that the number of previous treatments remains a major inconsistency between the modelled entrectinib population and both the likely treated population and modelled comparator arm of the model. The evidence provided to support the assumption of no prognostic effect of line of therapy is weak and based on a small number of patients. Importantly, there is a substantial risk that this comparison is confounded by other characteristics. For example, it is easy to imagine that patients with more aggressive tumours would be treated earlier in the pathway. The ERG also does not consider the company's conclusion of no difference to truly reflect the observed data. The Kaplan Meier (KM) data for both PFS and OS shows that observed outcomes for pre-treated patients are nearly always inferior to treatment naïve patients, although the ERG acknowledges that the observed difference is relatively modest. The ERG is also disappointed that this data was not incorporated into the economic analysis, as this would have allowed for a better understanding of the impact line of therapy has on both outcomes and estimated cost-effectiveness.

In addition to the above, the ERG also notes that a scenario analysis in which comparator therapies are matched to the position that entrectinib is given in the integrated analysis would have potentially altered model comparator costs as well as efficacy outcomes. Therefore the ERG does not consider the data presented as suitable justification for not implementing such a scenario.

2.3 *NTRK* fusion testing strategies (Issue 3)

There are a variety of testing strategies available for detecting *NTRK* fusions for individuals. The hierarchical approach recommended by the European Society for Medical Oncology (ESMO) proposes that testing for *NTRK* fusions should differ depending on the frequency of *NTRK* fusions in each tumour type, and current testing available. The use of first-line next generation sequencing (NGS), including whole genome sequencing or RNA-based NGS is an alternative approach to identifying *NTRK* fusions. After significant investment in genomic services across the country, and with the potential to be used to detect other oncogenic mutations, NHSE considers the use of NGS the most appropriate testing approach for detecting *NTRK* fusions. The company believes that due to the limited provision of NGS currently available on the NHS, immunohistochemistry (IHC) should be used as an interim test until genomic services are optimised for wide-scale genomic testing. The ERG

are concerned with the company's proposition of using interim IHC until NGS is available for a number of reasons.

First, the implementation of IHC for *NTRK* fusion testing would require a considerable investment into histopathology infrastructure, training and staffing, which would be disregarded after a relatively short time period when NGS is fully operational. Secondly, with the company's proposal of 'proactive entry' into the CDF, it is important that the testing strategy used to identify eligible patients within this setting reflects the future genomic testing that will be used in practice after the drug's period in the CDF. This is important as alternative strategies may impact upon the distribution of tumour types (due to the availability of testing) and may also impact on the number of false positives, with knock on consequences for efficacy of entrectinib. Finally, the total number of individuals that would require testing to identify those who are eligible for treatment is likely to differ between the two testing approaches, which will inevitably impact the investment required in laboratory infrastructure, additional staff and provision of training. Under the hierarchical testing approach, whereby IHC is followed by confirmatory NGS, it is likely that testing will be implemented upon eligibility for therapy and hence, fewer individuals will require testing. Under this approach, an estimated [redacted] individuals would require IHC and [redacted] would require NGS every year assuming testing is limited to sites known to express *NTRK* fusion (ERG report, Section 2.2.2.3), and will be higher still if genomic testing for *NTRK* fusions is adopted across all advanced and metastatic cancers.

The position of genomic testing is likely to depend on the strategy adopted. Under the hierarchical approach proposed by the company, it is implicitly assumed that testing is done upon eligibility for therapy, which may or may be at diagnosis depending upon the positing of entrectinib. In the context of a hierarchical approach this assumption of testing upon eligibility for therapy is reasonable and clinical advice received by the ERG suggested this matched the testing approach for other targeted therapies. If pan-genomic testing for all advanced or metastatic patients is implemented, as outlined by NHSE, this is unlikely to be at the point of eligibility for entrectinib, but rather at diagnosis. This has consequences for the costs of identifying a single patient with *NTRK* fusions because it means that where entrectinib is positioned as second or later line of therapy, not all *NTRK*-positive patients will receive entrectinib, as a proportion will die or become ineligible for treatment after earlier lines of therapy. The ERG have explored how this will affect the ICER (See Section 3.4.2)

2.4 *NTRK* Gene Fusion Testing Costs (Issue 4)

The company considers that all costs associated with testing should be excluded from the appraisal of entrectinib. The ERG disagrees with the company's position, and consider that the incremental costs of testing to identify *NTRK* fusions or Trk-proteins (in the case of IHC) should be included in the economic evaluation of entrectinib. The aim of the model is to estimate incremental costs of implementing entrectinib, which includes all additional costs associated with identifying *NTRK* fusion

positive patients. With the exception of whole-genome sequencing (WGS) and fluorescence in situ hybridisation (FISH) (currently used to detect *NTRK* fusions in mammary-analogue secretory cancer (MASC) and secretory breast carcinoma patients), which are currently reimbursed on the NHS, *NTRK* testing is not routinely carried out as part of NHS practice and therefore the implementation of *NTRK* testing for all other tumour types would represent additional testing undertaken by the NHS. The extent to which the cost of testing will affect the ICER will largely depend on the testing strategy utilised.

2.4.1 Potential testing strategies

Hierarchical Approach

Under the hierarchical approach (see Section 2.3), separate IHC assays are required to detect each type of protein i.e. IHC tests are protein specific and cannot be used to detect multiple mutations/fusions. In estimating the incremental cost associated with *NTRK* fusion testing, testing for other targets should not be included in the comparator arm of the model as these IHC tests will not be displaced by the introduction of *NTRK* fusion testing. Therefore, the total cost of IHC should be included for all tumour types in the analysis. As RNA-based NGS, is the gold standard of testing, the ERG believe that this should be used to confirm the presence of an *NTRK* fusion. The total cost of RNA-based NGS that would be required for each tumour type needs to be accounted for as, to the ERG's knowledge, RNA-based NGS is not currently reimbursed by the NHS, with the exception of NSCLC.

Up-front Genomic Testing

If, as NHSE propose, NGS is employed to detect *NTRK* fusions across tumour types, the extent to which the provision of *NTRK* fusion testing will displace current services is contingent on whether RNA-based NGS or DNA-based NGS is utilised.

First-line RNA-based NGS

An alternative approach would be to employ RNA-based NGS for the majority of patients with advanced and metastatic cancer. For patients currently receiving WGS, RNA-based NGS would be used to confirm Trk-protein expression. Therefore, for these patients, the incremental costs of RNA-based NGS would be included for patients currently receiving WGS who are *NTRK* fusion positive. For the tumour types where no multi-panel genomic testing is currently available, the total cost of RNA-based NGS should be used. The ERG are uncertain about the extent to which RNA-based NGS could displace current DNA-based NGS panels. Clinical advisors informed the ERG that RNA-based NGS cannot be used to detect all targeted genes currently tested for on DNA-based NGS panels. Therefore, the ERG consider that some incremental cost should be applied for the tumours where DNA-based NGS is available. To reflect the development of genomic testing, the ERG considers that the use of FISH to detect *NTRK* fusions in patients with MASC and Secretory Breast will be displaced

by RNA-based NGS. As RNA-based NGS is currently reimbursed for NSCLC, there are no incremental costs for detecting NTRK fusions in this tumour type.

DNA-based NGS with confirmatory RNA-based NGS

As described in Section 2.3, *NTRK* fusions could be identified using DNA-based NGS, with RNA-based NGS used to confirm that *NTRK* positive cases express the Trk-protein. DNA-based NGS is currently available for several tumour types. Patients with soft-tissue sarcoma or paediatric cancer are eligible or WGS. The ERG assume no incremental costs for identifying *NTRK* fusions for WGS. Furthermore, multi-panel testing is currently available for patients with melanoma, lung, colorectal, thyroid, breast and ovarian cancers. As these are currently reimbursed by the NHS, a nominal, incremental cost of adding *NTRK1/2/3* target markers would be applied for these tumour types. For the tumours where no genomic testing is currently available (such as pancreatic, prostate and gliomas), the total cost of a DNA-multi panel test is required.

As an *NTRK* fusion may be expressed, but not be transcribed into an oncogenic protein, RNA-based NGS should be used to validate the presence of Trk-protein for all positive tests using DNA-based NGS. As this is currently not available on the NHS, the total cost of confirmatory RNA-based NGS is required in the model.

Table 2 Summary of the ERG's position on incremental testing costs for each testing approach

Current Testing Availability	Hierarchical Approach	DNA-based NGS with confirmatory RNA-based NGS	Up-front RNA-based NGS
FISH	• No incremental cost	• No incremental cost	• Incremental cost of displacing FISH with RNA-based NGS
WGS	• No incremental cost of WGS. • Total cost of confirmatory RNA-based NGS for *NTRK* positive WGS.	• No incremental cost of WGS. • Total cost of confirmatory RNA-based NGS for *NTRK* positive WGS.	• No incremental cost of WGS. • Total cost of confirmatory RNA-based NGS for *NTRK* positive WGS.
Multi-Panel DNA testing	• Total cost of IHC • Total cost of confirmatory RNA-based NGS for patients with Trk expression using IHC.	• No incremental cost of DNA-based NGS. • Total cost of confirmatory RNA-based NGS for *NTRK* positive DNA-based NGS	• Incremental cost of displacing DNA-based NGS with RNA-based NGS
RNA-based NGS for NSCLC	• No incremental cost	• No incremental cost	• No incremental cost
No Genomic Testing	• Total cost of IHC • Total cost of confirmatory RNA-based NGS for patients with Trk expression using IHC	• Total cost of DNA-based NGS. • Total cost of confirmatory RNA-based NGS for *NTRK* positive DNA-based NGS	• Total cost of RNA-based NGS.

2.4.2 Testing Costs for Unrepresented Tumour Types

The ERG considers that it is important to include the costs of testing for the tumour types that are unrepresented in the integrated efficacy evaluable population so that the incremental costs most accurately reflect the anticipated histology-independent marketing authorisation. There is significant heterogeneity in the costs of testing across the unrepresented tumour types, which is dependent on the prevalence of *NTRK* fusions and the current testing availability. For example, in patients with renal cell carcinoma, based on a prevalence of [redacted]% (Foundation Medicine Inc (FMI) data), [redacted] individuals would require screening in order to identify one *NTRK*-fusion positive patient which will have a significant impact on overall costs.

2.4.3 Proportion of testing costs that should be attributed to entrectinib

The ERG has concerns regarding the appropriateness of scenarios presented by the company in which only a proportion of the testing costs are attributed to entrectinib. The ERG considers this to be conceptually flawed and not to reflect the decision problem being addressed. The ERG considers that all incremental costs associated with testing should be accounted for in the model as these represent significant costs associated with the implementation of entrectinib and also account for an important part of the heterogeneity in cost-effectiveness of entrectinib across tumour types. The ERG acknowledges that the implementation of *NTRK* testing may, for some tumour types, displace current testing. When that is the case, the ERG agrees that these costs should be taken into account such that testing costs truly reflects only the incremental costs of *NTRK*-testing. The ERG, however, does not consider it appropriate to apportion only part of the test costs associated with *NTRK* testing to entrectinib as this does not reflect the choice faced by the NHS; *NTRK* testing should only be implemented where it generates sufficient benefits to justify the costs of testing. The ERG notes the NHSE plan to move to towards pan-cancer genomic testing, and the company's interpretation that this implies that testing costs should be subsumed within the NHS budget and therefore removed from the economic analysis. The ERG considers that, any decision by NHSE to implement genomic testing should in theory be subject to the same considerations of cost-effectiveness as any other technology. Further, the ERG accepts that the introduction of pan cancer genomic testing represents a public good, which may generate positive externalities, particular as new targeted therapies become available for other genetic targets. However, currently *NTRK* is the only potential pan-cancer target and as such, any benefits of introducing genomic testing are limited to those generated by entrectinib until further therapies targeting different actionable mutations become available. To consider the potential for benefits from future target therapies within the current economic analysis the ERG implements a threshold analysis in which it considers the magnitude of benefits required for pan cancer genomic testing to be cost-effective and interprets this in the context of the prevalence of alternative genetic targets across the whole advanced and metastatic cancer population (see Section 3.5).

2.5 Identification of *NTRK* gene fusions – diagnostic accuracy (Issue 5)

Due to the recent development of Trk-inhibitors, there is limited, substantial evidence regarding the diagnostic accuracy of NGS for *NTRK* fusions. Furthermore, the analytical validity of NGS is likely to differ between each *NTRK* gene, fusion partner and tumour type. This leads to considerable uncertainty in the specificity and sensitivity of NGS for detecting *NTRK* fusions.

DNA-based NGS, in the form of multi-panel targeted assays or WGS, analyses genomic DNA from a tumour sample to identify mutations in multiple genes simultaneously.[1] The company present the diagnostic accuracy estimates for the MSK-IMPACT DNA-based NGS assay, which shows 81.1% sensitivity (Solomon et al[2]). The sensitivity of DNA-based NGS – that is, the ability of the test to detect individuals with an *NTRK* fusion – is moderate, especially in *NTRK2/3* fusions which have a sensitivity of 0% (0/4 *NTRK2* fusions detected) and 76.9% (30/39 *NTRK3* fusions detected). *NTRK2/3* fusion breakpoints (i.e. the location on the chromosome where DNA is mutated) often occurs in the intronic sequence of DNA, which are not covered by some DNA-based NGS assays currently available (e.g. MSK IMPACT and FoundationONE CDx). The utilisation of DNA-based NGS to detect individuals with an *NTRK* fusion may result in large numbers of false negatives. Thus, incorrectly identified *NTRK* fusion positive patients will miss out on entrectinib due to the poor sensitivity of the test.

The specificity of DNA-based NGS is high, with 99.9% of individuals being correctly identified as not having an *NTRK* gene rearrangement.[1] However, DNA-based NGS may detect DNA-level fusions that may not result in Trk protein expression, hence may not be an oncogenic driver. This is especially the case for rare or novel *NTRK* fusions.[3]

RNA-based NGS is the 'gold standard' of fusion testing. RNA-based NGS is a more thorough NGS strategy to use in order to detect *NTRK* fusions. In addition to its ability to confirm the presence of a Trk protein, there is no need to have prior knowledge of the fusion partners, or chromosomal breakpoints. As *NTRK* fusion testing is relatively new and there is likely to be novel and rare fusion partners, RNA-based NGS may be the most appropriate strategy in order to identify novel fusions and hence, correctly identify more individuals eligible for entrectinib. Therefore, the ERG considers that if DNA-based NGS is used, RNA-based NGS is required to confirm a transcribed *NTRK* fusion, and consequently Trk protein expression. This will limit the impact of false positive results and in turn, may improve the efficacy of entrectinib.

The diagnostic accuracy of RNA-based NGS is, however, somewhat uncertain. The company presents evidence supporting accuracy of Oncomine Focus Assay,[4] which has reported 100% sensitivity and specificity.[4] The cited study, however, has only a very limited sample of fusion positive patients. Furthermore, is important to note that the reliability of RNA-based tests is highly dependent upon the

quality of the RNA sample available and degradation in sample quality can result in increased numbers of false negatives.[3] This is important to consider in the context of the genomic strategy proposed by NHSE and NHSI, as there is a risk that the increased demand for tests may increase the likelihood of errors, which could incur additional time and resource without added benefit.

Alternative testing strategies will result in different numbers of false positives depending on different tests' sensitivity and the specificity. The hierarchical testing strategy will be associated with significant numbers of false-positive individuals requiring RNA-based NGS confirmation following the IHC pre-screen. As previously reported, [redacted] would require confirmatory RNA-based NGS following IHC screening every year. Using the ERG's estimate of the annual eligible population [redacted] only [redacted] of individuals who receive confirmatory NGS would be eligible for entrectinib therapy. However, under the hierarchical testing approach, the number of false positives who would go on to receive entrectinib treatment would be low.

In single-test strategies for *NTRK* fusion detection, the lack of confirmatory testing may increase the number of false-positive patients who will go on to receive entrectinib. As discussed above, the limitations of DNA-based NGS, and the potential for the detection of DNA-level fusions that do not result in Trk protein expression may result in false positives, especially in *NTRK2/3* fusions. RNA-based tests, however, are not limited in this way and therefore may confer an advantage over a DNA-based NGS strategy. The evidence supporting the analytical validity of RNA based tests is, however, limited to small samples, and therefore there is uncertainty surrounding the proportion of false positives that would occur using an RNA-based NGS approach.

Moreover, the number of false positive individuals is likely to differ between tumour types, depending on the prevalence of *NTRK* fusions. The predictive value of a test to correctly identify those with the fusion is dependent on the prevalence of the mutation; for tumours where *NTRK* fusions are rare, the likelihood of false positives is higher.

2.6 Distribution of tumour types (Issue 6)

The ERG's critique of the company's original submission outlined concerns that the distribution of tumour types represented in the integrated efficacy analysis is not representative of the distribution of tumour types that will be treated in practice. The ERG's report further states that the cost-effectiveness of entrectinib is likely to be influenced by the distribution of patients across tumour types because prognosis and costs vary substantially across tumour types. As part of the exploratory work undertaken by the ERG, the FMI database is put forward as a more plausible distribution of patients than the distribution integrated efficacy analysis. These concerns prompted the technical team to reflect on what is the most appropriate distribution of patients across tumour types. The company's response asserts that there is uncertainty in the distribution of tumour types included in this appraisal,

while also asserting that the proportion of tumour types included in the trial population may reasonably reflect the proportions seen in clinical practice. The company further states that the likely distribution may only be definitively defined following comprehensive testing across all cancer types.

The ERG disputes that the trial can be considered a representative sample and true reflection of the likely eligible population that will be treated in the NHS – the trial population, though recruited through a testing programme involving 10,000 patients, is significantly smaller than the FMI data set which is based on ~166,000 samples. Further, the reported incidence of MASC patients in the trial is clearly higher than what would be observed in practice, given the rarity of this tumour type in the UK. The company also acknowledges, in their response to **Issue 1**, that the FMI dataset represents the single largest and therefore most robust study of *NTRK* gene fusion prevalence.

In addition to the above, the ERG also questions whether further data collection, through placement of entrectinib in the CDF will allow more accurate estimates of prevalence to be determined. While such an exercise would generate data specific to the UK setting, it is doubtful that any data collection exercise could collect a larger sample than currently in the FMI database, in the limited period entrectinib would be in the CDF. The annual incidence of advanced and metastatic cancer across all tumour in England is approximately 100,000 patients meaning that nearly all advanced and metastatic patients would need to be tested to achieve comparable numbers.

As part of their response, the company further critiques the ERG's scenario analysis in which comparator outcomes as well as testing costs are reweighted to align with the distribution reported in the FMI database, noting that the ERG analysis did not simultaneously reweight outcomes in the entrectinib arm of the model due to the lack of patient level data. To address this issue the company provide revised KM curves that have been reweighted to the distribution of patients in the FMI database. The company notes that this reweighting results in only a marginal difference in entrectinib outcomes, while also noting that this approach is flawed due to the limited data sources.

The ERG appreciates the company's attempt to respond to this request, and agrees this reweighting appears to have limited impact on predicted PFS and OS outcomes lending credence to the scenario analysis presented by ERG. The ERG, however, considers that it would have been helpful to have implemented this analysis in the economic model to demonstrate the impact on the ICER even if that impact is only modest. The ERG is also not fully clear how the KM data was reweighted as this is not described in the company's response leading to a degree of uncertainty in the validity of the provided KM curves.

The ERG also notes the company's clarification of the reason this scenario leads to an increase in the ICER. The ERG can confirm that the changes to the ICER are driven primarily by additional WGS

costs, which are not modified in the entrectinib arm and therefore arguably unfairly inflate the ICER. However, what this analysis illustrates is that QALYs and costs are determined in part by the distribution of tumour types used. To further explore the impact of the distribution on the estimated cost-effectiveness the ERG carries further analysis in Section 3.

2.7 Unrepresented tumour types (Issue 7)

The technical team noted that the results from the entrectinib studies only included 13 different tumour sites harbouring *NTRK* fusions, out of all the possible tumour types that may be covered by the proposed histology-independent licensing indication. The ERG report identified at least 30 types of tumour that have been shown to harbour *NTRK* fusions and clinical advice has suggested there could be even more.

The company stated that the waterfall plots demonstrate compelling efficacy of entrectinib across tumour types. Waterfall plots have been criticised as being misleading and poor representations of response in oncology trials.[5] In addition, waterfall plots only represent the patients, and therefore the tumour sites, included in the studies and provide absolutely no information on tumours that were not included. The ERG also notes that the small number of patients with some tumour types does not allow accurate assessment of the expected response in those tumour types. The ERG conducted a formal analysis which pooled the response rates across tumour types, whilst allowing for potential heterogeneity in the response rates in different tumour types (ERG Report Section 4.3.1). Heterogeneity was found to be moderate with mean response rates across tumour sites varying from [redacted] to [redacted]%. Note that these values are less extreme than the observed range of response rates, which varied between [redacted] and [redacted]% due to the borrowing of information across tumours allowed by the model (ERG report, Table 26).

The company further argue that *NTRK* fusions are thought to play a similar role across tumours types. However, as noted above, whilst a meaningful level of response may be expected across many tumours, the current evidence of response in observed tumours is limited to small numbers of patients, and there is currently no evidence that this is the case for tumours that have not yet been included in trials. The examples of EGFR and anaplastic lymphoma kinase (ALK) inhibitors cited by the company, only concern a single histology (lung cancer and non-small cell lung cancer (NSCLC), respectively) and therefore are not relevant to the assumption that a targeted drug should work similarly in patients across multiple histologies. In fact, the ERG report (Section 4.2.1) cited three examples where different responses to targeted therapies were observed across different tumour types:

1. A trial of vemurafenib in 122 patients with *BRAF* V600–mutated cancers across multiple tumour types (including colorectal cancer, NSCLC, Erdheim–Chester disease and Langerhans'-cell histiocytosis, primary brain tumours, cholangiocarcinoma, anaplastic

thyroid cancer) found evidence of response in some tumour types including NSCLC and Erdheim–Chester disease and Langerhans'-cell histiocytosis, but not in colorectal cancer.[6]

2. A trial of imatinib, a tyrosine kinase inhibitor, that included 196 patients across 40 different tumour subtypes, found evidence of activity of imatinib in only five malignancies.[7]
3. A basket trial of imatinib that included 185 patients with 10 histologic subtypes of advanced sarcoma concluded that imatinib was not an active agent in these subtypes, although it had previously shown effectiveness in gastrointestinal stromal tumour, another subtype of soft tissue sarcoma.[8]

The ERG also notes that there is evidence that larotrectinib, a similar, recently approved, histology-independent Trk-inhibitor for patients with *NTRK* fusions, displayed very different response rates across the tumour sites included in the studies submitted for regulatory approval. The Bayesian hierarchical model (BHM) framework described in the ERG report (ERG report, Section 4.3.1) was used to estimate the heterogeneity in response across the different tumour types included in the published larotrectinib studies and the probability of response for each tumour type. The observed number of patients with response and the number of patients with each of the tumour types included in the larotrectinib efficacy evaluable data set[9] are shown in Table 3.

Table 3 Response by tumour type for patients in the published larotrectinib studies.

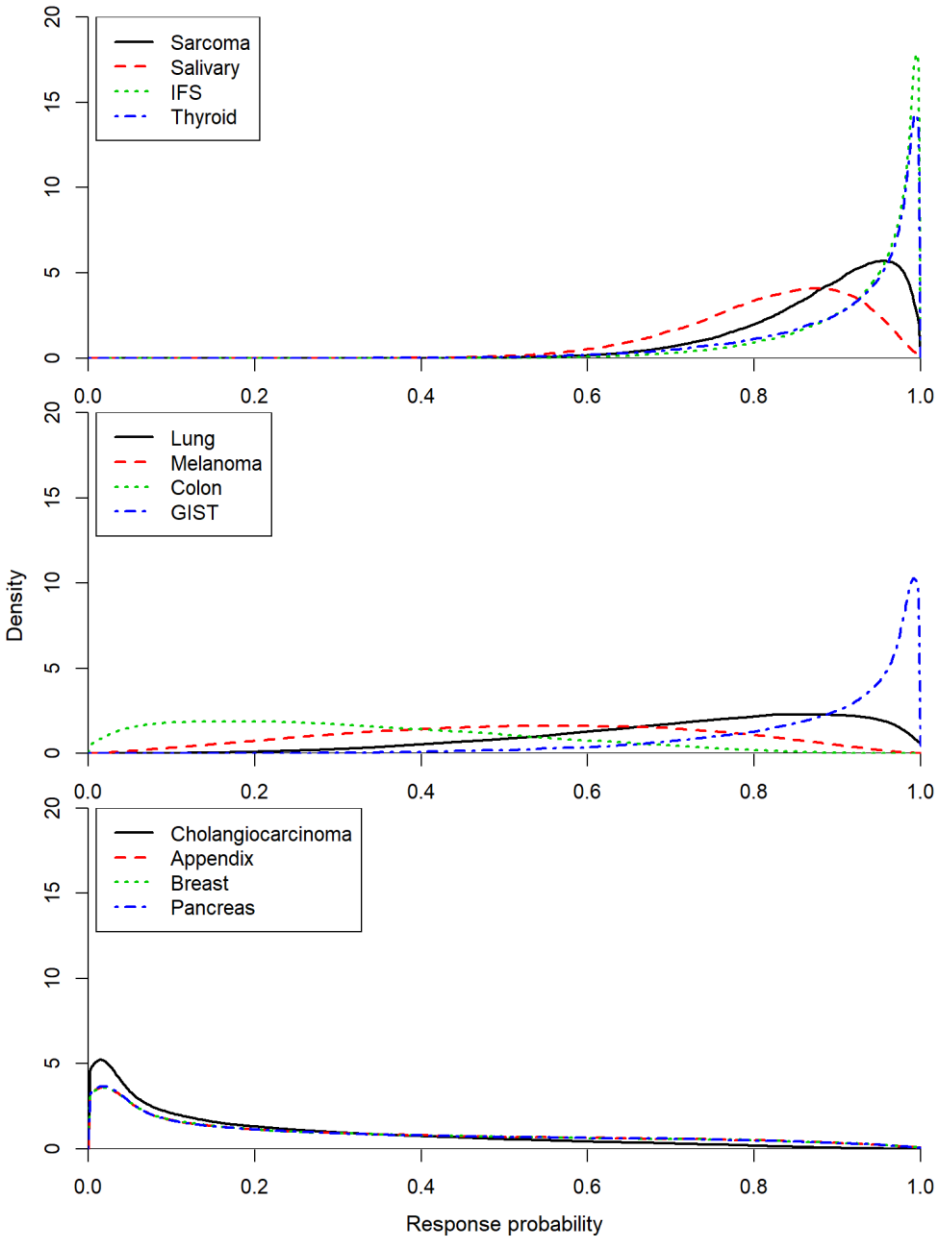
Tumour ID	Tumour Type	Patients (x)	Responders (n)	Observed response (%)	Estimated mean response based on BHM (%)
1	Soft tissue sarcoma	11	2	91%	88.1%
2	Salivary gland	12	2	83%	81.8%
3	IFS	7	9	100%	93.3%
4	Thyroid	5	2	100%	91.6%
5	Lung	4	6	75%	72.6%
6	Melanoma	4	5	50%	52.5%
7	Colon	4	1	25%	32.0%
8	GIST	3	6	100%	88.3%
9	Cholangiocarcinoma	2	0	0%	21.0%
10	Appendix	1	0	0%	30.0%
11	Breast	1	0	0%	30.0%
12	Pancreas	1	0	0%	29.8%
	Total	55	41	74.5%	60.9%

IFS, infantile fibrosarcoma; GIST, gastrointestinal stromal tumours

The BHM estimated a large between-tumour heterogeneity (posterior median 2.82, on the log-odds scale, 95% credible interval (CrI) 0.92 to 4.83), which suggests considerable variability in response

rates across tumour types, even though they are all *NTRK* fusion positive. Observed response rates range from 0 to 100% and BHM estimated response rates range from 21 to 93% (Table 3). Figure 1 shows the posterior distributions of the probabilities of response for each of the tumour types included in the published larotrectinib data. These suggest different probabilities of response across tumour types, with distributions for some tumour types supporting only large response rates and others supporting only small response rates. This further illustrates the lack of support for an assumption of homogeneity of response across different tumour types, even with the same target mutation. In addition, data presented to the FDA as part of an NDA Multidisciplinary Review and Evaluation of larotrectinib in patients with *NTRK*-gene fusions,[9] suggests that patients with *NTRK2* gene fusions had a lower overall response rate than those with *NTRK1* and *3* gene fusions, which may suggest differential response to Trk-inhibitors in this population.[9]

Figure 1 Posterior distribution for the probabilities of response in each tumours type for patients in the published larotrectinib studies.



The ERG also notes that the fact that different agencies have licensed entrectinib for a tumour agnostic indication, does not necessarily imply that entrectinib is equally effective, or cost-effective, across all tumours, only that those agencies have considered the level of uncertainty around the risk:benefit profile of entrectinib to be acceptable for licensing.

In summary, whilst the observed and estimated response rates on entrectinib may be considered promising across the tumour types included in the studies, there is no evidence that they are identical or that similar response rates should be expected in unrepresented tumours. Direct evidence of

effectiveness in unrepresented tumour types is not available and it is still uncertain how much can be inferred from the evidence presented on the tumour types included in entrectinib studies.

Therefore, the ERG maintains that there is considerable uncertainty in response rates for included tumour types, which implies even greater uncertainty in the response rates that may be observed in the unrepresented tumours. The ERG maintains that, given current evidence, the best estimate of response rates in unrepresented tumour types is the predictive probability estimated by the BHM: mean 57% with 95% CrI 18% to 89%. This has important implications when considering the cost-effectiveness of entrectinib across all tumours covered by the proposed licensing indication.

2.8 Heterogeneity of response across different tumour types (Issue 9)

The technical team asked whether the assumption of a uniform level of response to entrectinib across the different tumour sites was reasonable and whether the ERG's estimates of response across tumours (including unrepresented tumours) were appropriate. The company notes that current evidence for entrectinib and the literature on the role of *NTRK* fusions across tumour types supports the assumption that entrectinib is active across the observed tumour types and as yet unobserved tumour types. They then proceed to state that the assumption of a uniform level of response is also therefore reasonable. The ERG agrees that current evidence suggests a promising level of response to entrectinib across the tumour types included in the studies. However, evidence showing that the response is uniform, or that response rates would be as the same in tumour types not included in the entrectinib studies but covered by the proposed licensing indication has not been presented (see also response to **Issue 7**, Section 2.7).

The ERG does not agree with the company's view that the data are not robust enough to estimate the between-tumour heterogeneity because

1. Estimation of between-tumour heterogeneity depends mostly on the number of subgroups (i.e. included tumour types), and not so much on the number of patients in each subgroup.[10-12] Since there are 13 tumour sub-types included in the analysis we believe that there is enough information to estimate the between-tumour heterogeneity. This is further demonstrated by the comparison of the prior and posterior distributions for the between-study heterogeneity and the sensitivity analyses to different prior distributions conducted by the ERG (ERG report Figure 8 and Appendix E).
2. The company considered visual impressions from a waterfall plot and literature on single tumour sites as sufficient evidence to demonstrate homogeneity of response across tumour types (company's response to **Issue 7**). The ERG consider these to provide much weaker evidence than a formal analysis of heterogeneity based on the available data.

As stated in the ERG report and in response to **Issue 7** (Section 2.7), there was evidence of moderate heterogeneity in response rates across the included tumour types and although there is some uncertainty in the level of variability of response, this was well estimated by the data (ERG report Figure 8 and Appendix E).

In additional clarification, the company stated that they *"investigated the possibility of fitting an exponential frailty model to the PFS data including tumour type as the clustering term with a gamma frailty, but given low amounts of data the model did not converge"*. The ERG accepts that there is not enough evidence to fit such a model, for the same reason that it would not be meaningful to fit a BHM to the survival data by tumour type, namely due to small number of patients per tumour type.

The ERG believes that, in the absence of compelling evidence of a common tumour response to entrectinib, the default assumption should be that response rates differ across tumours, until it can be shown that they are sufficiently homogeneous to pool, rather than assuming they are homogeneous until they can be convincingly proved to be heterogeneous, as suggested by the company. The ERG's *a priori* assumption of heterogeneity is also supported by examples in the literature (Section 2.7) where variability of response across different tumour types was observed: vemurafenib in patients with BRAF V600–mutated cancers[6], imatinib[7, 8] and larotrectinib (Table 3, Figure 1). The ERG notes however, that although there is no evidence to support the assumption of a common response to entrectinib across different tumour types, the estimated heterogeneity is only moderate and a meaningful response is expected, on average, across all included tumour types.

The company states that the predictive interval for a response in an unobserved tumour type is wide and questions its usefulness in decision-making. However, the ERG maintains that this is currently the best estimate of the plausible response rates in unrepresented tumour types, given the available evidence. The company's proposal that response in unrepresented tumour types should be assumed to be the same as the average response observed across the included tumours is based on no evidence at all. The predictive distribution obtained by the ERG suggests that it is likely that response will be promising for unrepresented tumour types (ERG report, Table 25 and Figure 9) whilst appropriately characterising the uncertainty in this potential response. The predictive interval is useful for decision-making because it appropriately characterises the current uncertainty in response rates in tumours that have not yet been studied but which are covered by the proposed licensing indication. Decision-makers should be aware that a wide range of response rates are possible, and acknowledge the potential consequences in their decision. Whilst further data collection may reduce this uncertainty, deciding to collect more data is itself a decision that has consequences, and therefore a full characterisation of the uncertainty at the point that decision is made is valuable.

As noted in our response to **Issue 7**, the expectation of a promising response rate across tumour types does not automatically imply that entrectinib is more effective or cost-effective than current available therapies across tumour types. The ERG notes that the considerable uncertainty in effectiveness and cost-effectiveness of entrectinib in the tumour types represented in the studies implies an even greater uncertainty in effectiveness and cost-effectiveness for the unrepresented tumours. The fact that the time-to-event data is not yet robust enough to explore variability in PFS and OS across the different tumour types or to predict PFS and OS in the unrepresented tumour types, is a further key area of uncertainty. Given this lack of robust data, the ERG presents the response-based model as a plausible alternative model structure which, unlike the current partition survival approach adopted by the company, is able to model heterogeneity in outcomes amongst both observed as well as unobserved tumour types.

Analyses of the heterogeneity in the probability of being progression free at key landmark points, e.g. 6 and 12 months, and similarly for the probability of survival at key time points, would allow a comparison of the rates of PFS and OS at these time points across tumour types, which could provide valuable information on the heterogeneity of survival outcomes across tumour types.

2.9 Constructing a comparator arm (Issue 10)

The technical team asked whether the company's comparator arm was suitable for decision making and whether the ERG's use of non-responders to form the comparator arm, was appropriate. We comment on the company's response to these two issues in Sections 2.9.1 and 2.9.2, respectively.

2.9.1 Suitability of the company's submitted comparator arm

The company stated that the relatively large prevalence of CNS metastases and presence of *NTRK* fusions may worsen patient outcomes compared to patients with no brain metastases or oncogenic driver mutations. As discussed in the ERG report (ERG report Section 3.3), the prevalence of CNS metastases in the original comparator data, which was made up of an average of the median PFS and OS in previous NICE Technology Appraisals (TAs) of various tumours, is largely uncertain. Most comparator trials did not report baseline prevalence of CNS metastases or whether CNS metastases were excluded. Therefore the extent to which the relatively high prevalence of CNS metastases in the entrectinib trial evidence may have confounded analyses against comparator data is unknown. The naïve unadjusted comparisons presented by the company also do not account for other potentially important prognostic factors, such as age, performance status, or any other tumour mutations. Therefore it is not possible to predict the direction or magnitude of confounding associated with the naïve efficacy comparisons across tumour types included in the company submission. Overall, the ERG believes that the comparator data used to inform the company model is highly unreliable and may not be suitable for decision making.

To validate the originally submitted comparator arm survival estimates, the company used an intra-patient analysis[13] implemented in a deterministic fashion but did not attempt to fit a survival curve to the data. In this exploratory analysis, the company used [redacted]. The ERG considers this a promising approach, which may produce a more reliable estimate of mean/median PFS in the comparator arm than the original modelled comparator, or the ERG's non-responder analysis. A particular advantage of this approach over the non-responder model as implemented in the ERG's report is that it implicitly accounts for differences in patient characteristics as patients are acting as their own control. This approach, however, makes a number of assumptions:[13]

1. That the benefit of treatment is in delaying disease progression, with post-progression survival equal between entrectinib and the treatments in the previous line of therapy.
2. That survival risk is treatment independent.
3. That re-treatment with drugs that have failed on a previous line of therapy would yield the same time to further disease progression as at the previous line of therapy.
4. That the risk of death in the pre-progression state is the same across the lines of therapy, and negligible.

Assumption 1 may be reasonable given the limited availability of therapy post-progression and the limited evidence on post progression therapy in the integrated efficacy analysis. The assumption of equal survival risk (Assumption 2) may also be reasonable given that entrectinib has a similar adverse event profile as the treatments used in previous lines of therapy, although this may result in a conservative estimate of entrectinib benefit if its adverse event profile is better than therapies used in the previous line. However, Assumptions 3 and 4 are very strong and hard to verify and are a limitation of this approach.

Further, while [redacted], it is unclear whether extrapolation of a curve fitted to the intra-patient comparator would affect the ICER (median PFS is not the only driver of cost-effectiveness results). Further, the ERG notes that one of the ERG's criticisms of the estimates of PFS and OS obtained in the original modelled comparator arm was that their uncertainty was not properly quantified. The company have still not adequately described the uncertainty in the comparator PFS estimates.

2.9.2 Suitability of using non-responders as a proxy for the comparator arm

The company stated that they consider the use of non-responders as a proxy comparator arm as "deeply flawed". The ERG agrees that an intra-patient analysis may provide a better basis for a comparator arm (see Section 2.9.1) and notes that median PFS under the response model is substantially longer than predicted by the intra-patient analysis (8.3 vs 4.6 months). However, the ERG used the non-responders as proxy for patients not receiving an active therapy for two key reasons:

1. This was the only breakdown of the data available to the ERG. Data allowing an intra-patient analysis based on response to a previous line of therapy was not made available.
2. Separate modelling of PFS and OS for non-responders and responders allows linking of the probabilities of response for each tumour type to PFS and OS, making it is easier to generate ICERs specific to each tumour type (ERG report Section 6.5)

The company noted several reasons why using non-responders to emulate the comparator arm may be flawed. Whilst the ERG agrees that the approach has limitations, we note that some of the flaws identified by the company cannot be surpassed by any other currently available method, and also apply to the intra-patient analysis described in Section 2.9.1. In addition, the ERG notes that the "usual model" critiqued by Anderson *et al* (1983)[14] was not the model used by the ERG, neither did the ERG perform hypothesis tests or derive confidence intervals which is the focus of the critique in that paper.

The company notes that only a small number of patients is included in the responder analysis. The ERG agrees with this point but notes that (i) the company's alternative approach (discussed in Section 2.9.1) is based on [redacted] patients, which is less than [redacted] and (ii) the small number of patients included in the entrectinib studies is a concern throughout this appraisal. The company also notes issues with how response is measured. However, 'response' is used throughout the submission to justify effectiveness of entrectinib across tumour types and therefore the ERG considers it a good proxy for treatment activity. The ERG also notes a similar issue in the intra-patient analysis, namely that [redacted]. The company further notes that non-responding patients may still benefit from entrectinib or may have a reduction in tumour size which does not reach the threshold for response but still experience a benefit. The ERG notes that patients may also derive some benefit from the comparators as some are active therapies. As noted above, the only data available to the ERG were the KM curves for responders and non-responders. Therefore, PFS and OS subdivided by patients with a best response of progressive disease as a comparator arm was not available. A similar concern applies to the intra-patient analysis (Section 2.9.1), which is likely to over-estimate performance in the comparator arm since patients are at an earlier line of therapy (i.e. likely to be fitter), benefiting from

appropriate active treatments at the time of assessment and would be unlikely to benefit in the same from being given the same treatment post-progression.[13] For these reasons both the company and ERG's estimated comparator arms are likely to provide a conservative estimate of the effect of entrectinib.[13, 15] The company further note that the heterogeneity in prognostic factors in the patient population and the lack of adjustment for covariates may result in an unreliable comparison. The ERG agrees that heterogeneity in patients is a concern and that this may affect response. An adjusted analysis would have been preferable if the original data were available.[15]

The company makes the fundamental point of data immaturity and limited follow-up which limits the analysis. The ERG completely agrees with this point and notes that immaturity of survival data and the small number of patients included in the entrectinib studies is a key area of uncertainty throughout this appraisal.

Overall, the ERG agrees that the intra-patient analysis may be preferable to using the non-responders without further adjustment for covariates, and notes the more favourable estimates of median PFS when using non-responders (8.3 vs 4.6 months). However, the intra-patient analysis is not without limitations (see Section 2.9.1) and does not allow the incorporation of different potential responses to entrectinib depending on tumour type, which was a key aim of the ERG's exploratory analysis (ERG report Section 6.5).

2.10 Prognostic factors (Issue 12)

The ERG considers the company's attempts to adjust for differences in the prevalence of CNS metastases between the entrectinib and comparator arms a reasonable scenario. However, care should be taken not to over interpret the results. The company report that the prevalence of CNS in the integrated analysis was significantly higher than the highest proportion reported in the comparator TAs (20.4 vs 14%). As discussed in Section 2.9.2, the ERG considers this interpretation potentially misleading, as most of the TAs from which comparators were selected did not report the proportion of patients with CNS metastases. It is therefore not certain what the relative proportion of CNS patients was in the comparator studies. The ERG also questions the reasonableness of selectively adjusting for only one of the potentially important prognostic difference, while ignoring the existence of other differences between the trial and comparator datasets, which may act in other directions upon the results.

In addition to the adjustment for CNS metastases, the company reiterate the position stated in their original submission that *NTRK* gene fusions are in themselves linked to a worse prognosis than in patients with similar tumours not harbouring this alteration. As discussed in the ERG report (ERG report, Sections 1.5 and 5.2.6.1), there are a number of reasons why alternative assumptions concerning the prognosis of patients with *NTRK* fusions may apply, and thus the focus on a scenario

which assumes consistently worse outcomes for these patients may misrepresent the uncertainty around this issue. In summary, there is evidence to suggest that the prognostic value of *NTRK* fusions varies across cancer sites. Furthermore, it is unclear from the evidence available whether *NTRK* fusions are in themselves prognostic, or whether it is their association which other prognostic factors in the presented studies that drives the observed difference in prognosis.

The ERG also notes that the study from which Figure 8 in the company's Technical Engagement response was reproduced did not claim to demonstrate a statistically significant difference in outcomes between patients with *NTRK1* rearrangements and other patient subgroups.

2.11 Model structure (Issue 14)

The company followed the typical approach adopted in cancer appraisals of directly using extrapolated PFS and OS to populate a partitioned survival model. However, this is limited in the present decision problem by the lack of mature and reliable PFS and OS data for both the intervention assessed and the comparator. The ERG raised concerns regarding the representativeness of the population recruited to the entrectinib trials, the uncertainties around positioning of entrectinib, potential confounding due to secondary therapy received (ERG report, Section 5.2.6). In addition, the suitability of the comparator arm was questioned (**Issue 10**, Section 2.9).

Comparator OS and PFS data for each tumour type were generated from multiple NICE TAs, which were then weighted by the distribution of tumour types in the integrated efficacy analysis. The OS and PFS data were extrapolated assuming an exponential survival function. However, the ERG notes that the complex nature of the comparator arm is unlikely to be well represented by an exponential function, given that this is one of the least flexible parametric functions used to model time-to-event data. In response to **Issue 10**, the company provides an intra-patient analysis which gives [redacted] (Section 2.9.1). Whilst this [redacted], the fact that no parametric model was fitted to the new comparator arm, means that it is not possible to fully determine how using this alternative comparator would affect the ICER, or to quantify the uncertainty in this median.

The company's model was designed to provide an estimate of a single "full population" ICER and does not capture the heterogeneity in the patient population. In their response to **Issue 9**, the company states that since there is no robust evidence of heterogeneity in PFS or OS by tumour type, these can be assumed homogeneous and a single ICER provided for a single "full population" decision. However, the ERG notes that the reason for no reliable evidence of variability on PFS or OS is the small number of patients with some tumour types and the immature nature of the PFS and OS data available. In the absence of direct evidence, the ERG believes it is preferable to make the more conservative assumption that there is heterogeneity in PFS and OS across tumour types, particularly

as there is evidence of heterogeneity in response across tumour types (**Issue 9**, Section 2.8). The ERG considered a response-based model to explore the impact of heterogeneity in PFS and OS between tumour types (ERG report, Section 6.5.1) and noted that this had a considerable impact on the ICERs which varied across tumour types (ERG report, Table 55). Although the ERG acknowledges that there are drawbacks to this type of analysis (see response to **Issue 10**, Section 2.9), the current simplistic structure of the model does not fully account for the uncertainty in effectiveness across tumour types and is limited by the small sample size and immaturity of the available survival data. In addition, the current modelling framework does not account for the costs and benefits in unrepresented tumour types, even though a decision is still being made for these tumour types.

2.12 Extrapolation of overall and progression-free survival (Issue 15)

The exponential model as selected by the company represents the most optimistic extrapolation of PFS and OS of those presented in the company's landmark analysis (excluding the unrealistic log-normal and log-logistic functions). The ERG favour the Weibull curve, as the estimates of pre- and post-progression survival are more clinically plausible than those generated by the exponential distribution. The exponential was the only function to predict that post-progression survival was longer than pre-progression survival – the ERG considered this inappropriate given that only a small proportion of patients received treatment following progression, and that few effective treatment options would be available at this position in the treatment pathway. The ERG notes that the Weibull remains the second-most optimistic distribution, and has only a negligible difference in AIC and BIC fit statistics compared to the exponential.

2.13 Administration costs and resource use (Issue 17)

The ERG considers it a limitation that the administration costs for each of the comparators were not adequately captured in the model. The company simplified the comparators into categories and assumed equivalent costs for comparators within the same category, yet there are considerable and demonstrable differences between comparators. For example, eribulin is classed as a simple Intravenous (IV) therapy with infusion time of 2-5 minutes, while trabectidin, also classed as a simple IV therapy, is administered over a period of 24 hours. The company failed to provide the requested scenario analysis using individual administration costs owing to time constraints and the anticipated small impact of administration costs on the ICER; however, as this scenario has yet to be presented, the exact impact on the ICER is unknown and thus remains an area of uncertainty.

The ERG does welcome the inclusion of the chemotherapy tariff cost of £120 per visit for oral treatments, which was missing from the economic model. The resulting impact is an increase in the incremental cost of entrectinib from [redacted] and an increase in the base case ICER of entrectinib of £2,118.

The ERG also welcomes the scenario which included revised progressed-disease health state costs. As anticipated, the impact on the ICER is small with an increase in the ICER of £289.

2.14 Implementation and training costs (Issue 18)

The company recognises that in order for successful implementation of histology-independent drugs into the health care system, additional training and education for oncologists and pathologists is required. The company propose to work with the NHS to support training in these areas.

However, the company have not addressed the issue of the additional infrastructure that would be required if *NTRK* fusion testing across cancers were to be implemented. As discussed in Section 5.2.3.7 of the ERG report, an increased number of pathological referrals would require additional infrastructure and a larger workforce. The infrastructure requirements and training needs are likely to depend on the testing strategy used. If, as the company suggest, IHC is used as an interim test, a considerable investment in histopathological services would be required in order to manage the considerable number of referrals, given its current capacity constraints. The infrastructure requirements and training needs are likely to depend on the position in the disease pathway where testing will occur and this is still uncertain.

2.15 Uncertainty around the cost-effectiveness results (Issue 20)

The ERG's principal concern regarding the company's approach to modelling uncertainty relates to the presentation of a single ICER. The ERG considers the use of a single ICER to misrepresent the uncertainties in this appraisal. A single ICER is unable to capture between-tumour type heterogeneity in costs and efficacy. The ERG considers it important that this uncertainty is properly reflected in the economic analysis to allow the committee to consider the risk associated with a histology independent recommendation and to allow the committee to consider an optimised recommendation. Further, this uncertainty is potentially important to consider in the context of future data collection, as it will determine the value of collecting further data.

2.16 End of life (Issue 21)

As in **Issue 20** above, the ERG's primary concern is the effect of heterogeneity between tumour sites on decision uncertainty and how to interpret the End of Life criteria (EoL) given this heterogeneity. While the ERG report stated that the average mean OS across tumour sites was 23.64 months for SoC, and thus met the first EoL criterion on the whole, there are a number of tumour sites which would not meet this criterion. According to the distribution of tumour sites proposed in the ERG report, approximately 31% of the incident population eligible for entrectinib would have tumour types that do not meet the first criterion (i.e. neuroendocrine tumours, cholangiocarcinoma, thyroid tumours, and infantile fibrosarcoma). It is also uncertain whether extension to life would fall above the 3 months

required across all affected tumour sites, as a homogeneous treatment effect cannot be assumed (see Section 2.8).

The committee may view a 'histology independent' decision as a number of smaller decisions affecting many treatment pathways in which entrectinib may not always be the most cost-effective option. Therefore, it may not be considered appropriate to apply the end-of-life WTP threshold for tumour types in which it would not be met in a single-indication appraisal.

2.17 Cancer Drugs Fund (Issue 23)

The company asserts the proposed position of entrectinib should be within the CDF and reiterated that entrectinib meets the criteria for inclusion into the CDF. The company's revised base-case ICER of £49,358 per QALY suggests there is plausible potential for cost-effectiveness. However the ERG's base-case ICER presented in the ERG report of £77,120 per QALY, and the revised base-case of £76,322 per QALY (see Section 3, Table 4) is substantially above the £50,000 threshold considered when a technology meets EoL criteria. The ERG acknowledges these ICERs are inclusive of testing costs, but as outlined in **Issue 4**, the ERG considers it important that any costs associated with testing should be properly accounted for as they represent additional expenditure that would be required to implement entrectinib in the NHS.

The ERG also has concerns that the company's and NICE technical team's assessment that this intervention meets the criteria for inclusion in the CDF is based on a single, composite ICER. This single ICER is calculated as a weighted average across the 13 tumour types included in the efficacy evaluable data set. Due to the histology-independent nature of the proposed marketing authorisation, the use of a single ICER means a decision is being made about the use of entrectinib in every possible tumour type treated in adult and paediatric patients. The ERG notes that such a decision carries a very considerable uncertainty, as it will inherently include tumour types that were not included in the efficacy evaluable data set, and therefore have no evidence of plausible cost-effectiveness, plus tumour types in which there is evidence suggesting entrectinib is not cost-effective, and could not be cost-effective, given the cost of identifying patients. For example, analyses conducted by the ERG and presented in the ERG report (Section 2.2.2, Appendix B) show that for pancreatic cancer, upwards of 6,500 patients would need to be screened annually for *NTRK* fusions, to identify 15 patients eligible for entrectinib. Based on the frequency of *NTRK*-fusions estimated at [redacted] in pancreatic cancer (FMI data set), the cost of identifying a single patient with pancreatic cancer assuming IHC followed by confirmatory NGS is [redacted]. By using a single, composite ICER, tumour types such as pancreatic cancer will be included in the CDF despite no evidence indicating the use of entrectinib could plausibly be cost-effective in these tumour types. The ERG reiterates its concerns regarding a non-optimised decision, due to the failure to address or elucidate the heterogeneity underlying this single ICER. That is, a single ICER weighted according to the distribution of tumour types in the trial fails

to adequately consider the possibility of heterogeneity in the cost-effectiveness for different tumour types.

The ERG also notes that the company's assertion that entrectinib shows plausible potential for satisfying the criteria for routine use assumes the average extension to life across tumour types satisfies end-of-life criteria. As noted in response to **Issue 21** (Section 2.16), the ERG questions this assumption as it fails to fully account for the heterogeneity in both the magnitude of the extension to life, and the current survival expectations across different tumour types. It may not be appropriate to apply the end-of-life WTP threshold for tumour types in which it would not be met in a single-indication appraisal, and therefore it may not be appropriate to include these tumour types in the CDF.

The company has proposed its data collection plan to be undertaken during a period in the CDF, which the ERG accepts may help to address some of this uncertainty. This plan includes collecting data on the exact position entrectinib is being used in the treatment pathway, the distribution of *NTRK* fusion positive cancers in England, and the positioning of NGS screening in the treatment pathway. However, the ERG questions whether the proposed data collection will reduce the key areas of uncertainty surrounding a non-optimised decision. Notably, data collection in the CDF in the form of continued trial follow-up will provide more mature PFS and OS data, but will not necessarily offer survival data across an increased breadth of tumour types, including those unrepresented in the efficacy evaluable data set.

3 ERG additional analyses

3.1 ERG's updated preferred base-case

Table 4 presents the results of the ERG original base-case analysis. These incorporate a number of changes to key model parameters and assumptions to the company base-case analysis, previously reported in the ERG report, Section 6.4.

Table 4 Original ERG base-case analysis

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,853	1.61	1.03	[redacted]	[redacted]	[redacted]	£77,109
Entrectinib	[redacted]	[redacted]	[redacted]				

3.1.1 Revised Testing Calculations

The ERG have revised the estimation of the number requiring confirmatory testing to better account for the individual prevalence of each tumour. In addition, the ERG have updated the incidence of thyroid tumour (NOS). Table 5 presents the ERG's alternative base case with the revised testing calculations.

Table 5 Scenario analysis exploring revised testing calculations testing on ERG alternative base-case analysis

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,853	1.61	1.03	[redacted]	[redacted]	[redacted]	£77,618*
Entrectinib	[redacted]	[redacted]	[redacted]				

* corrected for a minor error in the ERG model. Committee were aware of these corrected values in their decision making.

Revising the testing calculations to better reflect the number of individuals requiring confirmatory testing has reduced the ICER.

3.2 Scenario Analyses

3.2.1 Removal of Inappropriate Costs and Outcomes in the Comparator Arm

Table 6 presents the scenario where the costs and outcomes of comparators considered to be inappropriate by NHSE and NHSI have been removed from the ERG's alternative base-case, with updated testing estimates.

Table 6 Scenario analysis exploring NHSE and NHSI's preferred comparators and revised testing calculations on ERG alternative base-case analysis

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,110	1.59	1.02	[redacted]	[redacted]	[redacted]	£76,745*
Entrectinib	[redacted]	[redacted]	[redacted]				

* corrected for a minor error in the ERG model. Committee were aware of these corrected values in their decision making.

The total QALYs for established management was reduced due the removal of costs and outcomes of comparators deemed inappropriate by NHSE and NHSI.

3.2.2 Revised Administration Costs

Table 7 presents the scenario which explores the impact of including the oral chemotherapy tariff cost for both entrectinib and comparator on the revised ERG base-case analysis, with updated testing estimates.

Table 7 Scenario analysis exploring the revised administration costs and revised testing calculations on ERG alternative base-case analysis.

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,989	1.61	1.03	[redacted]	[redacted]	[redacted]	£80,280*
Entrectinib	[redacted]	[redacted]	[redacted]				

* corrected for a minor error in the ERG model. Committee were aware of these corrected values in their decision making.

The inclusion of the tariff cost of oral chemotherapy for both the Entrectinib and comparator arm resulted in an increase to the ICER.

3.3 ERG Revised, Alternative Base-Case Analysis

Table 8 presents the results of the ERG revised alternative base-case analysis. These incorporate a number of changes to the model input parameters, which were previously explored individually in Table 5 to Table 7. The ERG revised alternative base-case analysis includes the following changes to the ERG's original base-case analysis (Table 4):

- Revision of testing calculations
- The removal of costs and outcomes of comparators that NHSE and NHSI deemed inappropriate
- Inclusion of the oral chemotherapy tariff cost to both entrectinib and the comparator oral chemotherapies

Table 8: ERG revised alternative base-case analysis

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,209	1.59	1.02	[redacted]	[redacted]	[redacted]	£79,391*
Entrectinib	[redacted]	[redacted]	[redacted]				

* corrected for a minor error in the ERG model. Committee presented with these correct values in the committee presentation.

The ERG's revised, base-case analysis has increased the ICER following changes to the ERG's original, alternative base-case. The total QALYs gained were reduced by the removal of costs and outcomes of comparators deemed inappropriate by NHSE and NHSI.

3.4 Impact of Testing on the Revised Base-Case

Table 9 to Table 11 present the results of scenario analyses that explore how a variety of testing strategies based on NHSE's proposed, up-front NGS testing influence the ERG's revised base-case analysis. Clinical advisors to the ERG provided the costs of RNA-based NGS (£350) and DNA-based NGS (£250).

3.4.1 Up-front RNA-based NGS: testing at eligibility for treatment

Table 9 explores a scenario where up-front RNA-based NGS is utilised to identify *NTRK* fusion positive patients. The ERG has made a number of assumptions:

- RNA-based NGS will displace FISH testing for patients with MASC or Secretory Breast Carcinoma.
- There is an incremental cost of providing confirmatory RNA-based NGS for patients with *NTRK* positive WGS results
- The total cost of RNA-based NGS for all other tumour sites is incurred, regardless of current test provision.

Table 9: Scenario analysis exploring up-front RNA-based NGS testing

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,209	1.59	1.02	[redacted]	[redacted]	[redacted]	£135,453
Entrectinib	[redacted]	[redacted]	[redacted]				

Due to the higher cost, the use RNA-based NGS at eligibility for treatment increases the ICER.

3.4.2 Up-front RNA-based NGS: testing at diagnosis of advanced or metastatic cancer

Table 10 explores a scenario where up-front RNA-based NGS is utilised to identify *NTRK* fusion positive patients. The ERG has made a number of assumptions:

- RNA-based NGS will displace FISH testing for patients with MASC or Secretory Breast Carcinoma.
- The incremental cost of providing confirmatory RNA-based NGS for patients with *NTRK* positive WGS results is included
- Total cost of RNA-based NGS for all other tumour sites is incurred, regardless of current test provision.

Table 10: Scenario analysis exploring up-front RNA-based NGS testing at diagnosis

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,209	1.59	1.02	[redacted]	[redacted]	[redacted]	£221,783
Entrectinib	[redacted]	[redacted]	[redacted]				

Due to the higher cost and an increase in the number of individuals who would require testing, the use of RNA-based NGS at eligibility for treatment increases the ICER considerably. As RNA-based NGS is carried out at diagnosis of advanced/metastatic cancer, it is likely that a proportion of individuals with an *NTRK* fusion will not be eligible for treatment or will have died when they reach the appropriate position in the treatment pathway. Therefore, the overall cost to identify one patient eligible for treatment increases.

3.4.3 Up-front RNA-based NGS, displacement of DNA-based NGS: testing at diagnosis of advanced or metastatic cancer

Table 11 explores a scenario where up-front RNA-based NGS is utilised to identify *NTRK* fusion positive patients. However, to account for current testing provision for the tumours where multi-panel testing is currently available, the ERG assumes that RNA-based NGS will displace the current provision. The ERG has made a number of assumptions:

- RNA-based NGS will displace FISH testing for patients with MASC or Secretory Breast Carcinoma.
- The incremental cost of providing confirmatory RNA-based NGS for patients with *NTRK* positive WGS results is included.
- Total cost of RNA-based NGS for tumours where genomic testing is not available is incurred
- In the tumour types where there is current DNA-based testing, it is assumed that RNA-based NGS will displace DNA-based NGS.

Table 11: Scenario exploring up-front RNA-based NGS sequencing at diagnosis of advanced or metastatic cancer, where DNA-based NGS is displaced with RNA-based NGS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	£19,209	1.59	1.02	[redacted]	[redacted]	[redacted]	£168,123
Entrectinib	[redacted]	[redacted]	[redacted]				

Due to the displacement of DNA-based NGS, and therefore a reduction in the incremental cost for tumours where DNA-based NGS is available, the total ICER is lower compared to the two previous scenarios. However, the total ICER is higher than the ERG's revised base-case, as there are still significant costs associated with RNA-based NGS in the tumours where no genomic testing is available.

3.4.4 DNA-based NGS with confirmatory RNA-based NGS: testing at diagnosis of advanced or metastatic cancer

Table 12 explores an exhaustive testing approach proposed by ESMO, where first-line DNA-based NGS is used, followed by confirmatory RNA-based NGS. The ERG has made a number of assumptions:

- No incremental cost for patients with MASC or secretory breast carcinoma, where FISH testing is currently reimbursed by the NHS.
- The incremental cost of providing confirmatory RNA-based NGS for patients with *NTRK* positive WGS results is included.
- The incremental costs of providing confirmatory RNA-based NGS for patients with *NTRK* positive DNA-based NGS results is included
- Total cost of DNA-based NGS and RNA-based NGS is incurred for tumours where genomic testing is not available

Table 12: Scenario exploring up-front DNA-based NGS sequencing at diagnosis of advanced or metastatic cancer, with confirmatory RNA-based NGAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER incremental (£/QALY)
Established management	[redacted]	1.59	1.02	[redacted]	[redacted]	[redacted]	£126,628
Entrectinib	[redacted]	[redacted]	[redacted]				

The use of DNA-based NGS followed by RNA-based NGS is one of the most exhaustive methods to identify *NTRK* fusions. The utilisation of this testing method, compared to the hierarchical approach has increased the ICER, but to a lesser extent than the previous scenarios (Table 9 to Table 11). This is because there are no incremental costs associated with testing for tumours where DNA-based NGS is available. In addition, for the tumour types where there is no genomic testing available, the cost of DNA-based NGS is lower.

3.5 Additional benefits required from introducing molecular testing

To explore the potential for positive externalities resulting from the introduction of molecular testing across all advanced and metastatic cancer patients, the ERG presents a number of scenario analyses which consider the magnitude of the benefits that would need to be generated and what this would mean in terms of the frequency of actionable gene targets across all advance and metastatic cancers.

To do this, we consider a scenario in which testing costs are excluded from the revised base-case presented in Section 3.3 and estimate the net monetary benefit (NMB) per treated patient assuming a willingness to pay threshold of £50,000 per QALY. This is then used to estimate the population NMB by multiplying this by the size of the total population eligible to receive entrectinib. This provides a population level estimate of the total benefits to the NHS of implementing entrectinb and can be considered an estimate of entrectinib's contribution to covering the cost of molecular testing. We then estimate the total costs of implementing genomic testing in the NHS. For this analysis, we consider the four scenarios presented in Section 3.4, and similarly multiply test costs per treated patient by the total population eligible for entrectinib, to generate population estimates of testing costs. Note these estimates will be somewhat of an underestimate of total testing costs as they presume that testing will be limited to those tumour sites in which NTRK fusions have been observed.

Table 13 Inputs used in exploratory threshold analysis

Input	Value used
Net monetary benefit per patient treated with entrectinib	£6,641
Annual population eligible for treatment with entrectinib	194
Population Net monetary benefit generated by entrectinib	£1,288,354
Total cost of testing: Up-front RNA-based NGS: testing at eligibility for treatment	£13,210,225
Total cost of testing: Up-front RNA-based NGS: testing at diagnosis of advanced or metastatic cancer	£25,256,159
Total cost of testing: Up-front RNA-based NGS, displacement of DNA-based NGS: testing at diagnosis of advanced or metastatic cancer	£17,768,834
Total cost of testing: DNA-based NGS with confirmatory RNA-based NGS: testing at diagnosis of advanced or metastatic cancer	£11,978,826
Annual advanced and metastatic cancer population	97,201

As can be seen from Table 13 the population NMB generated by entrectinib is much smaller than the total costs to the NHS of implementing molecular testing across all alterative testing strategies. The introduction of entrectinib alone is therefore insufficient for molecular testing to be consider cost-effective. However, as discussed in Section 2.4.3, the introduction of new targeted therapies may justify this expenditure on molecular testing. If we make the simplifying assumption that the NMB per patient treated with a new targeted therapy is the same as that generated by entrectinib we can consider how many patients would need to be treated with the new targeted therapy for molecular testing to be cost-effective. Table 14 summarises the number of patients that would need to receive a new target therapy to justify the costs of introducing molecular testing and the prevalence of actionable targets required given the size of the incidence of advanced and metastatic cancer in England. Note these figures are net of the contribution entrectinib makes to cover molecular testing costs.

Table 14 Number of patients required to be treated with new targeted therapies

Testing strategy	Annual number of patients required	Prevalence of actionable targets
Total cost of testing: Up-front RNA-based NGS: testing at eligibility for treatment	1795	1.85%
Total cost of testing: Up-front RNA-based NGS: testing at diagnosis of advanced or metastatic cancer	3609	3.71%
Total cost of testing: Up-front RNA-based NGS, displacement of DNA-based NGS: testing at diagnosis of advanced or metastatic cancer	2482	2.55%
Total cost of testing: DNA-based NGS with confirmatory RNA-based NGS: testing at diagnosis of advanced or metastatic cancer	1610	1.66%

The results of this analysis show that the number of patients that would need to be treated with a new target therapy ranges from 1,603 to 4,869 patients or 1.66% to 3.71% of the annual incident advanced and metastatic cancer population. Note these figures are net of the contribution entrectinib makes to cover molecular testing costs and so these would be for a new therapy.

4 Conclusions

Reflecting the ERG's critique of the company submission, the technical engagement report lists a substantive number of issues that centre on the challenges associated with assessing cost-effectiveness in a histology independent indications, and the uncertainties associated with the limited evidence of effectiveness available. As outlined in the ERG's original report, the fundamental challenges associated with this appraisal relate to the heterogeneity across tumour types and how to consider the

uncertainties that arise as consequence of this heterogeneity. The company's response seeks to address a number of these issue in particular those relating to potential heterogeneity in the treatment effect, the generalisability of the efficacy data for entrectinib and its comparability with the modelled comparator data, as well as uncertainties regarding the costs of identifying *NTRK*-fusion positive patients.

The company's stated position regarding the heterogeneity of the treatment effect is that the assumption of a uniform level of response is reasonable given the available evidence of treatment activity across a range of tumours. However, no new evidence was presented to support this position and the BHM analysis presented by the ERG demonstrates considerable uncertainty in response rates for included tumour sites (Section 2.8). Further, there is no evidence that response rates in unobserved tumours would be the same as those in the observed tumours.

The company provides new evidence on the appropriateness of the model comparator using data from [redacted]. The ERG considers this approach promising and it is reassuring that the predicted PFS is similar to that in the company base-case. It is, however, disappointing that this approach was not implemented in the economic model. It is also disappointing that the company did not seek to address some of the issues with the response based analysis implemented by the ERG as this approach has the advantage that it allows for heterogeneity in response to be incorporated within the economic model.

The company also presents a number of new and important scenarios in relation to testing costs. In these new scenarios a new testing strategy based upon using NGS as a first-line screening tool is modelled and alternative assumptions regarding the proportion of the testing costs that may be attributed to entrectinib explored. The company also put forward arguments which suggest that testing costs should not be considered, given NHSE's planned role out of genomic testing. The ERG does not consider these arguments valid and considers it important that all incremental costs associated with the implementation of entrectinib are accounted for. The ERG does not consider it appropriate to attempt to apportion only part of the testing costs to entrectinib without evidence to support the wider benefits of testing.

The ERG's scenario analysis centred on the issue of testing costs and a revised base-case was generated in which the testing strategy was revised to better reflect NHSE's current planned role of pan-cancer genomic testing. The incorporation of these revised testing costs increases the ICER substantially. Depending on the testing strategy adopted, the ICER rises from £82,311 per QALY using a hierarchical testing approach to between £126,722 and £282,282 per QALY.

The ERG also emphasises the limitations of the company's approach to estimating a single "full population" ICER. This approach does not capture the heterogeneity in the patient population, and, as highlighted in the ERG report, this heterogeneity is likely to lead to substantive differences in the cost-effectiveness of entrectinib across different tumour types. The ERG considers it important that this heterogeneity in the cost-effectiveness be appropriately reflected in the economic modelling, along with any associated uncertainty. As highlighted in Section 2.11, the ERG considers the model structure based on a partition survival approach to be limited in this regard. A response based model as presented in the ERG report may be preferable as it confers a number of advantages in that it can accommodate heterogeneity across the population and can better accommodate the uncertainty associated with a histology independent decision.

References

1. Solomon JP, Hechtman JF. Detection of NTRK Fusions: Merits and Limitations of Current Diagnostic Platforms. *Cancer Res* 2019.

2. Solomon JP, Linkov I, Rosado A, Mullaney K, Rosen EY, Frosina D, et al. NTRK fusion detection across multiple assays and 33,997 cases: diagnostic implications and pitfalls. *Mod Pathol* 2019. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31375766

3. Marchiò C, Scaltriti M, Ladanyi M, Iafrate AJ, Bibeau F, Dietel M, et al. ESMO recommendations on the standard methods to detect NTRK fusions in daily practice and clinical research. *Ann Oncol* 2019.

4. Williams HL, Walsh K, Diamond A, Oniscu A, Deans ZC. Validation of the Oncomine™ focus panel for next-generation sequencing of clinical tumour samples. *Virchows Archiv: European Journal of Pathology*;**473**:489-503.

5. Mercier F, Consalvo N, Frey N, Phipps A, Ribba B. From waterfall plots to spaghetti plots in early oncology clinical development. *Pharmaceutical Statistics* 2019;**18**:526-32. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/pst.1944

6. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay J-Y, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. 2015;**373**:726-36.

7. Heinrich MC, Joensuu H, Demetri GD, Corless CL, Apperley J, Fletcher JA, et al. Phase II, open-label study evaluating the activity of imatinib in treating life-threatening malignancies known to be associated with imatinib-sensitive tyrosine kinases. 2008;**14**:2717-25.

8. Chugh R, Wathen JK, Maki RG, Benjamin RS, Patel SR, Myers PA, et al. Phase II Multicenter Trial of Imatinib in 10 Histologic Subtypes of Sarcoma Using a Bayesian Hierarchical Statistical Model. *J Clin Oncol* 2009;**27**:3148-53. Available from: https://ascopubs.org/doi/abs/10.1200/JCO.2008.20.5054

9. Center for Drug Evaluation and Research. *NDA Multidisciplinary Review and Evaluation NDA 210861 and NDA 211710 VITRAKVI (larotrectinib)*: U.S. Food and Drug Administration, ; 2016. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/210861Orig1s000_211710Orig1s000MultidisciplineR.pdf

10. Spiegelhalter DJ, Abrams KR, Myles J. *Bayesian approaches to clinical trials and Health-Care Evaluation*. New York: Wiley; 2004.

11. Freidlin B, Korn EL. Borrowing information across subgroups in phase II trials: is it useful? *Journal of Clinical Cancer Research* 2013;**19**:1326-34.

12. Cunanan KM, Iasonos A, Shen R, Gonen M. Variance prior specification for a basket trial design using Bayesian hierarchical modeling. *Clin Trials* 2019;**16**:142-53.

13. Hatswell AJ, Sullivan WG. Creating historical controls using data from a previous line of treatment - Two non-standard approaches. *Stat Methods Med Res* 2019:962280219826609. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30698076

14. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. *J Clin Oncol* 1983;**1**:710-9. Available from: https://ascopubs.org/doi/abs/10.1200/JCO.1983.1.11.710

15. Hatswell AJ, Thompson GJ, Maroudas PA, Sofrygin O, Delea TE. Estimating outcomes and cost effectiveness using a single-arm clinical trial: ofatumumab for double-refractory chronic lymphocytic leukemia. *Cost Effectiveness and Resource Allocation* 2017;**15**:8. Available from: https://doi.org/10.1186/s12962-017-0071-x

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Technical report

Entrectinib for treating NTRK fusion-positive solid tumours

This document is the technical report for this appraisal. It has been prepared by the technical team with input from the lead team and chair of the appraisal committee.

The technical report and stakeholder's responses to it are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the appraisal committee meeting.

The technical report includes:

- topic background based on the company's submission
- a commentary on the evidence received taking into account written statements
- technical judgements on the evidence by the technical team
- reflections on NICE's structured decision-making framework.

This report is based on:

- the evidence and views submitted by the company, consultees and their nominated clinical experts and patient experts and
- the evidence review group (ERG) report.

The technical report should be read with the full supporting documents for this appraisal.

1. Topic background

1.1 Disease background: Neurotrophic tyrosine kinase (NTRK) fusion-positive solid tumours

- There are 3 NTRK gene fusions, NTRK1/2/3
- NTRK gene fusions are oncogenic drivers and are found in a wide variety of cancers including non-small cell lung cancer, breast cancer, pancreatic cancer and rare tumour types such as sarcoma and papillary thyroid cancer
- An estimated prevalence of [redacted]% has been reported by the company but prevalence of NTRK gene fusion varies across different tumour types, ranging from less than 1% prevalence (for example in head and neck cancer) to 91% to 100% prevalence (for example in secretory carcinoma of the salivary gland and congenital fibrosarcoma)
- Central nervous system (CNS) metastases are common in tumour types that are associated with NTRK gene fusions
- Treatment for rare, advanced cancers is often limited to standard chemotherapy with associated toxicity

1.2 Appraisal background: tumour site agnostic treatments

- This is one of the first technologies to be appraised for a histology-independent indication, with treatment determined by the presence of a specific type of genomic alteration, rather than the location of the tumour
- The anticipated indication included in the company's submission was for [redacted]
- [redacted]

[redacted]

- This appraisal considers any tumour type exhibiting the NTRK 1, 2 or 3 fusions
- Genomic testing is required to identify solid tumours with NTRK 1, 2 or 3 fusions. Testing procedures are not standardised across all tumour types at present

1.3 Treatment pathway and positioning of entrectinib

- There is no established treatment pathway specifically for patients with NTRK fusion-positive tumours. Treatment is guided by tumour-specific care guidelines
- The position where entrectinib would be offered is likely to vary by the availability of effective treatments in each tumour
- The company's proposed positioning of entrectinib is outlined in Table 1

Table 1: Company's proposed positioning of entrectinib for the treatment of NTRK fusion-positive, locally advanced or metastatic solid tumours (reproduced from company submission, Table 6)

Position of entrectinib in line of systemic therapy	
First-line*	**Second-line and beyond†**
Mammary analogue secretory cancer (MASC)	Non-small cell lung cancer
Soft-tissue sarcoma	Breast
Pancreatic cancer	Thyroid cancer
Cholangiocarcinoma	Colorectal cancer
Gynaecological cancers	Neuroendocrine tumours

*Patients ineligible for curative surgery or radiotherapy with no immunotherapy or targeted therapy options

†Some patients may receive first-line entrectinib treatment if not eligible for targeted or immunotherapies

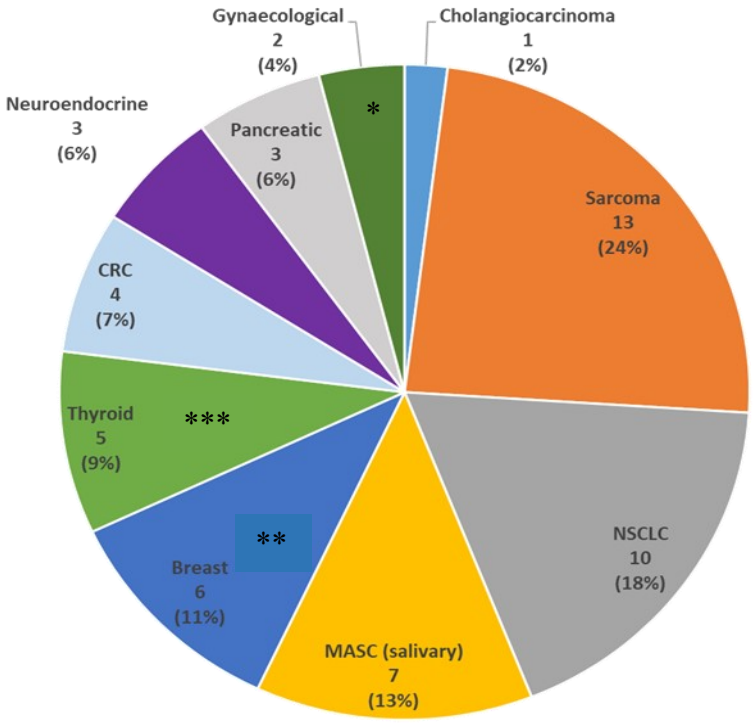
1.4 Clinical evidence

- The following information describes the clinical evidence provided in the company submission and at the clarification stage. This data is included in the economic model.

[redacted] The company provided clinical effectiveness data for 54 adult patients with solid tumours enrolled in the ALKA, STARTRK-1 and STARTRK-2 clinical trials, combined into the integrated efficacy evaluable dataset. This did not include primary CNS tumours or paediatric tumours. The proportions of people included in the

integrated efficacy analysis dataset with each solid tumour are given in Figure 1

Figure 1: Tumour types in the company's integrated efficacy evaluable analysis, n=54 (reproduced from company submission, Figure 7)



*gynaecological cancers represent ovarian and endometrioid cancers

**breast cancer represents non-secretory and secretory breast cancer

***thyroid cancer represents papillary and anaplastic thyroid cancer

Abbreviations: CRC, colorectal cancer; MASC, mammary analogue secretory carcinoma; NSCLC, non-small cell lung cancer

- Within the integrated efficacy evaluable dataset there were 2 subgroups based on investigator assessment, CNS metastases present at baseline (n=12) and no CNS metastases at baseline (n=42)

- At clarification, the company provided the data from the latest datacut ([redacted]) for the integrated efficacy evaluable population and

investigator-assessed response data for adults with primary CNS tumours (n=5) and paediatric patients (n=7). All adults had glioma and the tumour types for the paediatric patients were melanoma ([redacted]), infantile fibrosarcoma ([redacted]), glioma ([redacted]) and CNS embryonal tumour ([redacted]). The primary CNS tumours were excluded from the company's original integrated efficacy analysis dataset because these patients were assessed using Response Assessment in Neuro-Oncology (RANO) criteria rather than Response Evaluation Criteria In Solid Tumours (RECIST) v1.1 criteria

- The company did not include paediatric patients in their efficacy evaluable dataset as they consider the inclusion of paediatric tumours to be very challenging because of the absence of any robust comparator data, particularly those with infantile fibrosarcoma

- The company have included a scenario analysis in their economic model that includes the clinical effectiveness data for the integrated efficacy evaluable set, paediatric patients and adults with primary CNS tumours. After technical engagement, this became the company's base case population

1.5 Key trial results

Median survival follow-up among all adults in the efficacy evaluable analysis set (May 31st 2018 datacut) was 12.9 months (range: 0.6 to 24.7 months). Median survival follow-up not available for [redacted] datacut.

Table 2: Clinical effectiveness results for entrectinib for the company's efficacy evaluable population, [redacted] datacut. Does not include people with primary CNS tumours or paediatric patients

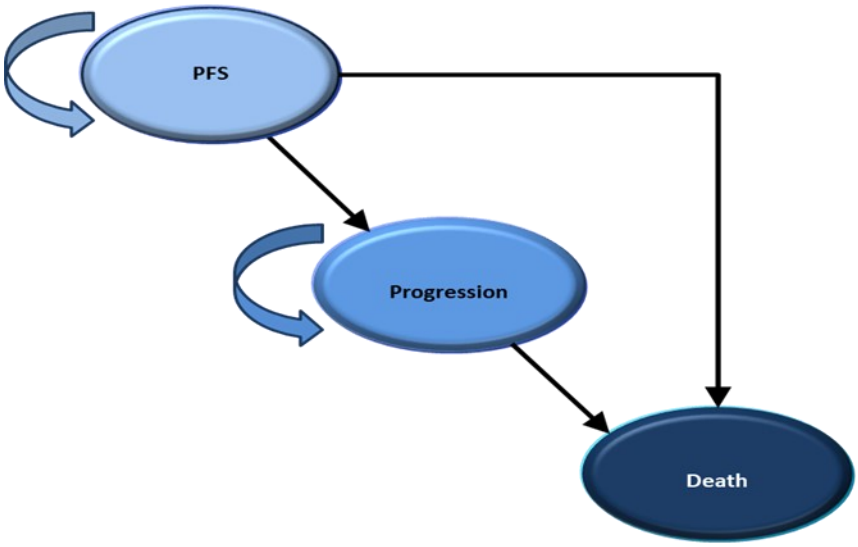
Outcome	Company's efficacy evaluable data set (n=54)
Objective response	
Responders, n (%)	[redacted]
Complete response	[redacted]
Partial response	[redacted]
Objective response rate (95% CI)	[redacted]
Duration of response	
Median DOR, months (95% CI)	[redacted]
Overall survival	
People with event, n (%)	[redacted]
Median OS, months (95% CI)	[redacted]

Progression free survival	
People with event, n (%)	[redacted]
Progressive disease, n (%)	[redacted]
Death, n (%)	[redacted]
Median PFS, months (95% CI)	[redacted]

People receiving entrectinib had clinically equivalent responses regardless of the presence or absence of CNS disease. The objective response rate for people with CNS metastases (n=12) was [redacted], and [redacted] for people without baseline CNS metastases (n=42) ([redacted] datacut).

Response data were assessed by the investigator only for adults with primary CNS tumours and paediatric patients. For primary CNS tumours, response was measured using the Response Assessment in Neuro-Oncology (RANO) criteria and not the Response Evaluation Criteria In Solid Tumours (RECIST v1.1) criteria. Of the 7 paediatric patients included in the economic model, tumour response included [redacted] complete response, [redacted] partial responses and [redacted] stable disease. Response data was [redacted]. Data included for the 5 adult primary CNS tumours included [redacted] partial response, [redacted] stable disease and [redacted] progressed disease ([redacted] datacut).

1.6 Company's model structure



Abbreviations: PFS, progression-free survival

1.7 Data in economic model

- Entrectinib arm: efficacy and safety and tolerability were based on results from the integrated efficacy evaluable dataset
- Established management: Median PFS and OS data were extracted for NICE-recommended treatments for each tumour type within the entrectinib trials apart from cholangiocarcinoma, gynaecological tumours and MASC. An average of the medians was calculated when data were extracted from more than 1 technology appraisal. The average median PFS and OS by tumour type are given in Table 3. These PFS and OS values were converted to mean values. This conversion was based on an exponential extrapolation assumption, in order to simulate an exponential area under the curve
- Data for people with cholangiocarcinoma, gynaecological tumours and infantile fibrosarcoma were calculated as the average of known OS and PFS from the other tumour types included in the entrectinib trials, given in Table 3. For MASC, NICE has not produced any relevant guidance, therefore OS data from a failed phase II study of platinum and

gemcitabine in patients with advanced salivary gland cancer was used. PFS was not reported, so an average of other sources was applied

- Common comparators have been assumed for adult and paediatric primary CNS tumours
- One of the 4 primary CNS paediatric patients was a CNS embryonal tumour rather than a glioma but this patient has been grouped with the glioma patients in the model

Table 3: Company's comparator data by tumour type (from company clarification response, Tables 3 & 5)

Solid tumour site	Average of median OS, months	Average of median PFS, months
Breast	12.2	3.0
Colorectal cancer	9.1	2.6
Mammary analogue secretory carcinoma (MASC)	13.8	4.1
Neuroendocrine	39.6	8.0
Non-small cell lung cancer	10.7	3.8
Pancreatic	8.8	5.2
Sarcoma	14.3	3.9
Thyroid	31.0	4.6
Other (gynaecological & cholangiocarcinoma)	15.8	4.1
Glioma	7.95	3.16
Infantile fibrosarcoma	15.8	4.1
Melanoma	6.40	1.50

1.8 Key model assumptions

Area	Assumption	Company justification
Time horizon	30 years	This period is expected to allow for consideration of all costs and outcomes for the relevant population
Clinical effectiveness: PFS & OS (entrectinib)	Exponential	Best statistical fit to the entrectinib data, representing a conservative but clinically plausible assumption
Clinical effectiveness: PFS & OS (comparator)	Exponential	NICE TSD14 highlights conversion of published medians to exponential mean as reasonable in the absence of patient-level data
Clinical effectiveness: prognostic factors (1)	No adjustment made for NTRK fusion positive status	Limited data available, therefore no adjustment is made for the prognostic factors of NTRK fusion-positive status. This is tested within scenario analyses and comprises part of the Cancer Drugs Fund data collection proposal
Clinical effectiveness: prognostic factors (2)	No adjustment for central nervous system (CNS) metastases	Although randomised trials typically do not recruit patients with baseline CNS metastases, variable levels of reporting made it infeasible to adjust the comparator outcomes in the base case. An indicative scenario analysis was conducted to test the influence of a matched proportion of CNS metastases patients on the incremental cost-effectiveness ratio (ICER)
Treatment duration	Entrectinib treatment duration is equivalent to PFS	[redacted]
Treatment waning	None	No treatment waning is a plausible but conservative assumption based on the method of administration and mechanism of action of entrectinib

Health-related quality of life (HRQoL)	Progression-free quality of life data for entrectinib based on EQ-5D data collected by the company in STARTRK-2	Consistent with previous appraisals
	Progression-free quality of life data for comparator based on weighted average of data from previous NICE appraisals	Selected data were identified and accepted within previous NICE technology appraisals
	Post-progression utility is equivalent for comparator and entrectinib	Conservative assumption; limited data for entrectinib post-progression are significantly higher than weighted comparator ([redacted]) and may not represent the full progressed disease health state
	Omission of AE disutilities in the base case analysis	Disutility associated with AEs was assumed to have been captured in the EQ-5D responses in STARTRK-2. This is in-line with the approach taken in past appraisals in oncology

Safety	"Weight increased" is considered to be the only adverse event which has a higher frequency for entrectinib than comparator. Anaemia, fatigue and neutropenia were also included in the economic model but rates were equal between arms	Based on clinical expert feedback. Safety analysis also does not account for chemotherapy adverse events in the base case, which can have a significant impact on quality of life and costs
Subsequent treatment	[redacted] of people who received entrectinib had subsequent treatment	Based on trial data. Examples of the subsequent therapies include [redacted]
	0% of people who received the comparator therapy had subsequent treatment post-progression	Assumption

Screening	Cost-effective approach to screening with immunohistochemistry (IHC) and next generation sequencing (NGS) panel is proposed	The proposed screening approach aims to minimise the cost of screening solid tumours while utilising current screening methods. It represents a conservative approach as it does not account for the benefits incurred outside of this evaluation (e.g. from identifying eligible patients for clinical trials)
	100% of incremental screening applied to entrectinib	Conservative approach; a second NTRK fusion-targeted medicine, larotrectinib, is being appraised on parallel timelines to entrectinib. This approach therefore risks double-counting screening costs in the event that both products are available to the NHS

2. Summary of the technical report

After technical engagement the technical team has collated the comments received and, if relevant, updated the judgement made by the technical team and rationale. Judgements that have been updated after engagement are highlighted in bold below.

2.1 In summary, the technical team considered the following:

- Prevalence of NTRK gene fusions in each tumour site is not widely known. A database (Foundation Medicine Inc. data set) provided by the company including circa 166,000 samples is the best available source of evidence for this appraisal (see issue 1)
- Issues with the generalisability of the trial to clinical practice increases the uncertainty in the cost-effectiveness estimates (see issues 2 and 7)
- Screening pathway and testing costs to identify NTRK gene fusion positive solid tumours remain uncertain and depend on the provisions set up by NHS England in a timeframe that aligns with this appraisal (see issues 3 and 4)
- The diagnostic accuracy of Next Generation Sequencing (NGS) to identify NTRK fusions is uncertain (see issue 5)

- Estimates of the distribution of the tumour types based on the same dataset used to calculate the prevalence of NTRK gene fusion positive tumours for each tumour site are more reflective of what would be seen in clinical practice in England than the distributions seen in the company's clinical trials. The distribution of tumour types likely to be seen in clinical practice in England is an outstanding uncertainty (see issue 6)
- Primary CNS tumours and paediatric tumours should be included in the analysis (see issue 8)
- Homogeneity of response across different tumours types is uncertain. Heterogeneity should be captured in the model (see issue 9)
- **Issues with the robustness of the control arm increase the uncertainty of the cost-effectiveness estimates (see issues 10 and 11)**
- Prognostic factors such as the presence of CNS metastases should be adjusted for in the model if there is robust evidence (see issue 12)
- **The duration of treatment with subsequent therapies should be shorter than until death. A treatment duration of 6 months, representing approximately half the post-progression survival duration, is a reasonable assumption. There is some outstanding uncertainty around the use of subsequent therapy and the treatments given because of the uncertainties around the position of entrectinib in the treatment pathway (see issue 13)**
- **A responder-based model structure is a plausible alternative to modelling the decision problem but a partitioned survival approach is also reasonable. Time to next treatment is a reasonable confirmatory analysis (see issue 14)**
- A Weibull extrapolation of overall and progression-free survival is appropriate (see issue 15)

- Drug wastage should be included in the model and the electronic market information tool (eMIT) should be used to source drug acquisition costs for the comparator (see issue 16)
- **The oral chemotherapy tariff and monitoring costs in the resource use for the progressed disease state should be included in the model. There is outstanding uncertainty around the administration costs included for the comparators (see issue 17)**
- Associated implementation and training costs should be considered (see issue 18)
- The company's assumption of using different utility values for the pre-progression health state might be plausible but the magnitude of difference is uncertain and is not well supported by evidence **(see issue 19)**
- A proportion of the tumour types included in the analysis do not meet the end-of-life criteria. Committee will consider this in its judgement. This is compounded by uncertainty in the positioning of entrectinib and judgements around tumours types that are unrepresented in the evidence base (see issue 21)
- Entrectinib is innovative but its adoption is dependent on innovation around testing (see issue 22)
- Entrectinib meets the criteria for inclusion in the Cancer Drugs Fund. However, depending on considerations about the end of life criteria and because of uncertainty around testing costs, entrectinib may not have plausible potential to be cost-effective at the price incorporating the patient access scheme (see issue 23)

2.2 The technical team recognised that the following uncertainties would remain in the analyses and could not be resolved within the standard technology appraisal timelines:

- Follow-up of the entrectinib trials is short and overall and progression-free survival data are immature. This could be resolved through further data collection if entrectinib is recommended in the CDF (see issue 23)
- The relative safety of entrectinib compared with established management is highly uncertain as it could not be determined. The small size of the clinical trial may have limited the opportunity to collect data on rarer adverse events

2.3 The cost-effectiveness results include a patient access scheme for entrectinib.

2.4 Taking these aspects into account, the technical team's preferred assumptions result in an incremental cost-effectiveness ratio (ICER) of £79,330 per QALY gained (see table 11). This estimate does not include the commercial arrangements for eribulin, everolimus, nab-paclitaxel, nintedanib, trabectedin and trifluridine-tipiracil, because these are confidential and cannot be reported here. Cost-effectiveness estimates that included these commercial arrangements would be higher than those reported above.

2.5 Based on the company's modelling assumptions, entrectinib is likely to meet the end-of-life criteria (see issue 21). However, based on the ERG's base case, the mean overall survival is much greater than 24 months for 2 tumour types (neuroendocrine and thyroid tumours) and is around 24 months for cholangiocarcinoma and infantile fibrosarcoma. Neuroendocrine and thyroid tumours represent approximately [redacted] of the incident NTRK fusion population.

2.6 The technology is likely to be considered innovative (see issue 22).

2.7 No equality issues were identified by the company. The technical team are concerned that because entrectinib is administered as an oral capsule, only people who have the ability to swallow will be able to use

entrectinib. At clarification the company noted that paediatric patients in the STARTRK-NG study who were unable to swallow capsules were administered with an experimental formulation which could be sprinkled over food. The company stated that they are currently testing gastrointestinal tube administration of the commercial formulation and is developing a new age-appropriate formulation. This formulation does not yet have a UK market authorisation. The ERG highlight that the application of the higher willingness to pay threshold for a technology that meets the end-of-life criteria may potentially raise issue about the equity of access to treatment. In a case where the end-of-life criteria is considered to be met for the whole population, despite some of the sub-populations not meeting the criteria for end-of-life, this implies that people are able to access therapy that would otherwise be considered cost-ineffective based on conventional thresholds. In this situation, a QALY generated in NTRK positive and NTRK negative patients are being valued differently.

2.8 The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund noted that the 7 Genomic Laboratory Hubs are at different stages of being able to implement Next Generation Sequencing (NGS) multigene panel testing and this variation will be resolved in the next 1 to 2 years. They also note that whole genome sequencing (WGS) will take time to embed within clinical treatment pathways, particularly in respect of the need for the collection and processing of fresh tissue.

3. Key issues for consideration

Issue 1 – Prevalence of NTRK gene fusion

<table>
<tr><td>Questions for engagement</td><td>1. Which dataset more accurately reflects the prevalence of NTRK fusion for each tumour type?</td></tr>
<tr><td>Background/description of issue</td><td>
Prevalence of NTRK gene fusions is uncertain with literature reporting estimates between 0.25% to 0.31% in the adult population and 0.34% to 0.49% in the paediatric/adolescent population.
The company estimates that NTRK gene fusions are present in [redacted] of all cancers. They also report an estimated prevalence of [redacted] based on next generation sequencing (NGS) profiling of 116,398 adult and paediatric tumour samples using the Foundation Medicine Inc. (FMI) NGS platform.
The company calculated the prevalence of NTRK gene fusions by tumour type based on a weighting of literature prevalence estimates (Amatu et al.) and data on file. This analysis was limited to tumour types included in the entrectinib clinical trials.
The ERG considers the company's estimate of [redacted] to be significantly higher than other figures reported in the literature. The ERG estimate the prevalence of NTRK fusion in each tumour type using a larger FMI database that included circa 166,000 samples and was provided by the company at clarification. The prevalence estimate remained at [redacted] across all solid tumours in the analysis of the larger Foundation Medicine dataset.
Table 7 shows the prevalence of NTRK fusion by tumour type as estimated by the company and the ERG.

Table 7: prevalence of NTRK fusion by tumour type
<table>
<tr><th>Tumour Type</th><th>Company's estimates</th><th>ERG's estimates</th></tr>
<tr><td>Salivary gland (MASC)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>NSCLC</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Breast cancer (not specified)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Secretory breast carcinoma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Papillary thyroid tumour</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Thyroid Tumour (NOS)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Colon/colorectal</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Neuroendocrine (NOS)</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Cholangiocarcinoma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Pancreatic</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Uterine</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Ovarian</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Cervix</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Soft tissue sarcoma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>High grade glioma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Paediatric high grade glioma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Congenital mesoblastic nephroma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Paediatric melanoma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Infantile fibrosarcoma</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>Paediatric low grade glioma</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td></tr>
<tr><td>Why this issue is important</td><td>Prevalence of NTRK gene fusions is linked to other inputs in the economic model. It is included in the calculation of the number needed to screen and so impacts on the screening costs (see issue 4). Screening costs are a main driver of the cost-effectiveness results.</td></tr>
</table>

Technical team preliminary judgement and rationale	The technical team agree with the ERG's estimates of the prevalence of NTRK gene fusion as they are sourced from a larger database which seems more appropriate given the rarity of the gene fusions and more likely to be generalisable. The technical team would like to see the analyses including the ERG's prevalence estimates.
Summary of comments	**Comments received from the company:** Agrees with the ERG's preference for using the FMI data set as the source of NTRK prevalence data. Considers the FMI data set to represent the single largest and therefore most robust study of NTRK gene fusion prevalence. Using the ERG's preferred NTRK prevalence estimates decreases the company's updated base case ICER by around £3,400 (£45,941 instead of £49,358). Prevalence is an area of uncertainty that could be resolved through further data collection if entrectinib is recommended in the CDF. **Comments received from NCRI-ACP-RCP-RCR:** True prevalence may be underestimated. Foundation Medicine assay detects NTRK fusion as a DNA-based only test with no RNA component. FMI data set would seem to be the most robust data available currently. **ERG considerations on new evidence received during technical engagement:** No comments.
Technical team judgement after engagement	The technical team consider prevalence estimates derived from the FMI data set to be appropriate for decision making.

Issue 2 – Treatment pathway and positioning

Questions for engagement	2. How will 'acceptability' in the context of this appraisal be defined in clinical practice? 3. For each tumour type, at what points in the respective treatment pathways will entrectinib be used in clinical practice in England? 4. Is the evidence for entrectinib generalisable to clinical practice in England?
Background/description of issue	There is no defined clinical pathway specifically for people with solid tumours expressing the NTRK gene fusion. Treatment is currently guided by tumour-site specific care guidelines. The anticipated marketing authorisation of entrectinib wording does not specify the point in the treatment pathway that entrectinib can be used. There is a mismatch between the positioning of entrectinib in the clinical trials and where entrectinib may be used in clinical practice, in line with the circa 166,000 samples anticipated marketing authorisation. In the entrectinib trials, [redacted] of patients received entrectinib as first line systematic therapy, [redacted] as second line, and [redacted] as third line or beyond. These data include the company's efficacy evaluable population plus 5 patients with primary CNS tumours and 7 paediatric patients. **The company** propose that entrectinib will be used as an alternative to standard chemotherapy when one or more other options have been exhausted, or as first-line option where there are no acceptable alternatives. In its clarification response, the company stated that they anticipate entrectinib to be used in later lines of treatment in the majority of cases, at the point where therapeutic options are very limited or exhausted altogether. The company's proposed positioning of entrectinib for the tumour types included in their clinical trials is given in table 1 above. **The ERG** consider the term [redacted] in the proposed marketing authorisation to be ambiguous. They highlight that there were 'acceptable' alternative standard therapies available for some people who received entrectinib first-line in the clinical trial. The ERG's clinical advisors said that 'acceptability' would be a subjective assessment of the response rates and adverse event burden associated with existing options. They noted that the threshold at which a decision to offer entrectinib would be made is likely to vary between tumour types. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** also considers the wording of the company's anticipated marking authorisation to be open to potentially variable interpretation. They consider that entrectinib was given in the entrectinib clinical trials before other alternative therapies had been exhausted and note that this could have led to potential considerable bias in the entectinib clinical trials. **Table 4: Proposed positioning of entrectinib for each tumour type included in the entrectinib clinical trials and positioning of entrectinib treatment in the clinical trial**

Tumour type	Positioning in the entrectinib clinical trial (from table 8 in ERG report)	Company's proposed positioning (from table 6 in company submission)	NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund's proposed positioning

		MASC	1st line: [redacted] 3rd line or beyond: [redacted]	First-line*	Agrees with company's positioning
		Soft-tissue sarcoma	1st line: [redacted] 2nd line: [redacted] 3rd line or beyond: [redacted]	First-line*	First-line for chemo-resistant sarcomas Second-line for chemo-sensitive sarcoma
		Pancreatic cancer	1st line: [redacted] 2nd line: [redacted] 3rd line or beyond: [redacted]	First-line*	Uncertain, either first- or second-line
		Cholangiocarcinoma	3rd line or beyond: [redacted]	First-line*	Uncertain, either first- or second-line
		Gynaecological cancers (ovarian and endometrioid)	3rd line or beyond: [redacted]	First-line*	Second-line
		NSCLC	1st line: [redacted] 2nd line: [redacted] 3rd line or beyond: [redacted]	Second-line and beyond†	After any immunotherapy and 1st line cytotoxic chemotherapy
		Breast	1st line: [redacted] 2nd line: [redacted] 3rd line or beyond: [redacted]	Second-line and beyond†	Third-line
		Thyroid cancer	1st line: [redacted] 2nd line: [redacted] 3rd line or beyond: [redacted]	Second-line and beyond†	Second-line
		Colorectal cancer	1st line: [redacted] 3rd line or beyond: [redacted]	Second-line and beyond†	Third-line
		Neuroendocrine carcinomas	1st line: [redacted] 2nd line: [redacted]	Second-line and beyond†	Third-line
	*Patients ineligible for curative surgery or radiotherapy with no immunotherapy or targeted therapy options †Some patients may receive first-line entrectinib treatment if not eligible for targeted or immunotherapies				
Why this issue is important	The population who received entrectinib in the clinical trial does not match the population in the NICE scope. The evidence for entrectinib may not be generalisable to clinical practice in England. The higher proportion of people having entrectinib in earlier lines of therapy across tumour types may lead to overestimation of the survival estimates in the entrectinib arm. This also has implications for end-of-life considerations (see issue 21).				
Technical team preliminary judgement and rationale	There is a huge amount of uncertainty about treatment positioning. The company's clinical trial evidence does not align with the company's expected positioning for entrectinib nor NHS England's suggested positioning for all tumour types included in the entrectinib clinical trial apart from MASC. The technical team consider the wording of the proposed marketing authorisation to be open to interpretation and this limits the technical team's understanding of where it should be positioned in the treatment pathway. To make a recommendation in line with entrectinib's marketing authorisation, committee will need to know where entrectinib will be positioned in each of the treatment pathways for each of the tumour types. It is likely that the decision to use NTRK inhibitors would depend on clinician judgement. However, clinical judgement may				

	be difficult to elicit because of the rarity and diversity of these tumour types. This issue is considered to be unresolvable at present. More information is needed and could be collected through the Cancer Drugs Fund.
Summary of comments	**Comments received from the company:** Maintain position given in the original submission (shown in table 4 above) for the point in the treatment pathway at which entrectinib may be used in respective tumour treatment pathways. However, recognise the uncertainty around the positioning of entrectinib in clinical practice. Agree that this issue may only be resolved through further data collection if entrectinib is recommended in the CDF. Clinical expert feedback to the company around the definition of 'acceptable' therapy suggested some further points to consider: the lack of a target biomarker for the therapy could dictate choice, while an ORR of less than 20% and a PFS of less than four months could indicate an 'unacceptable' outcome, though these criteria are clearly dependent on context and a number of variables such as tumour type. Provided a scenario analysis using the NHSE and NHSI's alternative recommended positioning as 3rd line treatment for breast and colorectal cancer and neuroendocrine tumours. This scenario was executed in the economic model by revising the comparator treatments included for these tumour types, see issue 11. Considers the entrectinib clinical evidence to be generalisable to clinical practice in England. Does not consider the proportion of treatment-naïve patients receiving entrectinib to confound comparative results in favour of entrectinib. Numerous reasons given to explain why line of therapy is not a determinant of prognosis, including that the main determinant of prognosis for entrectinib is the patient's fitness, as defined by their Eastern Cooperative Oncology Group Performance Status (ECOG PS), not their line of therapy. Provided [redacted] **Comments received from NCRI-ACP-RCP-RCR:** Definition of 'acceptable' will vary between tumour types primarily determined by the alternative standard of care options. In general, entrectinib should be given second-line and beyond until further data is available. The risk-benefit ratio could favour entrectinib over first-line standard of care treatment in some circumstances for the following tumour types: MASC, chemo-resistant soft tissue sarcoma, pancreatic cancer, NSCLC and paediatric tumours with high tumour TRK fusion prevalence, for example infantile fibrosarcoma. Suggest the following amendments to NHS England's suggested positioning: • Soft tissue sarcoma: second-line and beyond for chemo-sensitive sarcomas • Pancreatic: first or second-line • Cholangiocarcinoma: likely second-line and beyond • NSCLC: first-line and beyond • Breast: second-line and beyond • Thyroid: second-line and beyond but with further data collection, anticipate moving into first-line • Colorectal: third-line and beyond • Neuroendocrine: second-line and beyond. Consider data to be generalisable in the context of a person with NTRK fusion positive solid tumour rather than by disease type and line of therapy. Data within the trials are highly encouraging for responses across different diseases. Agree that there is discrepancy between the trial data and proposed positioning. Further data collection in the context of the CDF vital in assessing the outcomes and providing supportive data to bring into earlier lines of therapy. **Comments received from GIST Support UK:** Entrectinib should be given as first line therapy if the GIST tumour is NTRK fusion positive. **ERG considerations on new evidence received during technical engagement:** Maintain that there are substantive uncertainties regarding the positioning of entrectinib. Considers that entrectinib's positioning will be influenced significantly by the timing of NTRK testing in the treatment pathway. Notes that the continuing data collection programme outlined by the company and the potential placement of entrectinib within the CDF are likely to mean that the evidence supporting the efficacy of entrectinib will evolve significantly. The availability of improved efficacy data may impact on where entrectinib is used in the treatment pathway. Company did not present further economic analysis where the comparator data is matched to the position that entrectinib is given in the integrated efficacy analysis. Instead subgroup analysis was provided. Subgroup analysis not considered to be supportive of the company's conclusion [redacted]. Maintain that the number of previous treatments remains a major inconsistency between the modelled entrectinib population and both the likely treated population and modelled comparator arm. Company's subgroup analysis not included in the economic model. This limits the understanding of the impact line of therapy has on both outcomes and estimated cost-effectiveness.

Technical team judgement after engagement	The technical team maintain that this issue is unresolvable at present. [redacted] More information is needed and could be collected if entrectinib is recommended in the CDF.

Issue 3 – NTRK gene fusion screening pathway

Questions for engagement	5. What is the likely screening pathway to identify NTRK positive solid tumours? 6. At what point in the treatment pathway for each tumour type is/will NTRK gene fusion testing carried out?
Background/description of issue	All solid tumours types can potentially harbour NTRK gene fusions. This means that the number of people who require testing for NTRK gene fusions is very high. A national service for cancer genomic testing has been created by NHS England and is regionally organised by 7 Genomic Laboratory Hubs. The hubs process tissue samples for whole genome sequencing (WGS) and pass these to the Genomic Medicine Service for analysis, perform next generation sequencing (NGS) and interpret all NGS and WGS results before returning them to the requesting clinician. NGS provides the technology for multigene panels (which provide testing for anything between 5 and 500 genes). **The company** stated that in current clinical practice WGS is already funded for adult sarcomas and paediatric tumours, MASC undergo NTRK fusion testing and other biomarker screening is done for colorectal cancer, NSCLC, breast cancer and thyroid cancer. The company state that no molecular testing is currently conducted in clinical practice for neuroendocrine, pancreatic and gynaecological tumours and cholangiocarcinoma. For the cancers where a genetic test is already conducted in clinical practice or where no molecular testing is conducted, the company includes a two-stage approach to testing in its economic model in the entrectinib arm. This comprises an immunohistochemical (IHC) test followed by a NGS test if the IHC result is positive. The company cite clinical data provided by an investigator involved in the entrectinib clinical trial development programme that suggests that IHC testing removes 89% of NTRK negative samples. Therefore, 11% of tumour samples also have a NGS test. **The ERG** considers the company's proposed approach to testing to be broadly plausible. For sarcoma and paediatric tumours, the ERG's clinical advisers have said that in clinical practice a RNA-based NGS test would be needed after WGS to confirm an NTRK-fusion positive tumour. The ERG also noted that the NHS reimburses NTRK-fusion positive screening assays for people with MASC. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlight that for paediatric cancer and sarcoma, funding for WGS is in place. However, it is noted that NGS may be necessary for NTRK fusion testing in the short term until WGS is fully operational. Funding is also currently in place for MASC and the secretory variant of breast cancer through the National Genomic Test Directory for 2019. For all other adult solid cancers, NTRK gene fusion testing is not currently required by the National Genomic Test Directory and is not systematically performed. However, by the end of the 2019/20 financial year, the Genomic Laboratory Hubs plan to introduce NGS gene panels for solid tumour testing, which will include the capability to identify NTRK gene fusions. This could be for example with a 50 to 60 gene panel or a 500 gene panel. The clinical lead notes that for patients and clinicians to be able to best use the information of NGS panel testing, it has to be done prior to starting any systemic therapy for the locally advanced/metastatic disease.
Why this issue is important	The number of people who will be tested for NTRK gene fusions will be high and also costly. The screening pathway is currently uncertain. There is a potential equality issue as service provisional has not yet been rolled out nationally.
Technical team preliminary judgement and rationale	The technical team understands the difficulties in determining the potential screening pathway. It notes that using IHC as a screen is problematic as it will need to be standardised but recognises that NHS England is not intending to use IHC screening to identify NTRK fusions as its focus is on genomic testing. The technical team considers the screening pathway to depend on the provisions set up by NHS England in a timeframe that aligns with the NTRK appraisals. It is recognised that this is a rapidly changing field. If entrectinib is recommended within the Cancer Drugs Fund for a period of managed access, this will be a key issue when entrectinib exits the Cancer Drugs Fund.
Summary of comments	**Comments received from the company:** European Society for Medical Oncology (ESMO) guidance includes the IHC pre-screen in circumstances where NGS availability is limited. Including the IHC pre-screen is seen as an interim step to ensure patient access to entrectinib and avoid the implementation of NGS-based testing becoming the rate limiting step. This approach could also reduce inequity of access to testing and tumour agnostic therapy. Consider the most optimal testing route to be the wide-scale implementation of appropriate NGS-based testing capable of the sensitive and specific detection of NTRK fusion-positive solid tumours as early in the treatment pathway as possible. Notes the NGS-based testing implementation available to all metastatic cancer patients is not anticipated until 2021 at the earliest. A recommendation for use within the CDF gives a window in which to optimise the proposed screening pathway around early use of appropriate NGS screening costs and could be supported by interim use of the IHC pre-screen.

	Technical consideration should be given to the availability of biopsy tissue in certain tumour types with testing material being very limited in many cases. Using parallel marker identification via NGS on limited tissue early in the diagnostic pathway maximises the benefits to be gained. **Comments received from NCRI-ACP-RCP-RCR:** Testing should be performed early in the diagnostic pathway and by IHC followed by confirmatory NGS or upfront NGS, determined by resource. Panel testing is ideal to identify other potentially actionable alternations. Defer to NHSE on the roll out of NGS DNA/RNA panel testing. A validated RNA-based NGS may be preferable over a DNA-based approach. WGS will be available for paediatric and sarcoma patients but patients would still likely require a separate RNA fusion panel to identify a TRK fusion. For disease types with known high prevalence of TRK fusions (including MASC, secretory carcinoma of the breast), upfront testing (pre-first line treatment) is essential and already included in the test directory. **Comments received from GIST Support UK:** Provided the standard screening pathway for GIST, please see GIST Support UK's response to technical engagement. Point of NTRK gene fusion testing will vary for each patient. Testing will only happen where surgery or a biopsy has provided a tissue sample. Testing will be carried out when it has been identified that the patient has tested negative for all of the other known GIST mutations. **ERG considerations on new evidence received during technical engagement:** ESMO recommend a hierarchical approach with testing for NTRK fusions differing depending on the frequency of NTRK fusions in each tumour type and the current available testing. Concerns with the company's proposition for using interim IHC until NGS is available include: • Implementation of IHC would require considerable investment which would then be disregarded after a relatively short time, when NGS is fully operational • If entrectinib is recommended for use in the CDF, the testing strategy must reflect the future genomic testing. Alternative testing strategies could impact on the distribution of tumour types and possibly the number of false positives • The total number of individuals who would require testing to identify eligibility for entrectinib is likely to differ between the testing approaches as hierarchical testing will likely be done just prior to treatment if eligible whereas NGS will be done at diagnosis. This will impact on the investment required. If testing is conducted at the point of diagnosis of locally advanced or metastatic disease, this will impact on the costs of identifying a single NTRK positive patient as some people will be tested but not treated with entrectinib.
Technical team judgement after engagement	The technical team is anticipating receiving further guidance on the diagnostic testing pathway from NHS England and the Genomic Medicine Service ahead of the committee meeting.

Issue 4 – NTRK gene fusion testing costs

Questions for engagement	7. Should testing costs for molecular testing that is already done in the NHS be included in the economic model, even if it is not to detect NTRK gene fusions? 8. Are the costs of adding a panel to an RNA-based NGS test negligible? 9. Is it appropriate to include the costs of confirmatory NGS for people who have WGS? 10. Should the economic model include testing costs for unrepresented tumour types? 11. What proportion of the screening costs should be included in the economic model for this appraisal? Or is it appropriate to exclude all testing costs from the analyses?
Background/description of issue	Testing for NTRK positive tumours is likely to be very costly. It is one of the main drivers of the cost-effectiveness for this appraisal. The process of reimbursing a new target within the NHS Genomic Testing Directory has not yet been established. **The company** includes equal testing costs between arms for adult sarcoma and MASC in their base case and paediatric tumours in scenario analysis as WGS (assumed to cost £800) is already available in NHS clinical practice for adult sarcoma and paediatric tumours and NTRK testing (assumed to cost £75) is funded for MASC. In tumour types where a genetic test is already conducted in clinical practice (colorectal cancer, NSCLC, breast cancer, thyroid cancer), a two-stage approach to testing is included in the entrectinib arm. This comprises an IHC test for pan-TRK followed, if a positive result for NTRK gene fusions, by a NGS test for 11% of people who test positive for IHC (see issue 3). The company assumed that IHC cost was £75 and a NGS panel cost [redacted]. For the comparator, the company applied the cost of an IHC test (assumed to cost £75). In tumour types where no genetic testing is conducted currently, the company did not include screening costs in the comparator arm but in the entrectinib arm included the cost of IHC testing and a confirmatory NGS in 11% of people who had a positive IHC test.

	The company have only included testing costs for tumour types included in their efficacy evaluable dataset (i.e. 13 tumour types) and have not accounted for other tumour types that could be NTRK positive (see issue 7). The company ran scenario analyses that included 50%, 25% and 0% of screening costs attributed to entrectinib. **The ERG** considers the company's proposed approach to testing to be broadly plausible. For sarcoma and paediatric tumours, the ERG's clinical advisers have said that in clinical practice a RNA-based NGS would be needed after WGS to confirm an NTRK-fusion positive tumour. They consider it appropriate to only include testing costs in the comparator arm if the current testing for that tumour type is able to identify NTRK fusions. Therefore, the ERG considers testing costs for MASC, adult sarcoma and paediatric tumours appropriate to include as equal between arms as the NHS already reimburses testing that can identify NTRK fusions for these tumour types. The ERG highlight that as the company have not included the testing costs for the tumour types not represented in their efficacy evaluable dataset (see issue 7), the costs of testing are unlikely to represent testing costs for the population eligible for entrectinib. The ERG ran numerous scenario analyses exploring testing costs: 1. Testing costs were estimated based on all tumour types known to harbor NTRK fusions. They were assumed to be tested using IHC, followed by confirmatory NGS, unless WGS is already available on the NHS. 2. Zero incremental costs included for sarcoma and MASC as WGS is currently funded for sarcoma in clinical practice and current testing for MASC already identifies NTRK fusions. 3. Removed the cost of NGS for lung cancer. This is because the cost of adding a new NTRK panel to an RNA-based NGS test is currently negligible. RNA-based NGS is already used in clinical practice for a subgroup of people with NSCLC. 4. Added the cost of a confirmatory NGS test in people who have already had WGS. This was applied to 11% of the number needed to screen as it is assumed that WGS will remove 89% of NTRK fusion negative samples. In its base case, the ERG estimates testing costs using the number needed to screen based on the whole NTRK population (see issue 1), removes testing costs from the comparator arm if the tests do not identify NTRK fusions, does not include NGS test costs for lung cancer and includes a confirmatory NGS test after the WGS test. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlight that for paediatric cancer and sarcoma, WGS is currently in place. However, it is noted that NGS may be necessary for NTRK fusion testing in the short term until WGS is fully operational. Funding is also currently in place for MASC and the secretory variant of breast cancer through the National Genomic Test Directory for 2019. For all other adult solid cancers, NTRK gene fusion testing is not currently required by the National Genomic Test Directory and is not systematically performed. However, by the end of the 2019/20 financial year, the Genomic Laboratory Hubs plan to introduce NGS gene panels for solid tumour testing, which will include the capability to identify NTRK gene fusions. This could be for example with a 50-60 gene panel (cost ~£250) or a 500 gene panel (cost ~£400). The clinical lead considers that it is appropriate that at least part of the costs for multi-gene panel testing be covered by each company that benefits from the new service provision. They do not support the company's approach for screening given their move to genomic testing. They highlight that the weighted average cost of testing calculated by the company carries very significant uncertainty and is sensitive to the prevalence of NTRK gene fusions in each tumour type included in the calculation. NHS England and NHS Improvement would like to see scenario analyses in which various percentages of the costs of NGS multi-gene panel testing are borne by the company: 100%, 50%, 33% and 0%.
Why this issue is important	The number of people who will be tested for NTRK gene fusions will be high and also costly. Testing costs are a significant driver of cost-effectiveness. **Company's scenario analyses:** Attributing 50% of testing costs to entrectinib decreases the company's base case ICER by around £7,800 (£44,762 instead of £52,609). Attributing 25% of testing costs to entrectinib decreases the company's base case ICER by around £12,000 (£40,838 instead of £52,609). Removing all testing costs from the economic model decreases the company's base case ICER by around £15,700 (£36,914 instead of £52,609). **ERG's scenario analyses:** Calculating testing costs using the prevalence of NTRK fusions based on the whole NTRK population (see issue 1) increased the company's base case ICER by around £13,400 (£65,981 instead of £52,609). Removing testing costs in the comparator arm for tests that do not identify NTRK fusions increases the company's base case ICER by around £10,700 (£63,329 instead of £52,609). Removing the testing costs of NGS for people with lung cancer, as well as the testing costs in the comparator arm for tests that do not identify NTRK fusions, increases the company's base case ICER by around £6,900 (£59,465 instead of £52,609). Including the cost of confirmatory NGS for 11% of people who have WGS, as well as the testing costs in the comparator arm for tests that do not identify NTRK fusions, increases the company's base case ICER by around £12,000 (£64,608 instead of £52,609).

Technical team preliminary judgement and rationale

Incremental testing costs should not be included in the economic model for paediatric tumours, sarcoma, MASC and the secretory variant of breast cancer as testing that will detect NTRK gene fusions is already funded in clinical practice. The technical team recognises that if analyses are run including testing costs for unrepresented tumour types then the analysis is only appropriate if the health effects generated from treating the NTRK positive solid tumours in these unrepresented tumour types are also included. The proportion of overall testing costs that should be included in the analysis for the entrectinib appraisal is not a judgement that the technical team can make at this stage without more input from the company and NHS England. However, the technical team would like to see the scenario analyses in which various percentages of the costs of NGS multi-gene panel testing are borne by the company: 100%, 50%, 33% and 0%, as suggested by The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund.

Summary of comments

Comments received from the company:

Testing costs should be excluded from the analysis for the following reasons:

- Testing for NTRK gene fusions is unlike any previous new biomarker testing in oncology because of the anticipated tumour agnostic licence. Testing cost per patient identified is much higher than has been seen previously.
- Applying testing costs penalises the innovation of the technology as the first therapies using this testing methodology will be forced to bear significant costs, while therapies of the same class entering the market later are likely to fit into existing pathways.
- Numerous other tumour agnostic therapies are in development and therefore it is inappropriate to attempt to apply a proportion of testing costs.
- NGS provides a wealth of additional genetic information, well beyond the value of identifying NTRK gene fusion positive patient.
- Removing the testing component from the appraisal significantly reduces the uncertainty burden.
- NHS England's drive towards pan-tumour, up-front NGS-based testing for all cancer patients is taking place regardless of the advent of tumour agnostic therapies. The cost of implementing this initiative will therefore ultimately be borne by the NHS budget, regardless of the emergence of new therapies that could play a role in this area.

Where testing costs are included, to avoid double counting, the costs of testing for any biomarker or oncogenic driver mutation should be included in the comparator arm, whether or not it is for NTRK gene fusions. NTRK testing including NGS-based testing is likely to replace existing testing by other methods.

There is lab-to-lab variation in the cost of adding a single gene to an RNA-based NGS test but theoretically the costs are negligible.

The use of NGS hybrid panels capable of the parallel detection of both DNA and RNA targets minimises the need for confirmatory follow-up testing.

Availability of test material may be limited in some tumour types and where biopsy material is low in tumour content or insufficient, the ability to do confirmatory testing may be significantly limited.

The company consider it inappropriate to include testing costs for unrepresented tumour types in the absence of outcome data in these tumour types.

The company ran scenario analyses exploring the impact of up-front NGS-based testing at varying proportions, Table 5. The incremental testing costs for MASC, secretory breast cancer and soft tissue sarcoma was held at £0.

Table 5: Company's scenario analyses: testing costs

Proportion of testing costs attributed to entrectinib	Company ICER (£/QALY)	Change from company's updated base case ICER
100%	£96,641	+£47,280
50%	£66,205	+£16,850
33%	£56,060	+£6,700
25%	£50,988	+£1,630
10%	£41,857	-£7,500
0%	£35,770	-£13,600

Comments received from NCRI-ACP-RCP-RCR:

Cost of adding a single gene to an RNA-based NGS test is negligible. Reasonable to include confirmatory NGS costs following WGS. WGS testing may not efficiently detect fusions, unless the assays/bioinformatics are further optimised by the time of roll-out.

Reasonable to include testing costs for unrepresented tumour types in the economic model.

	ERG considerations on new evidence received during technical engagement: Maintains that the incremental costs of testing to identify *NTRK* fusions or Trk-proteins (in the case of IHC) should be included in the economic evaluation of entrectinib as well as the testing costs for unrepresented tumour types. ERG has concerns regarding the appropriateness of scenarios in which only a proportion of the testing costs are attributed to entrectinib. The extent to which the cost of testing will affect the ICER will largely depend on the testing strategy utilised. Provide detail of potential testing strategies including first-line RNA-based NGS and upfront DNA-based NGS with confirmatory RNA-based NGS (see section 2.4.1 of ERG's technical engagement response). Acknowledge that the introduction of pan-cancer genomic testing represents a public good, which may generate positive externalities, particularly as new targeted therapies become available for other genetic targets. Present a threshold analysis showing the magnitude of benefits required for pan-cancer genomic testing to be cost-effective. This gives the opportunity for benefits from future target therapies to be included within the current economic analysis. Details are given in section 3.5 of the ERG's technical engagement response. This analysis shows that the number of patients that would need to be treated with a new target therapy ranges from 1,603 to 4,869 patients or 1.66% to 3.71% of the annual incident advanced and metastatic cancer population to justify the costs of introducing molecular testing. These figures are net of the contribution entrectinib makes to cover molecular testing costs.
Technical team judgement after engagement	The technical team is awaiting further guidance on the diagnostic testing pathway from NHS England and the Genomic Medicine Service ahead of the committee meeting, as per issue 3. Until there is more information available, the technical team's preferred ICER includes the company's hierarchical modelling approach at the point of eligibility for entrectinib that includes an IHC pre-screen with a confirmatory NGS-based test.

Issue 5 – Identification of NTRK gene fusions - diagnostic accuracy

Questions for engagement	12. What is the expected diagnostic accuracy of NGS testing? 13. Is there a testing sequence that could avoid a substantial number false positive results in low NTRK fusion-positive tumour types? 14. Is it appropriate to limit testing to avoid false positive results and the associated costs?
Background/description of issue	In addition to issues identifying the prevalence (see issue 1) and uncertainty about the testing pathway (Issue 2), there is uncertainty about the ability to correctly identify NTRK fusions through diagnostic accuracy. **The company** does not explore issues of diagnostic accuracy. They state that the sensitivity and specificity rates reported for a representative NGS assay (Oncomine Focus Assay) are 100% for gene fusions.
Why this issue is important	The issue of diagnostic accuracy may explain some of the heterogeneity of response across different solid tumour types (see issue 9), tumour types with low NTRK fusion prevalence are more likely to have a false positive NTRK fusion test, in which case response would not be expected. Diagnostic accuracy will also affect the practicalities of implementing tests to identify eligible patients.
Technical team preliminary judgement and rationale	There is uncertainty about the accuracy of testing for NTRK fusion. The technical team would like an estimate of the sensitivity and specificity of NGS testing of NTRK fusions as there is a likely to be a high level of uncertainty about the practicality of testing in low prevalence tumour types, even with tests of high sensitivity and specificity.
Summary of comments	**Comments received from the company:** Genomic Laboratory Hubs are still in the process of becoming operational so there is no single representative NGS-based offering demonstrating a given diagnostic accuracy. There is variation in potential diagnostic accuracy across NGS-based assays and this is an area of potential further evidence generation, particularly in the specific context of NTRK fusion detection in clinically derived samples. Recent literature is in agreement that the sensitivity and specificity of NGS-based approaches is generally considered to be high. Utilisation of a RNA/DNA hybrid panel NGS approach has the potential to maximise sensitivity and specificity, therefore minimising the risk of false negatives/positives. Further, the utilisation of an IHC pre-screen prior to NGS confirmation would also minimise false-negatives and is alignment with the most recent ESMO guidance. In light of the potential tumour agnostic indication being sought for entrectinib, it would not be appropriate to limit testing and exclude certain tumour types. **Comments received from NCRI-ACP-RCP-RCR:** Utility of NGS is primarily determined by the quality of the sample input. Experts note that analysis failure is seen in approximately 20% of cases that are screened by NGS. Sensitivity and specificity is very high if the analysis is successful. Specificity of NGS expected to be high (100%). The sensitivity is uncertain and from clinical judgement for DNA NGS may be in the region of 70% (unsubstantiated). RNA testing would provide higher sensitivity (if assays are optimised for FFPE tissue). The likelihood of false positives is very low for DNA- and RNA-based NGS.

	Disagree with the suggestion that the differential responses in the clinical trial may be due to false positive NTRK results. Note that the issue is about identifying false positives, rather than avoiding false negatives. A RNA-based NGS approach would be optimal for identification of fusions but would depend on good quality FFPE tissue. A combined DNA/RNA NGS approach could be used to reduce the risk of false negatives (and would also maximise the chance of identifying other molecular alterations that could be actionable [treated with a standard-of-care treatment or guide clinical trial selection]). An alternative would be a 2-step diagnostic approach such as IHC followed by confirmatory DNA/RNA testing but this would not be in line with NHS England ambitions to roll out NGS. **Comments received from GIST Support UK:** Diagnostic accuracy of NGS testing is very accurate. **ERG considerations on new evidence received during technical engagement:** Considerable uncertainty in the specificity and sensitivity of NGS for detecting NTRK fusions as the analytical validity of NGS is likely to differ between each NTRK gene, fusion partner and tumour type. DNA-based NGS may detect DNA-level fusions that may not result in Trk protein expression.If DNA-based NGS is used, RNA-based NGS is required to confirm a transcribed *NTRK* fusion, and consequently Trk protein expression. This will limit the impact of false positive results and in turn, may improve the efficacy of entrectinib. The study that the company cite that states that the RNA-based NGS Oncomine Focus Assay has 100% sensitivity and specificity is based on a very limited sample of fusion positive patients. The reliability of RNA-based tests is highly dependent upon the quality of the RNA sample available. Degradation in sample quality can increase the numbers of false negatives. Consider RNA-based NGS to likely be the most appropriate testing strategy to identify novel fusions and correctly identify more individuals eligible for entrectinib. Single-test strategies with no confirmatory test may increase the number of false-positive patients going on to have entrectinib. The prevalence of NTRK fusions in the different tumour types will also impact on the number of false positives; for tumours where *NTRK* fusions are rare, the likelihood of false positives is higher.
Technical team judgement after engagement	The technical team is anticipating receiving further guidance on the diagnostic accuracy of the various potential NGS options ahead of the committee meeting.

Issue 6 – Distribution of tumour types

Questions for engagement	15. Are the ERG's estimates of the distribution of tumour types reflective of what would be seen in clinical practice in England? If no, what is the likely distribution of NTRK positive tumour types in clinical practice?
Background/description of issue	The distribution of the tumour types in the trial is used in the comparator arm of the model (company's and ERG's) to weight the comparator effectiveness, utilities and costs for each tumour type. **The company** assumes that the distribution of the 10 tumour types in the efficacy evaluable dataset are reflective of the distributions seen in clinical practice in England. It uses this distribution to estimate a weighted set of outcomes for the comparator arm in its base-case analysis. **The ERG** is concerned that the estimate of cost-effectiveness is being driven by the proportion of tumour types included in the company's efficacy evaluable dataset. The ERG note that in applying the distribution of tumour types seen in the efficacy evaluable dataset, it is assumed that the trial population is reflective of clinical practice. The ERG determined an alternative distribution of tumour types using a dataset of over 166,000 tumour samples. This is the same dataset used by the ERG to calculate the prevalence of NTRK gene fusion (see issue 1). They consider this dataset to be more representative of clinical practice as it is based on a larger sample than the company's original estimate of distribution. The alternative distributions were used in a scenario analysis run by the ERG. The ERG were not provided with individual patient data and were unable to analyse the effect of this distribution on the entrectinib arm. **Table 6: Distribution of tumour types in the entrectinib integrated efficacy evaluable analysis set and an alternative ERG distribution (Reproduced from Table 29 in ERG report)**

Tumour Type	Proportion in company submission	ERG
Sarcoma	24%	
Non-small cell lung cancer	18%	
MASC	13%	
Breast	11%	
Thyroid	9%	

	<table><tr><td>Colorectal cancer</td><td>7%</td><td>[redacted]</td></tr><tr><td>Neuroendocrine</td><td>6%</td><td>[redacted]</td></tr><tr><td>Pancreatic</td><td>6%</td><td>[redacted]</td></tr><tr><td>Gynaecological</td><td>4%</td><td>[redacted]</td></tr><tr><td>Cholangiocarcinoma</td><td>2%</td><td>[redacted]</td></tr><tr><td colspan="3">ERG, evidence review group.; MASC mammary analogue secretory carcinoma</td></tr></table> **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlights that the treated population in England will significantly differ from the distribution of tumour types included in the entrectinib clinical trials. They note in particular the high proportion of people included in the entrectinib trial that have MASC.
Why this issue is important	The distribution of tumour sites is used to weight a number of variables in the cost-effectiveness analysis. Including the ERG's alternative distribution of tumour types increases the company's base case ICER by around £17,000 (£69,747 instead of £52,609). This scenario was associated with greater incremental costs and lower incremental QALYs than the base-case analysis. The difference in cost was mostly driven by the large decrease in the proportion of people with sarcoma. Sarcoma is associated with higher total costs than the other tumour types in the established management arm.
Technical team preliminary judgement and rationale	The technical team would like to see the analyses conducted with the ERG's distributions applied to the entrectinib arm. These distributions are considered to be more reflective of what would be seen in clinical practice in England than the distributions seen in the company's clinical trials. It also aligns with the technical team's preference to use the Foundation Medicine Inc. database to estimate the prevalence of NTRK fusions (see issue 1).
Summary of comments	**Comments received from the company:** Reasonable to expect that the tumour types included in the entrectinib trial population is reflective of clinical practice given that the entrectinib trial population was recruited through an extensive testing programme. Likely distribution of NTRK gene fusion-positive tumour types in England may only be definitively answered once comprehensive NGS-based testing is implemented in an unbiased fashion across all cancer patients. The entrectinib data has been re-weighted according to the number of patients with each tumour type calculated by the ERG, as the ERG had done for the comparator data. The company reported that [redacted]. Data presented in table 4 of the company's response to technical engagement. **Comments received from NCRI-ACP-RCP-RCR:** FMI data set may be the best guide as to the relative prevalence of NTRK fusion. The figures look mostly reasonable with the following exceptions: Breast cancer: would anticipate breast cancer patients as a whole to represent a smaller proportion of patients in the model, perhaps around [redacted] The tumour type included in the analysis is the rare secretory breast cancer subtype. Thyroid cancer: seems a high proportion of patients. Question whether the histological sub-type of patients is known in the FMI database and whether this is representative of the UK population. Colorectal cancer: would anticipate this proportion to be lower Cholangiocarcinoma: this is estimated to be [redacted] but this seems unlikely given the [redacted]% prevalence. **Comments received from GIST Support UK:** It is estimated that approximately 14% of wild-type GIST patients are "quadruple negative". It is estimated that this represents around 17 people per year with around 50% of these likely to be clear of the disease after surgery and will not require therapy. This gives 8 people who will be screened for NTRK fusions. **ERG considerations on new evidence received during technical engagement:** Does not agree with the company that the trial can be considered a representative sample and a true reflection of the likely eligible population in the NHS. Unsure whether further data collection will allow more accurate estimates of prevalence to be determined given that it is doubtful that any data collection exercises could collect a larger sample than the FMI database. Appreciates the company's attempt to reweight the costs and outcomes in the entrectinib arm based on the distribution in the FMI database, as the ERG did for the comparator arm. Although, have not been able to validate the reweighting of the KM data as it is not described in the company's response or included in the economic model.
Technical team judgement after engagement	The technical team maintain that the distributions seen in the company's clinical trials are not reflective of the likely distribution that will be seen in clinical practice in England. This could be addressed through further data collection if entrectinib is recommended in the CDF.

Issue 7 – Unrepresented NTRK gene fusion positive solid tumour types

Questions for engagement	16. Are the results of the entrectinib studies that include 13 different NTRK positive tumours sites generalisable to all NTRK-fusion positive tumour types?
Background/description of issue	There is currently limited evidence available on the tumour types that harbour NTRK fusions. The entrectinib clinical trials included 13 tumour types. **The company** included 13 tumour sites in their base case evidence base. **The ERG** report that there is no representation of some other tumour types in the evidence, because there are at least a [redacted] that will be covered by the anticipated marketing authorisation based on those tumour types in which an *NTRK* fusion has been identified. Clinical advisors to the ERG suggested that it is plausible that NTRK gene fusions could potentially be present in 400+ tumour types. The ERG have used a Bayesian Hierarchical Modelling (BHM) framework to group and analyse response data to explore the potential heterogeneity in effects across tumours (see section 4.3.1 of the ERG report). The BHM framework can also be used to predict the response probability that would be expected in a tumour type that is not represented in the data set (further description and results presented in issue 9).
Why this issue is important	The omission of these unrepresented tumour types has numerous implications for the model and potentially impacts on a number of inputs including clinical effectiveness of both entrectinib and the comparator, testing and treatment costs and health state utilities. This issue could limit the generalisability of the clinical trial results to clinical practice in England.
Technical team preliminary judgement and rationale	Lack of any data for a potentially high number of other tumour sites means there is substantial uncertainty about entrectinib's clinical and cost-effectiveness. The technical team would like to see scenarios where the impact of the unknown effectiveness of entrectinib in these other tumour sites is explored. Using the Bayesian hierarchical modelling framework set out by the ERG (see issue 9), the entrectinib individual patient data could potentially be used to inform estimates of PFS and OS for the unknown tumour sites. These data could then be included in the company's partitioned survival model to represent the proportion of unknown tumour sites alongside the data from the entrectinib clinical trials. Best and worst case scenarios could be run using the range of clinical effectiveness from the Bayesian hierarchical model.
Summary of comments	**Comments received from the company:** Acknowledge the uncertainty associated with the existing entrectinib clinical trial data in terms of the potentially unrepresented tumour types. Maintain that the existing data for entrectinib demonstrates compelling efficacy across tumour types. Did not investigate tumour type heterogeneity in time-to-event outcomes using the ERG's Bayesian Hierarchical modelling framework because consider the current data to not be robust enough to give any meaningful estimates using this approach. CDF data collection proposals include collecting data in patients with tumour types that are not currently represented in the clinical trial data. **Comments received from NCRI-ACP-RCP-RCR:** There is likely to be clinically significant variability in response rates between tumour sites which will have widely varied genomic alterations in addition to NTRK fusion. Data do not exist to support this but it would be reasonable to assume a similar efficacy rate would be seen given the results to date across 13 different tumour types. It is reasonable to consider entrectinib and larotrectinib to have a similar level of efficacy so data from other tumour types included in the larotrectinib studies could be informative. **ERG considerations on new evidence received during technical engagement:** Maintain that there is no evidence that response rates are identical between tumour types or that similar response rates should be expected in unrepresented tumours. However, acknowledge that the observed and estimated response rates on entrectinib may be considered promising across the tumour types included in the entrectinib studies.
Technical team judgement after engagement	The technical team maintain that there is substantial uncertainty about entrectinib's clinical and cost-effectiveness given the lack of any data for a potentially high number of other tumour sites and an observed between-tumour heterogeneity in the known tumour sites.

Issue 8 – Primary CNS tumours and paediatric tumours

This issue was resolved at technical engagement and is addressed in Table 13.

Issue 9 – Heterogeneity of response across different solid tumour types

Questions for engagement	17. Is a uniform level of response to entrectinib across the different tumour types with a NTRK fusion reasonable?

	18. Is the Bayesian Hierarchical Modelling (BHM) framework an appropriate method to capture the heterogeneity of response across different solid tumour types for this appraisal? 19. Would it be appropriate to apply the BHM framework to explore the heterogeneity in the time to event outcomes? 20. Are the response rates in tumour sites not represented in the trial data suitable for consideration?
Background/description of issue	Previous basket trials have shown heterogeneity of response across tumour types and previous trials have shown heterogeneity in effectiveness of chemotherapeutic agents across tumour types (see section 4.2.1 of ERG report). **The company** assume that each of the solid tumour types will have identical response rates when treated with entrectinib. This allows them to generate a pooled response estimate across each of the tumour types included in their efficacy evaluable dataset, [redacted] response. This approach does not take into account the potential for heterogeneity of response across different tumour types. The company have not explored alternatives to this assumption. **The ERG** considered each of the tumour types as a "basket" and analysed the response data using a Bayesian Hierarchical Modelling framework to explore the potential heterogeneity in effects across different tumour types (see section 4.3.1 of ERG report). The ERG used this method to estimate posterior probabilities of response (this is the probability of response after all evidence or background information has been taken into account) for each tumour type, as well as a pooled posterior probability of response across all tumour types, accounting for the potential lack of uniformity of effect across tumours. This is the same method that the ERG used to predict the response probability that would be expected in solid tumour types that were not included in the entrectinib trials (see issue 7). Using this method, the ERG identified that while all distributions overlap, the distributions of response for [redacted] and the 95% confidence interval suggest that response rates [redacted] are plausible. These tumour types also have the lowest probabilities of having a response rate greater than 30%. For unrepresented tumour types in the entrectinib trial data, the 95% confidence interval for the response probability is wide, showing that the response probability could be as low as [redacted], or as high as [redacted]. The ERG note that heterogeneity in time to event outcomes (PFS and OS) can be explored using the BHM framework in a similar way to what is described above, however it may not provide useful results for this appraisal (see ERG report section 4.3.2). This is because the survival data is immature (with [redacted] and [redacted] of patients having experienced a PFS and OS event respectively) and there is a small population with each tumour type. PFS and OS data by tumour type and line of therapy were not made available to the ERG but as more data become available, this method could be used to determine the extent of heterogeneity in PFS and OS across different tumour types. This would also allow predictive distributions of PFS and OS to be used to inform the survival of people with unrepresented tumour types.
Why this issue is important	While the ERG found that response rates obtained using the BHM framework were similar to those observed when equal response probabilities are assumed (company's assumption) [around [redacted] mean estimated response] they identified considerable uncertainty in the level of heterogeneity of response rates across tumour types. The ERG note that it is unclear how heterogeneity in response outcomes impacts on survival outcomes, and consequently cost-effective estimates, but these analyses illustrate the potential for heterogeneity between tumour types. This is important as the clinical trial evidence does not include all tumour types known to have NTRK fusions (see issue 7).
Technical team preliminary judgement and rationale	Lack of supporting evidence means there is considerable uncertainty about the assumption that entrectinib's treatment effect is homogeneous across all tumour types. TKIs generally have heterogeneity across tumour sites. The technical team would like to see more evidence to support the assumption of homogeneous response. They would like to see additional descriptive data from the company, for example Kaplan-Meier data for subgroups, 6 and 12 month PFS and OS rates by tumour type and line of therapy. This will help to determine where there appear to be similarities and differences in PFS and OS and subsequently will help determine whether hierarchical modelling is appropriate. If appropriate, the technical team would like to see the BHM framework used to explore the heterogeneity in the time to event outcomes (i.e. OS and PFS).
Summary of comments	**Comments received from the company:** Maintain that the evidence available currently suggests that entrectinib is active across the studied range of tumour types and will be active in as yet unknown tumour types. Acknowledge the uncertainty in the assumption. Consider existing subgroup data for entrectinib not robust enough at the specific tumour type level for reliable modelling to assess tumour or response heterogeneity and note that even within the largest subgroup, soft tissue sarcoma, the population is potentially very heterogeneous. Consider the time-to-event data to not be robust enough to use the BHM to explore heterogeneity in these outcomes across tumour types and the individual tumour type subgroups are too small to model outcomes using the BHM framework. PFS and OS KM curves provided by tumour type. Report that they show that any attempt to model time-to-event outcomes using a BHM framework is unlikely to result in useful information. Consider the ERG's BHM-predicted response rates in unknown tumour types of [redacted] to be so broad that its usefulness for decision making is limited. Note that they were [redacted] CDF data collection proposal includes collecting further data in tumour types that are currently represented in the available data and new data in patients with tumour types that are not currently represented in the clinical trial data.

	Comments received from NCRI-ACP-RCP-RCR: Currently uncertain whether responses would be heterogeneous and the proposed additional data and modelling are appropriate. Generally fusions are a strong oncogenic driver of disease (and often in the absence of other genomic alterations) compared with a somatic mutation. With a somatic fusion there can be considerable heterogeneity between disease types in terms of contribution to disease progression (and often with co-mutations contributing). Differential responses are frequently seen with tyrosine kinase inhibitors. Clinical expert opinion then suggests that inhibition of a signalling pathway (TRK), and thus response to treatment, being driven by a genomic fusion may be more homogeneous. **Comments received from GIST Support UK:** Paediatric patients with NTRK fusions should be allowed access to entrectinib. **ERG considerations on new evidence received during technical engagement:** Agree that current evidence suggests a promising level of response to entrectinib across the tumour types included in the studies. However, evidence showing that the response is uniform, or that response rates would be as the same in tumour types not included in the entrectinib studies has not been presented. Do not agree with the company's view that the data are not robust enough to estimate the between-tumour heterogeneity.
Technical team judgement after engagement	The technical team maintain that the lack of supporting evidence means there is considerable uncertainty about the assumption that entrectinib's treatment effect is homogeneous across all tumour types.

Issue 10 – Constructing a comparator arm

Questions for engagement	21. Is the company's comparator arm suitable for decision making? 22. Is it appropriate to use non-responders as a proxy for patients not having an active treatment to generate a comparator arm for this appraisal?
Background/description of issue	The entrectinib clinical trials did not include a control arm. There is difficulty in creating a basket comparator for established management without entrectinib. **The company** generated a comparator arm by identifying PFS and OS comparator data for established management by searching NICE pathways to identify NICE approved comparators for each of the tumour types represented in the entrectinib clinical trials. PFS and OS values were extracted from the clinical effectiveness data presented within the committee slides or the company submission. Median PFS and OS for each tumour type were averaged and then pooled to calculate mean overall PFS and OS across all tumour types, weighted by the distribution of each tumour type within the trial population (see issue 6). For cholangiocarcinoma and gynaecological cancers (ovarian and endometroid), data were not extracted from NICE recommended treatments and instead an average of the PFS and OS from the other sources for the other tumour types were used. For MASC where there are no NICE recommended chemotherapies, the company used surrogate trial data for best supportive care from a publication for OS and an average from the other sources for the other tumour types for PFS. At clarification the company state that there are no robust comparator data for paediatric tumours, particularly infantile fibrosarcoma. In their scenario analysis that includes primary CNS and paediatric tumours, the paediatric primary CNS patients were grouped with the adult primary CNS patients for the weighted comparator outcomes and costs since common comparators are assumed. The company did not account for tumour types unrepresented in the entrectinib clinical trials (see issue 7). **The ERG** deem the methods used to identify, select and combine comparator data to be inappropriate and consider the comparator data in the company model to be highly unreliable. The ERG consider it preferable for the company to extract PFS and OS estimates directly from the survival curves for each of the NICE recommended treatments rather than extracting them from the committee slides or company submission. Further, they consider it preferable to use inputs from a robust meta-analytical techniques rather than naively pooling unweighted means of medians. They consider this approach to be statistically inappropriate in circumstances where more than one clinical trial informed the technology appraisal for the NICE recommended treatment. The ERG highlight that the company's method of pooling the PFS and OS estimates from the different therapies implies an even distribution of people receiving each of the treatment options within a given line of therapy, which does not reflect clinical practice. The ERG highlight that the prognosis for people with a given tumour type such as cholangiocarcinoma, gynaecological cancers and MASC may differ significantly from the people with other unrelated tumour types and so using an average of the PFS and OS outcomes from the other tumour types is not appropriate. The ERG have suggested two further approaches. The company could: 1. Use effectiveness data on non-responders as a proxy for patients not receiving an active treatment. The rationale behind this approach is that people in which no response is observed represent those with a lack of treatment effect and therefore are representative of a counterfactual where no effective therapy exists. 2. Compare the outcomes for people on entrectinib with their outcomes on the previous line of therapy. This approach involves taking the average time to progression on the previous line of therapy and comparing it to the average patient time to progression when treated with entrectinib and estimating a ratio. The ERG note the disadvantages of these approaches but consider both approaches to be potentially valid.

	The ERG conducted the non-responder as controls method as an alternative approach to modelling the decision problem and present the analyses as exploratory analyses (see issue 14). They noted that the time, resource and data requirements are significantly higher to conduct the second approach. The ERG's responder-based model enabled the tumour types that were unrepresented in the entrectinib trials to be accounted for (see issue 7).
Why this issue is important	A lack of direct evidence adds uncertainty to the true comparative efficacy of entrectinib and established management. The lack of control group in the entrectinib trial evidence means that the relative effectiveness and safety of entrectinib compared with relevant alternative cancer therapies are highly uncertain. The ERG's alternative responder-based approach that helps to overcome the issue of constructing a control arm is discussed further in issue 14.
Technical team preliminary judgement and rationale	The technical team recognises the difficulty in constructing a comparator arm for this appraisal. It considers the company's pragmatic approach to be intuitive but recognises the following issues: • Comparator population is not consistent with the entrectinib population for CNS metastases and other potential prognostic factors (see issue 12); • Comparator population is not reflective of the population seen in clinical practice if the comparators have not been selected at the appropriate line of treatment. The ERG's responder-based method overcomes some of the issues with the company's pragmatic approach but also has some issues with the non-responders potentially having a different prognosis to the responders and receive different subsequent therapies and also generalisability to the population seen in clinical practice, particularly if the clinical trial population is not generalisable. This alternative model is also limited by the amount of data that is available given the small numbers in the entrectinib clinical trials and maturity of the data. The technical team consider the responder-based approach to be equally appropriate but note that the results from the company's analysis are similar to the results from the responder-based approach.
Summary of comments	**Comments received from the company:** Maintain that the modelled comparator arm gives a reasonable and pragmatic basis on which to assess the comparative effectiveness of entrectinib. Conducted an [redacted] Do not consider it appropriate to use non-responding patients as a proxy comparator arm and consider this approach to be deeply flawed and inappropriate for decision making. Reasons for this include that: • sample sizes in each arm are too small to give a meaningful comparison • different measures of response have been used for groups within the responder analysis • non-responding patients may still benefit from entrectinib but may not reach the threshold for response • key prognostic factors are not balanced between the groups • overall survival data for entrectinib is still immature. The company consider that further data collection through a period in the CDF will help to resolve some of the uncertainty in the assessment of entrectinib. **Comments received from NCRI-ACP-RCP-RCR:** Company's comparator survival data seems reasonable for the 'beyond standard of care' cohort. Agree with the ERG's assessment of the flaws associated with the company proposed comparator arms. Use of real-world data would be preferable but data availability likely to be highly limited. The non-responder analysis would have some value but may introduce bias as there will be biological reasons why patients do not respond which may in themselves be prognostic (likely associated with adverse prognosis). This approach is also potentially flawed as even in the absence of tumour response, there is still the possibility/likelihood that disease stabilisation may occur and thus impact on PFS and OS. Assessing time to PFS on previous treatment would be a valuable approach but acknowledge these data may not be available if not specifically collected in the supporting trials. **ERG considerations on new evidence received during technical engagement:** Maintain that the comparator data used to inform the company model is highly unreliable and may not be suitable for decision making. Company's intra-patient analysis may produce a more reliable estimate of mean/median PFS in the comparator arm than the original modelled comparator, or the ERG's non-responder analysis. There are several assumptions associated with this approach. Consider the following two assumptions to be very strong and hard to verify and note these as a limitation of this approach: 1. Re-treatment with drugs that have failed on a previous line of therapy would yield the same time to further disease progression as at the previous line of therapy. 2. Risk of death in the pre-progression state is the same across the lines of therapy, and negligible. Note that while [redacted], it is unclear whether extrapolation of a curve fitted to the intra-patient comparator would affect the ICER. The ERG highlight that the company have not adequately described the uncertainty in the comparator PFS estimates.

Technical team judgement after engagement	The technical team consider the 3 approaches to constructing a comparator arm (naive-weighted comparison, intra-patient analysis for TTNT and non-responder analysis) to be useful because they each have different biases. If the results of the different approaches concur then this potentially provides strong evidence. Otherwise, committee will base their decision on the uncertainty of each model.

Issue 11 – Comparator treatments

Questions for engagement

23. Is the company's modelled comparator population representative of the eligible population in England?
24. Should the comparator data be matched to the position that entrectinib is given in the efficacy evaluable dataset?

Background/description of issue

The comparator dataset is drawn from a population at a later line of therapy than entrectinib was given in the clinical trials (see issue 2).

The company's comparator dataset drew predominantly from people in later lines of therapy, with data from 7 of the 13 tumour types, representing 57% of people, being drawn from 2nd line or beyond. This differs to the entrectinib arm where [redacted] of the integrated efficacy population received entrectinib as first-line therapy (see issue 2). The company also include multiple comparator treatments for the same tumour type but at different lines of therapy.

The ERG notes that the company did not account for treatment being given at different lines for the same tumour type and does not attempt to match the comparator data used to the position that entrectinib was given in the integrated analysis dataset for each tumour type. The company's method of pooling the PFS and OS estimates from the different therapies at different lines of therapy assumes that there is an equal distribution of people receiving different lines of therapy within a tumour type. The ERG does not consider this to be in line with clinical practice.

The ERG highlighted that the company placed entrectinib as first line treatment for soft tissue sarcoma but comparator trials included people at second line and beyond.

The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund considers some of the comparators used in the company's blended comparator to not reflect the treatments that entrectinib would displace in clinical practice.

Table 8: Treatments that entrectinib will displace in clinical practice by tumour type if entrectinib receives a positive recommendation

Tumour type	Company (from Table 3, company response to clarification)	NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund
MASC	Not reported	Chemotherapy
Soft-tissue sarcoma	1st line & beyond: doxorubicin 2nd line & beyond: trabectedin	Chemo-resistant sarcoma: trabectedin Chemo-sensitive sarcoma: best supportive care
Pancreatic cancer	1st line: gemcitabine with nab-paclitaxel, gemcitabine monotherapy, FOLFIRINOX	If entrectinib used 1st line: chemotherapy If entrectinib used 2nd line: best supportive care
Cholangiocarcinoma	Not reported	If entrectinib used 1st line: chemotherapy If entrectinib used 2nd line: best supportive care
Gynaecological cancers	Not reported	Pegylated liposomal doxorubicin or topotecan
NSCLC	2nd line & beyond: docetaxel, nintedanib with docetaxel	Docetaxel ± nintedanib

<table>
<tr><td></td><td>
<table>
<tr><td>Breast</td><td>2nd line: capecitabine
3rd line & beyond: eribulin, vinorelbine, gemcitabine with paclitaxel</td><td>Best supportive care or vinorelbine or eribulin (depending on what has been used previously)</td></tr>
<tr><td>Thyroid cancer</td><td>2nd line & beyond: best supportive care</td><td>Best supportive care</td></tr>
<tr><td>Colorectal cancer</td><td>2nd line: FOLFIRI, irinotecan
3rd line: trifluridine-tipiracil, best supportive care</td><td>If entrectinib is used 3rd line: Trifluridine/tipiracil
If entrectinib is used 4th line: best supportive care</td></tr>
<tr><td>Neuroendocrine carcinoma</td><td>1st line: everolimus, best supportive care</td><td>Best supportive care</td></tr>
</table>
</td></tr>
<tr><td>Why this issue is important</td><td>Line of therapy is an important determinant of prognosis. The ERG consider it very likely that estimates of PFS and OS for the comparator are confounded in favour of entrectinib. Selecting the most appropriate comparator treatments is important as it also impacts on the costs and utility values included in the economic model for the comparator arm.</td></tr>
<tr><td>Technical team preliminary judgement and rationale</td><td>The technical team would like to see the comparator data matched to the position that entrectinib is likely to be given in clinical practice.</td></tr>
<tr><td>Summary of comments</td><td>Comments received from the company:
Acknowledge the uncertainty regarding appropriate comparators.
Provided a scenario analysis using the NHSE and NHSI’s alternative recommended positioning for entrectinib as third-line treatment for breast and colorectal cancer and neuroendocrine tumours (see issue 2). This was executed in the economic model by removing capecitabine for second-line breast cancer, FOLFIRI and irinotecan for second-line colorectal cancer, and everolimus for the treatment of neuroendocrine tumours. Using the NHSE and NHSI’s alternative recommended positioning for entrectinib decreases the company’s updated base case ICER by around £1,500 (£47,894 instead of £49,358).
Note that the best supportive care outcomes used for neuroendocrine tumours are heavily confounded by cross-over to active therapy, as acknowledged in NICE TA449 and TA539. Arbitrary reduction of overall survival for BSC in NETs to 24 months reduced the ICER by around a further £1,900.
Not appropriate to match the comparator data to the position that entrectinib is given in the entrectinib clinical trials. Line of therapy at which entrectinib is given is not a prognostic factor and this means that the entrectinib outcome data can be considered representative of outcomes seen at lines of therapy matched to the modelled comparators, see issue 2.
Consider that further data collection through a period in the CDF will help to resolve some of the uncertainty regarding the positioning of entrectinib and the most appropriate comparators to include in the assessment of entrectinib.
Comments received from NCRI-ACP-RCP-RCR:
Company’s modelled population seems reasonable although the lines of therapy used for the comparator arm by the company do not fully reflect current practice compared with the line of treatment entrectinib was administered in the clinical trials. Comparator data should be matched to the position that entrectinib is given in the efficacy evaluable dataset but acknowledge that the intended positioning of entrectinib may be different from that used in the clinical trials.
ERG considerations on new evidence received during technical engagement:
Company’s scenario analysis including NHS England and NHS Improvement’s preferred comparators was executed incorrectly. The company removed all comparator drug and administration costs for breast and colorectal cancer and neuroendocrine tumours. The ERG re-ran this analysis only removing capecitabine in second-line breast cancer, FOLFIRI and irinotecan in second-line colorectal cancer, and everolimus in neuroendocrine tumours.
Using the NHSE and NHSI’s alternative recommended positioning for entrectinib and suggested comparators decreases the company’s updated base case ICER by £64 (£49,294 instead of £49,358).
The ERG consider the company’s new scenario a reasonable reflection of the position of entrectinib. The ERG maintain that there are substantive uncertainties regarding the positioning of entrectinib, see issue 2 for details.</td></tr>
</table>

Technical team judgement after engagement	The technical team considers the ERG's correction to the company's scenario to be reasonable for decision making with the information that is available at this time. The position that entrectinib is used in the treatment pathway determines what comparators should be used in the analysis. The uncertainty around the positioning in the treatment pathway could be resolved through further data collection if entrectinib is recommended in the CDF.

Issue 12 – Prognostic factors

Questions for engagement	25. Is it appropriate to adjust the comparator data to account for worsened prognosis of CNS metastases? 26. Is it appropriate to adjust the comparator data to adjust for NTRK fusions? If yes, are there any additional evidence sources which could be used to inform whether presence of an NTRK fusion is prognostic?
Background/description of issue	All people included in the entrectinib trials were NTRK fusion positive and 20.4% of the entrectinib efficacy evaluable population had CNS metastases. It is unknown how prevalent CNS metastases are in the comparator population. Only a small proportion of people in the comparator dataset are likely to be NTRK fusion positive. There is evidence to suggest that NTRK fusions are prognostic, though available evidence is limited. **The company** notes that the NTRK fusion status and prevalence of CNS metastases may not be matched between the entrectinib and comparator arm. They claim that this may lead to the comparator OS and PFS values being overestimated. The company state in its submission that the prognosis of people with NTRK fusion positive solid tumours is worse than for people who have tumours that do not harbour an NTRK gene fusion. They also state that the presence of CNS metastases is associated with higher disease burden, reduced life expectancy and poorer quality of life compared with other types of metastases. In its base case, the company has not adjusted the effectiveness data for the comparator to reflect any possible impact of NTRK fusions or CNS metastases altering patient prognosis. The company presents a scenario analysis which uses evidence from people with colorectal cancer to adjust the comparator data to account for NTRK fusion positive status. A hazard ratio of 2.33 is applied to the comparator data as this paper reported that people with NTRK, ALK or ROS1 rearranged metastatic colorectal cancer tumours had a poorer median OS when compared with people who had tumours without the genomic alterations (15.6 months versus 33.7 months). They also present a scenario analysis to adjust for the impact of CNS metastases in the comparator arm. **The ERG** have identified literature that suggests that the prognostic value of NTRK may be different across tumour types. The ERG consider that the company's scenario analysis be interpreted with caution. The ERG found insufficient evidence to support the claim that comparator OS and PFS values may be overestimated because of NTRK impact on prognosis and CNS metastases. The ERG's clinical adviser suggested that variability in prognosis across the different tumour types is possible and prognosis is likely dependent on the role NTRK fusions play in that specific tumour type. The ERG investigated the prevalence of CNS metastases in the comparator data but found that most of the comparator trials did not report baseline prevalence of CNS metastases. The ERG also note that the company does not report any baseline characteristics for the comparator arm. It is uncertain whether the characteristics between the two arms are comparable for the characteristics that are commonly prognostic such as age and ECOG status. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlights that the paper cited by the company does not include a pure NTRK gene fusion group and the incidence of NTRK gene fusion in colorectal cancer is thought to be <1%. NHS England and NHS Improvement recognises that there may be a difference in outlook for people with incurable NTRK fusion positive metastatic cancer but there is no robust evidence to support this at present. They note that the presence of cerebral metastases has an adverse prognosis to people having systemic therapy.
Why this issue is important	Given that only a small proportion of patients in the comparator data set are likely to be NTRK fusion positive, it is a potentially important source of confounding bias. Survival outcomes for the comparator may not reflect the true survival outcomes for this group if NTRK fusions are themselves prognostic and if the impact of CNS metastases have not been appropriately captured. Including the company's adjustment to reflect poorer prognosis when tumours are NTRK positive decreases the company's base case ICER by around £16,900 (£35,732 instead of £52,609). Including an adjustment for CNS metastases in the comparator arm decreases the company's base case ICER by around £5,600 (£46,981 instead of £52,609). The company also consider this issue to be important when considering whether entrectinib meets the end-of-life criteria (see issue 21).
Technical team preliminary judgement and rationale	The technical team consider it appropriate to adjust for factors that are known to impact on prognosis. The model results are sensitive to the inclusion of an adjustment for CNS metastases in the comparator arm. The technical team would like to see an explanation of how the adjustment was implemented in the model. The technical team agree that there could be differences in prognosis for tumours with NTRK gene fusions but the current evidence is not robust enough to support this. Therefore, it is not appropriate to include an adjustment to reflect poorer prognosis when tumours are NTRK positive.
Summary of comments	**Comments received from the company:** Appropriate to adjust comparator data to account for the worse prognosis associated with the presence of CNS metastases given the relatively relatively high prevalence of CNS metastases in the original base-case population (20.4%).

	Original submission included a scenario adjusting for the possible impact of brain metastases in the comparator arm. Applying this adjustment to the new base-case population reduces the ICER by £4,350 (£46,981 instead of £52,609). Literature review conducted to identify any new literature on the prognostic implications of the presence of NTRK gene fusions. The search did not identify any further literature beyond that previously cited in the company's initial submission and ERG report. Appropriate to consider a scenario where the comparator population prognosis is worsened. In particular, two of the three studies show a worse prognosis of NTRK fusions in thyroid cancer, and suggest that this should be taken into account for the comparator arm and end-of-life criteria (Issue 21). Further data collection through a period in the CDF will help address this uncertainty. **Comments received from NCRI-ACP-RCP-RCR:** Insufficient data to determine the impact of the presence of CNS metastases on prognosis in the NTRK fusion population. Comparator data may not include people with active brain metastases as usually not eligible to participate in clinical trials. Alternatively, people with brain metastases can be included in trials if local treatment is administered first to control the CNS disease. This in itself (localised brain treatment) could impact on patient outcome (favourably) and could therefore impact survival data. Limited data to determine the impact of the NTRK fusions on prognosis so not appropriate to adjust for this factor. **ERG considerations on new evidence received during technical engagement:** Company's attempt to adjust for differences in the prevalence of CNS metastases between the entrectinib and control arm is reasonable. Although most of the Technology Appraisals from which comparators were selected did not report the proportion of people with CNS metastases so it is potentially misleading to state that the prevalence of CNS metastases in the integrated analysis was significantly higher than the highest proportion reported in the comparator Technology Appraisals. Other differences that exist between the trial and comparator datasets have not been taken into account. Maintain that it is unclear from the evidence available whether NTRK fusions are in themselves prognostic, or whether it is their association which other prognostic factors in the presented studies that drives the observed difference in prognosis. Company's literature review of NTRK fusion prognosis did not identify any additional evidence to that presented in the original company submission.
Technical team judgement after engagement	The technical team maintain that it is not appropriate to include an adjustment to reflect poorer prognosis when tumours are NTRK positive as no additional evidence has been presented to support its inclusion. The technical team consider it appropriate to adjust for factors that are known to impact on prognosis and consider the company's scenario analysis including adjustment for brain metastases to be a reasonable exploratory scenario. The technical team is aware that other potential prognostic factors have not been accounted for in the analysis.

Issue 13 – Subsequent therapies

Questions for engagement	27. What percentage of people, who receive entrectinib, would be expected to receive subsequent therapy following disease progression? 28. What percentage of people, who receive established management, would be expected to receive subsequent therapy following disease progression? 29. Which subsequent therapies would be used and in what proportions? Please answer for following entrectinib and following established management. 30. How long would these people be expected to be treated with subsequent therapy? Please answer for following entrectinib and following established management.
Background/description of issue	In its model, **the company** has based the proportion of people receiving subsequent therapy post-progression in the entrectinib arm on the entrectinib clinical trial. This resulted in [redacted] of people in the entrectinib arm receiving subsequent therapy. No subsequent therapy was assumed for people receiving established management. At clarification, the company provided data that showed that [redacted] in the efficacy evaluation population received a subsequent chemotherapy after progression, a number of these were targeted therapies. However, a breakdown of the total number who received subsequent targeted therapies, or which patients received subsequent targeted therapies was not provided. People were assumed to receive subsequent treatment from progression until death in the company's model. At clarification, the company submitted two scenarios where people were assumed to receive subsequent treatment for 3 months and 6 months following progression. The duration of treatment with subsequent therapies in the clinical trial is not clear. **The ERG** considers it reasonable to assume that treatments that are displaced by entrectinib could be used as subsequent therapies following progression. It highlights that the mix of therapies that were used as subsequent therapies in the entrectinib clinical trial is very different to the treatments that would be displaced by entrectinib in clinical practice in England. The ERG considers it to be overly pessimistic to assume that people will have subsequent therapy after treatment with entrectinib from progression until death. They suggest that many people will move to best supportive care before death. This may reflect the exhaustion of treatment options or lack of fitness to continue to receive therapy.

	The ERG considers the appropriate duration of treatment with subsequent therapies to be for 6 months from progression but notes that this is an arbitrary assumption. This duration represents roughly half of the post-progression duration. The ERG includes a 6 months duration of treatment with subsequent therapy in the entrectinib arm in its base case.
Why this issue is important	Subsequent treatments likely to be received following progression within the NHS should be accounted for and modelled appropriately. This can have a significant effect on cost-effectiveness estimates. Greater health gains and increased costs can be expected from treatment with an active subsequent therapy compared with standard chemotherapy. Including greater health gains associated with active subsequent therapy but not including the increased costs for these active therapies in the entrectinib arm could bias the cost-effectiveness result in favour of entrectinib. Limiting the duration of subsequent therapy to 6 months decreases the company's base case ICER by around £12,500 (£40,093 instead of £52,609). Limiting the duration of subsequent therapy to 3 months decreases the company's base case ICER by around £12,800 (£39,849 instead of £52,609).
Technical team preliminary judgement and rationale	The technical team is concerned that the effectiveness data from the entrectinib clinical trial includes a reasonably sized proportion of people having subsequent therapies, some of which would not be available in clinical practice in England. As some of these are targeted therapies, it is likely that health gains are greater than what would be seen in clinical practice. The extent to which overall survival in the entrectinib arm is driven by the efficacy of subsequent therapies is unknown. The technical team would like to see a scenario where the entrectinib trial data has been adjusted to remove the effect of subsequent therapies from the entrectinib arm where appropriate. The model results are sensitive to the duration of subsequent therapy. Although the ERG includes 6 months in their base case they note that this is an arbitrary assumption. The technical team agree that the duration of treatment with subsequent therapies before best supportive care is highly uncertain.
Summary of comments	**Comments received from the company:** Maintain that the most reasonable estimate for the proportion of people receiving subsequent therapy following treatment with entrectinib is [redacted] based on data from the STARTRK-2 trial. Not appropriate to adjust the entrectinib trial data to remove the effect of subsequent therapies from the entrectinib arm because the standard methods require either a randomized trial or large patient numbers. It is also likely that there is some subsequent therapy effect in clinical practice. The likely impact of subsequent therapies after entrectinib to be small given the limited follow-up available. The assumption that 0% of people receiving established management will go on to receive subsequent therapy is conservative in terms of the cost-effectiveness of entrectinib given that some patients will be fit enough for further treatment options or eligible for a clinical trial. The potential cost impact of subsequent therapies may be reduced for people with NTRK gene fusion positive solid tumours given the potential role of clinical trials. Agree that 6 months duration of subsequent therapy is a plausible alternative to subsequent therapy until death, though even this may be overly pessimistic given the typical 4 to 6 cycle duration of many standard chemotherapies. Applying a 6-month subsequent therapy duration decreases the company's updated base-case ICER by around £9,500 (£49,358 to £39,890). **Comments received from NCRI-ACP-RCP-RCR:** The subsequent therapies given, and the percentage of people who would be expected to receive subsequent therapy following disease progression, will depend on entrectinib's position in the treatment pathway for each of the disease types. **Comments received from GIST Support UK:** All GIST patients who progress on entrectinib or established management would expect to receive subsequent therapy. Following entrectinib, subsequent therapy options could include another NTRK fusion inhibitor, imatinib, sunitinib, regorafenib, clinical trial. Following established management, subsequent therapy options could include imatinib, sunitinib, regorafenib, clinical trial. **ERG considerations on new evidence received during technical engagement:** No comments.
Technical team judgement after engagement	The technical team considers it reasonable to limit the duration of subsequent therapy to 6 months as it is more appropriate than a duration of subsequent therapy until death but recognise that this is an arbitrary duration. The technical team recognise that there is some outstanding uncertainty around the use of subsequent therapy and the treatments given because of the uncertainties around the position of entrectinib in the treatment pathway.

Issue 14 – Model structure

Questions for engagement	31. What is the most appropriate model structure for this appraisal?
Background/description of issue	A partitioned survival model populated with extrapolated PFS and OS data requires the availability of reliable, mature PFS and OS data for both the intervention and the comparator. **The company** structured its economic model as a partitioned survival model with 3 mutually exclusive health states: progression free, progressed and dead.

	At clarification, the company provided responder analyses as well as individual patient level response data for the efficacy evaluable dataset and those with primary CNS tumours and paediatric tumours. They did not provide this data for PFS and OS. The company noted in their submission that it is not considered appropriate to use data for trial-based non-responders to create a comparator population as it may provide an overly-optimistic estimate of incremental cost-effectiveness and introduce unnecessary uncertainty. They also considered the sample size of the non-responders to be too small to provide a meaningful comparator sample. **The ERG** suggest that a model structure built around response may have been more suitable to address the decision problem and could represent a more robust approach upon which to base long term outcomes. A response-based model may also better lend itself to characterize uncertainty resulting from any heterogeneity in the treatment effect. This is because fewer observations are required on response outcomes to draw meaningful conclusions about differences between tumour types, than would be required by time-to-event outcomes, such as PFS and OS. The ERG built a response-based model as an alternative approach and presented it as exploratory analysis for illustrative purposes (see section 6.5 of ERG report). This approach determines PFS and OS according to ORR. The ORRs across tumour types were estimated by the ERG's Bayesian Hierarchical Model (see section 4.3.1 of ERG report). The survival of non-responder patients was used to estimate survival predictions in the established management arm. The entrectinib arm was based on a weighted average of responder and non-responder survival predictions.
Why this issue is important	The comparator arm used in the company's economic model may not be appropriate for decision making (see issues 10 and 11). The response-based model overcomes these issues. The ERG's scenario using a response-based model and incorporating the ERG's preferred assumptions increases the ICER to £95,709 per QALY. The costs and QALYs generated for the entrectinib arm were similar to that of the ERG base-case analysis but the costs and QALYs in the established management arm are higher. This is because the response-based model estimates a higher rate of survival for the established management arm than the partitioned survival model. Incorporating the ERG's preferred assumption in the company's partitioned survival model increases the base case ICER by around £24,500 (£77,109 instead of £52,609).
Technical team preliminary judgement and rationale	If individual patient level data are available then the Bayesian Hierarchical Model framework can be used to estimate OS and PFS (see issue 9). In this case, the technical team prefer the ERG's response based model. However, if not, the technical team consider the company's approach reasonable for decision making.
Summary of comments	**Comments received from the company:** Maintain that the most appropriate model structure for decision making is the partitioned survival model. Continued recruitment of new patients and follow-up of existing patients, alongside other data collection proposals, will provide further confidence in the base-case model by reducing the uncertainties identified during the appraisal process. **ERG considerations on new evidence received during technical engagement:** Acknowledge that there are drawbacks to the response-based model approach but maintain that the company's approach does not fully account for the uncertainty in effectiveness across tumour types and is limited by the small sample size and immaturity of the available sample data. Unrepresented tumour types are also not accounted for in the company's modelling framework.
Technical team judgement after engagement	The technical team acknowledge that there are limitations and uncertainties with each of the modelling approaches. The committee will base their decision on the uncertainty of each model, see issue 10.

Issue 15 – Extrapolation of overall and progression-free survival

Questions for engagement	32. Is the exponential or Weibull distribution most appropriate for extrapolating overall and progression-free survival?
Background/description of issue	**The company** considered two parametric curves to give clinically plausible overall survival predictions, Weibull and exponential. The company used the exponential curve in its base case on the basis that it had the best statistical fit across all scenarios. Using the exponential function, the company's model estimates overall survival to be longer when in the progressed health state than in progression-free health state. For PFS, the company considered 4 parametric functions to be clinically plausible: exponential, Weibull, Gompertz and Gamma. Exponential function was used in the company base case and the company deemed this a 'conservative' choice that gives a statistically and clinically plausible estimate of PFS. **The ERG** considers the exponential, Weibull, Gompertz and Gamma functions to have good statistical fit for overall and progression-free survival. For OS they consider that they all make reasonable predictions about long term survival extrapolations. The ERG notes that the exponential function is the only function to predict that post progression survival is longer than progression-free survival. The ERG questions the clinical plausibility of this given that entrectinib is discontinued on progression and only [redacted] of people received subsequent therapy. The ERG also highlight that the positioning of entrectinib as last line of therapy means that few active treatment options would be available to regain tumour control after progression, with consequences for post-progression mortality. The ERG considers the Weibull, Gompertz and Gamma to be more reasonable extrapolations of OS with the ERG favouring the Weibull function because of its marginally better statistical fit. For PFS, the ERG notes that all 4 curves produce very similar estimates of mean PFS and that the ICER is insensitive to this model parameter. The ERG considers the Weibull function as the preferred extrapolation function as it is likely the most appropriate given its good statistical fit and for consistency with the ERG's preferences

	regarding the extrapolation OS; combining the Weibull function for OS with an exponential function for PFS implies a decreasing hazard for progression events which is clinically unlikely. The Weibull distribution to extrapolate OS and PFS is used in the ERG's base case. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** considers that the Weibull extrapolation for both overall and progression-free survival is as clinically plausible as the exponential distribution.
Why this issue is important	Choice of function for the survival extrapolation impacts on the clinical plausibility of the estimated survival. The company's choice of the exponential function gives survival estimates that are longer in the post-progression health state than the progression-free health state. This is deemed to be clinically implausible by the ERG. Using the Weibull distribution for overall survival for entrectinib increased the company's base case ICER by around £11,540 (£64,149 instead of £52,609). Using the Weibull distribution for progression-free survival for entrectinib decreased the company's base case ICER by around £150 (£52,463 instead of £52,609).
Technical team preliminary judgement and rationale	The model results are sensitive to the choice of parametric function for overall survival in the model. The technical team appreciate the difficulty in obtaining clinical validation of the survival estimates using each distribution. The Weibull function should be used to extrapolate OS and PFS as it gives a more clinically meaningful survival estimate (i.e. survival estimate for progression-free health state is longer than the post-progression health state), is a good statistical fit and is more conservative than the exponential function.
Summary of comments	**Comments received from the company:** Maintain that the exponential distribution is the most appropriate for extrapolating overall survival. Scenario analyses provided where a proxy method for applying a Weibull distribution to the comparator arm is explored in an attempt to address the potential overestimation of the comparator OS in the ERG's original analysis. Using the Weibull distribution for overall survival for entrectinib increased the company's updated base case ICER by around £13,600 (£62,962 instead of £49,358). Using the Weibull distribution for overall survival for entrectinib and the company's proxy method for applying the Weibull distribution to the comparator arm increased the company's updated base case ICER by around £5,900 (£55,244 instead of £49,358). The company consider that further data collection through a period in the CDF will help address the uncertainty around the extrapolation of entrectinib overall survival data. **ERG considerations on new evidence received during technical engagement:** Maintain their preference for the Weibull distribution as it gives estimates of pre- and post-progression survival that are more clinically plausible than those generated by the exponential distribution. No comments on the company's proxy method for applying the Weibull distribution to the comparator arm.
Technical team judgement after engagement	The technical team maintain that the Weibull distribution should be used to extrapolate OS and PFS as it gives a more clinically meaningful survival estimate, is a good statistical fit and is more clinically plausible than the exponential function.

Issue 16 – Drug wastage and source of treatment costs

Questions for engagement	33. Is it appropriate to assume no drug wastage of entrectinib? 34. Should drug acquisition costs for the comparator therapies be sourced from the British National Formulary (BNF) or the electronic market information tool (eMIT)?
Background/description of issue	The BNF provides list prices of treatments. The eMIT provides information on the average price paid by the NHS for pharmaceuticals. Drug wastage can occur when people stop taking treatment that has been distributed to them. **The company** did not include the costs of drug wastage for entrectinib in its economic model or provide any explanation about drug wastage. The company sourced the drug acquisition costs, dosing frequency and route of administration information from the BNF for the comparator therapies included in its model. **The ERG** considered that drug wastage of entrectinib should be included in the economic analysis. This is considered appropriate as once a pack of tablets has been started these would not be reused should treatment be stopped part-way through a pack. The ERG considers eMIT to be a more accurate and up-to-date indicator of treatment costs. The ERG ran scenarios where drug wastage is included in the economic model and where eMIT costs are used to source the costs of comparator therapies, independently. The ERG included drug wastage in their base case and used eMIT as the source for comparator drug costs. **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** considers eMIT to be the correct source of treatment costs as these represent the costs borne by the NHS.
Why this issue is important	Including drug wastage for entrectinib increases the company's base case ICER by around £2,750 (£55,357 instead of £52,609).

	Using eMIT as the source of the costs of comparator therapies decreases the company's base case ICER by around £530 (£52,081 instead of £52,609).
Technical team preliminary judgement and rationale	Drug wastage should be included in the base case. Drug acquisition costs for the comparator therapies should be sourced from the eMIT.
Summary of comments	**Comments received from the company:** Acknowledge that there may be a small amount of drug wastage in clinical practice. Assumed 100% dose intensity for entrectinib in original submission as a conservative assumption when the mean observed dose in the original analysis (31st May 2018) was [redacted]. Including drug wastage for entrectinib increases the company's updated base case ICER by around £2,700 (£52,103 instead of £49,358). Applying the mean observed dose intensity for entrectinib decreases the company's updated base case ICER by around £4,500 (£53,819 instead of £49,358). Using eMIT as the source of the costs of comparator therapies decreases the company's updated base case ICER by around £255 (£49,103 instead of £49,358). **ERG considerations on new evidence received during technical engagement:** No comments.
Technical team judgement after engagement	The technical team consider the results including drug wastage and eMIT costs for comparator treatments to be the most appropriate for decision making.

Issue 17 – Administration costs and resource use

Questions for engagement	35. Have administration costs been adequately captured in the company's model? 36. Has resource use been adequately captured in the company's model?
Background/description of issue	**The company** used a simplifying assumption to cost the administration costs of treatment as they categorised the different types of treatment into three administration classes: oral, simple intravenous and complex intravenous. Each category is associated with an average monthly administration costs, which is applied to each intervention. Similarly, for healthcare resource use, the company estimated health state costs based on the administration categories as a simplifying assumption. For each category, the most recent NICE technology appraisal identified within the search to identify comparator technologies was used as the source of the resource use. For the progressed disease health state, the company assumed that in the comparator arm, palliative care would be given following disease progression and so it did not include any monitoring or testing costs. **The ERG** notes that the infusion time varies significantly within the categories used by the company in their simplifying assumption. Further, the ERG consider that some costs have not been included in the company's progressed-disease health state cost. To accurately reflect clinical practice, the ERG consider the company's analysis to be missing medication costs (e.g. steroids, NSAIDS, morphine, bisphosphonates and dietary supplementation), and tests and procedure costs (e.g. full blood count, serum chemistry, CT scan, home oxygen and x-ray). **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** highlighted that the company did not include the chemotherapy tariff cost for oral treatments in the entrectinib arm of the model. The appropriate cost for inclusion is SB11Z oral chemotherapy tariff (£120 per visit).
Why is this issue important	The ERG did not have the resource to derive and include each of the individual administration costs for each of the comparator interventions to be used. It is uncertain whether the company's approach under- or over-estimates the costs in the comparator arm. The ERG state that including the additional costs in the progressed-disease health state will likely have a small impact on the cost-effectiveness result but including these costs would increase the ICER. Entrectinib has a substantial mean treatment duration so the additional cost of the chemotherapy tariff cost for oral treatments will impact on the cost of treatment mainly in the entrectinib arm. Including this cost would increase the ICER.
Technical team preliminary judgement and rationale	The technical team would like to see a scenario with the individual administration costs for each of the comparator treatments included in the analysis as well as the additional costs included in the progressed-disease health state, as suggested by the ERG.
Summary of comments	**Comments received from the company:** Did not conduct the analysis with the individual administration costs for each of the comparator treatments included in the analysis. Included the oral chemotherapy tariff cost to both entrectinib and the comparator oral chemotherapies and ran a scenario that included monitoring costs in the progressed disease state resource use. Including the oral chemotherapy tariff in both arms increases the company's updated base case ICER by around £2,120 (£51,476 instead of £49,358).

	Including the monitoring costs in the progressed disease state resource use increases the company's updated base case ICER by around £290 (£49,647 instead of £49,358). **Comments received from NCRI-ACP-RCP-RCR:** Included administration costs and resource use seem reasonable. **ERG considerations on new evidence received during technical engagement:** Welcomes the inclusion of the oral chemotherapy tariff and the scenario analysis including the revised progressed-disease health state costs. It is a limitation that the administration costs for each of the comparators are not included in the company's model and this is an outstanding uncertainty.
Technical team judgement after engagement	The technical team consider the inclusion of the oral chemotherapy tariff and monitoring costs in the resource use for the progressed disease state to be appropriate. The technical team agrees with the ERG that there is outstanding uncertainty around the administration costs included for the comparators.

Issue 18 – Implementation and training costs

Questions for engagement	37. What additional infrastructure and training requirements could be considered for this appraisal?
Background/description of issue	Site-agnostic oncology treatments are a new concept in clinical practice in England. Oncologists will likely require training about what tumour types may be eligible for treatment with entrectinib and at what point in the treatment pathway entrectinib can be used. Further, training will likely be required around the collection and handling of the tissue biopsies for testing for NTRK gene fusions.
Technical team preliminary judgement and rationale	It is important to capture the impact of using entrectinib in clinical practice. Any likely constraints on the resources required to support the implementation of the appraised technology should be highlighted, and comment should be made on the impact this may have on the implementation timescale.
Summary of comments	**Comments received from the company:** Agree that the education of oncologists and pathologists in the new concepts associated with tumour-agnostic therapies is the primary additional training requirement. The company intends to collaborate closely with the NHS to support training in these areas. Additional training may be required around the appropriate material handling where RNA extraction is required for NGS-based testing should a DNA/RNA hybrid panel approach be implemented. This training is in line with the existing training requirements associated with wide scale implementation of NGS-based testing. **Comments received from NCRI-ACP-RCP-RCR:** No additional infrastructure or training requirements for the delivery of entrectinib. Depending on the strategy for identifying NTRK gene fusions, there will be infrastructure and training considerations. **ERG considerations on new evidence received during technical engagement:** Highlights that the company have not addressed the issue of additional infrastructure that would be required if NTRK fusion testing across cancers were to be implemented. If the hierarchical approach is used then considerable investment in histopathological services would be required for the IHC testing.
Technical team judgement after engagement	This issue will be discussed by the committee with input from NHS England and the Genomic Medicine Service. Any likely impact on the implementation timescale will be noted in the appraisal documents.

Issue 19 – Utility values

Questions for engagement	38. Is the utility value estimate for the entrectinib progression-free health state collected in the STARTRK-2 trial appropriate ([redacted])? 39. Is the utility value for the established management progression-free health state appropriate (0.73)? 40. Should the utility value for the progression-free state be equal between the entrectinib and established management arms? If yes, what should the utility value be? 41. Is it reasonable to assume that the utility value for the progressed disease health state be equal between the entrectinib and established management arms? If yes, is the value of 0.59 appropriate?
Background/description of issue	EQ-5D-3L data was collected in the STARTRK-2 entrectinib clinical trial. **The company** used a linear mixed model with a nested random effect by patient within tumour type to estimate utility values for the progression-free health state for people treated with entrectinib ([redacted]). The company considered its estimate for the progressed disease health state to be implausible as it is higher than the pre-progressed value. The company derived the utility values for the comparator arm by extracting utilities from a single NICE technology appraisal for each tumour type included in the entrectinib clinical trial. A single utility value was calculated for each health state by calculating the average utility value, weighted by the distribution of the tumour types included in the entrectinib trial (see issue 6) [0.73 for progression-free and 0.59 for progressed disease].

	The company assumed that the utility estimates for the progressed disease health state were equal between the two treatment arms (0.59 for both entrectinib and established management). The company uses different utility values for the pre-progressed health state for the two treatment arms ([redacted] for entrectinib and 0.73 for established management). They give the justification that entrectinib is an oral treatment with more convenient administration and relatively tolerable safety profile when compared with traditional cytotoxic chemotherapies which form the majority of the comparator products. **The ERG** considers the assumed quality of life benefit in the progression-free state to be plausible and reasonable given the safety profile of entrectinib but note the lack of evidence to justify the assumption on differential quality of life progression-free is a concern. The ERG considers there to be considerable uncertainty regarding the magnitude of any difference. The ERG is concerned that the company's choice of NICE technology appraisal to extract the utility value for the comparator arm may be biasing the cost-effectiveness analysis. The selected utilities reflect a specific line of therapy. This may not be the line of therapy where entrectinib is used in clinical practice or equivalent to the line of therapy where entrectinib was given in the entrectinib trial (see issue 2). **The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund** considers it appropriate to use the same utility value in the progression-free health state for the entrectinib and comparator arm.
Why this issue is important	Inaccurate utility values could bias estimates. Differential utility values for the pre-progressed health state between the two treatment arms biases the cost-effectiveness results in favour of entrectinib in the company base case. Given that people in earlier lines of therapy may have better HRQoL, sourcing the utility value for the entrectinib arm from the clinical trial where entrectinib may be given at earlier line of therapy than in clinical practice, and comparator utility values from a single NICE technology appraisal for each tumour type selected at a line of therapy that may not be equivalent to the entrectinib line of therapy, may lead to the difference in HRQoL between the two arms to be overestimated. This would bias the cost-effectiveness results in favour of entrectinib. Using the comparator utility value for the progression-free health state (0.73) for both treatment arms increases the company's base case ICER by around £6,300 (£58,946 instead of £52,609).
Technical team preliminary judgement and rationale	The technical team prefer the same value being used for both the entrectinib and comparator arms (s*entence updated following technical engagement to correct a factual inaccuracy).* This is because of the uncertainty associated with the data from the clinical trial resulting in a clinically implausible value for the post-progression health state. Further, if entrectinib was given at an earlier point in the treatment pathway in the clinical trial than it would be in clinical practice (see issue 2) then the values derived from the clinical trial are likely overestimated. Although, the technical team would consider more evidence to support the different utility values across the treatment arms and the magnitude of this difference.
Summary of comments	**Comments received from the company:** Maintain that the utility estimates used in the original submission are appropriate. Reiterate that the differential utility value between the progression-free health states for entrectinib and the comparator arm is appropriate because of entrectinib's oral administration and its relatively tolerable safety profile compared with traditional cytotoxic chemotherapies. Acknowledge that there is some uncertainty in the entrectinib progressed state utility value. **ERG considerations on new evidence received during technical engagement:** No comments.
Technical team judgement after engagement	The technical team maintains that there is uncertainty associated with the utility data. The company's values are potentially plausible for the pre-progressed heath state based on the position in the treatment pathway that entrectinib was given in the entrectinib trials but the magnitude of the difference in utility between arms is uncertain given the lack of evidence presented. The impact on the ICER of including the same utility value for both entrectinib and the comparator for the pre-progressed health state is acknowledged.

Issue 20 – Uncertainty around the cost-effectiveness results

Questions for engagement	42. Is the uncertainty around the cost-effectiveness results captured appropriately in the company's model?
Background/description of issue	**The company** varied the extrapolated mean around a normal distribution for published survival estimates for the comparator arm. This company noted the lack of covariance matrices and correlations reported and the fact that they used the exponential model for the extrapolation as justification for its approach. **The ERG** has concerns about the uncertainty of the probabilistic ICER included in the company submission. The ERG notes the narrow distributions of the comparator costs, total life years gained and total QALYs appear to be unrealistic. The ERG is also concerned about the standard errors around the survival estimates used to construct

	the established management comparator. The standard errors are assumed and were not extracted from the original sources by the company for comparator effectiveness, utilities or costs.
Why this issue is important	The ERG considers the uncertainty around the cost-effectiveness results of the comparator to be underestimated.
Technical team preliminary judgement and rationale	The technical team would like to see some further analysis that explores the uncertainty around the cost-effectiveness results. This could involve incorporating the standard errors from the original sources for comparator effectiveness, utilities and costs.
Summary of comments	**Comments received from the company:** Updated probabilistic sensitivity analysis with broader comparator standard errors (0.25 instead of 0.15) and wider (value not defined) confidence intervals for comparator life years gained, QALYs gained and costs, shown in table 9. **Table 9: Company's updated probabilistic sensitivity analysis with broader standard error and confidence intervals**

Technology	Total costs (£)	Total life years gained	Total QALYs	ICER (£/QALY)
Established management	[redacted]	[redacted]	[redacted]	£49,090
Entrectinib	[redacted]	[redacted]	[redacted]	

	ERG considerations on new evidence received during technical engagement: The principal concern regarding the company's approach to modelling uncertainty relates to the presentation of a single ICER. A single ICER could misrepresent the uncertainties of the appraisal such as the between tumour type heterogeneity in costs and efficacy. Important to reflect this uncertainty in the economic analysis to allow the committee to consider the risk associated with a histology independent recommendation and to allow the committee to consider an optimised recommendation. This uncertainty is also potentially important to consider in the context of future data collection, as it will determine the value of collecting further data.
Technical team judgement after engagement	The technical team welcomes the company's revised probabilistic sensitivity analysis and understands the ERG's concerns about the presentation of a single ICER. If entrectinib is to be recommended in the CDF this issue may be better understood and considered further after a period of data collection. There will be some unresolvable uncertainty and committee will be aware of this.

Issue 21 – End of life

Questions for engagement	43. What is the life expectancy of the patient group receiving established management? 44. What is the extension to life of the patient group receiving entrectinib?
Background/description of issue	**The company** suggest that entrectinib meets the end-of-life criteria (specified in NICE's guide to the methods of technology appraisal) compared to the current established management across all patients potentially eligible for entrectinib. The company's supporting evidence comes from its integrated efficacy analysis. The company's base case model (submitted at clarification) predicts a median overall survival of 14.48 months for people receiving established management (mean: 19.34 months). The company's model predicts a mean extension to life of 16.43 months for people treated with entrectinib. The company highlight in response to clarification that the prognostic implications of NTRK gene fusions are also a consideration for the end-of-life criteria. They note the limited data available that suggests that people with NTRK gene fusions perform less well on current standard of care than people without NTRK gene fusions. **The ERG** highlight that it is challenging to apply the conventional application of end-of-life criteria to a highly heterogeneous population with no established comparators. They give two reasons: 1) The end-of-life criteria may apply across some tumour types and not others. 2) There is uncertainty around the estimates of both life expectancy and extension to life which may vary widely by tumour type. This is exacerbated by the uncertainty around the positioning of entrectinib in the treatment pathway for each tumour type (see issue 2). The ERG do not consider the first of the end-of-life criterion to be met, despite the ERG's base case mean survival estimate being below the two years stipulated in Criterion 1, given in Table 9 below. **Table 9: Criterion 1 of NICE's end-of-life criteria and the ERG's estimates of overall survival**

Criterion	ERG's base case	ERG's response-based model for non-responders to entrectinib (Weibull function)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Mean = 20.9 months Median = 15.7 months	Mean = 24.7 months Median = 19.9 months

The reasons for the ERG not considering criterion 1 to be met are that:

1) The comparability of the comparator population with patients eligible for entrectinib in clinical practice is highly uncertain. Life expectancy data from NICE technology appraisals covering multiple lines of therapy within the same indication has been used by the company. However, the testing strategy will dictate when entrectinib is made available and this will always be at the same point in the pathway and the comparator life expectancy estimate does not reflect that.
2) The life expectancy of people who have NTRK fusion positive tumours at sites not included in the entrectinib trials are not included in the estimate. The life expectancy of these populations is unknown and may be significantly different to those included in the company submission.
3) The ERG does not consider existing literature to support the concept of NTRK as being consistently prognostic of a shorter life expectancy.
4) The average overall survival estimate does not reflect the heterogeneity of life expectancy across tumour types when they reach eligibility for treatment. Inspection of the mean overall survival estimates by tumour type used in the ERG base case analysis show that people with thyroid and neuroendocrine tumours would not meet the first end of life criteria, mean overall survival of 44.65 and 57.14 months, respectively. They represent approximately 31% of the incident NTRK fusion population.

The ERG considers it more likely that the second of the end-of-life criterion is met as extension to life following treatment with entrectinib is likely to be greater than 3 months, as shown in table 10.

Table 10: Criterion 2 of NICE's end-of-life criteria and the ERG's estimates of overall survival following treatment with entrectinib

Criterion	ERG's base case average mean overall survival	ERG's response-based model for responders to entrectinib (Weibull function) – Median overall survival
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	31.1 months (extension of 10.2 months)	25.2 months (extension of 5.3 months)

However, the ERG highlight that the benefit in unrepresented tumour types is unknown and cannot be assumed to be equal to those included in the entrectinib clinical trials. The ERG used the predictive distribution in its response based model (described in issue 13) to estimate the extension to life in unobserved tumour types. The extension to life could potentially range from [redacted] months to [redacted] months. This indicates that some tumour types may not meet the criterion 2 of end-of-life.

The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund considers entrectinib to meet the end-of-life criteria if using a weighted average of survival for the comparator is a reasonable approach. However, they consider that survival in the comparator arm may be overestimated as they consider the company to have assumed the use of entrectinib at earlier points in the treatment pathway in some diseases (see issue 11).

Why this issue is important	The appraisal committee's judgements about the acceptability of the technology as an effective use of NHS resources will take into account whether the technology meets the criteria for special consideration as a 'life-extending treatment at the end of life'. A technology which meets NICE's end of life criteria has an increased cost-effectiveness threshold.
Technical team preliminary judgement and rationale	Based on the ERG's base case, the mean overall survival is much greater than 24 months for 2 tumour types (neuroendocrine and thyroid tumours) and is around 24 months for cholangiocarcinoma and infantile fibrosarcoma. Neuroendocrine and thyroid tumours represent approximately 31% of the incident NTRK fusion population. Therefore, a proportion of the tumour types included in the analysis do not meet the end-of-life criteria and committee will consider this in their deliberations when making a judgement on whether entrectinib meets end-of-life. The positioning of entrectinib in the treatment pathway is important in these considerations (see issue 2), as well as the

	prevalence and distribution of tumour types (see issue 1 and 6) and unrepresented tumour types (see issue 7) as this impacts on the proportion of people that the end-of-life criteria apply. There is also uncertainty about the natural history of some of the tumours harbouring the NTRK gene fusion (see issue 12).
Summary of comments	**Comments received from the company:** In the company's revised base-case, the mean discounted OS for entrectinib and the weighted comparator are 35.8 and 19.3 months, respectively. The current standard of care results in mean OS of less than two years, and the mean OS gain conferred by entrectinib is 16.5 months. Using the NHSE and NHSI's preferred comparators results in a discounted mean OS of 18.9 months for the current standard of care, resulting in a mean OS gain of 16.9 months for entrectinib. Dispute the ERG's concerns about why criterion 1 is not met with reasons including: • New scenario analysis using NHSE and NHSI's preferred comparators shows that the mean OS for current standard of care is less than 24 months. • Uncertainty around the life expectancy of unrepresented tumour types and the prognostic impact of NTRK gene fusions may be addressed during a potential CDF data collection period. • The ERG quote high mean overall survival figures for thyroid and neuroendocrine tumours and the company acknowledges that these tumours may have relatively good prognosis but note that these figures are heavily confounded by cross-over to active therapy in the case of best supportive care. The company consider that further data collection if entrectinib is recommended in the CDF will help to resolve some of the uncertainty around the efficacy of entrectinib. **Comments received from NCRI-ACP-RCP-RCR:** Life expectancy of the patient group receiving established management will vary considerably depending on the tumour type and the line of therapy. Extension to life for people receiving entrectinib cannot be said with certainty and the extension gained will be significantly impacted on by the underlying biology/rate of progression and the availability of subsequent therapies. **ERG considerations on new evidence received during technical engagement:** Maintain that the primary concern is the effect of heterogeneity between tumour sites on decision uncertainty and how to interpret the end-of-life criteria given this heterogeneity.
Technical team judgement after engagement	The technical team recognise the uncertainties around applying the end-of-life criteria to histology-independent treatments such as entrectinib. A proportion of the tumour types included in the analysis do not meet the end-of-life criteria. Committee will consider this in its judgement. This is compounded by uncertainty in the positioning of entrectinib and judgements around tumours types that are unrepresented in the evidence base

Issue 22 – Innovation

Questions for engagement	45. Is entrectinib an innovative treatment?
Background/description of issue	**The company** claim that entrectinib is an innovative treatment. They describe entrectinib as a step-change in the treatment of cancer, as the focus is shifted from the origin of the primary cancer to the underlying oncogenic driver, regardless of histology. Further, they state that entrectinib is a CNS-active NTRK inhibitor. Utilising novel genomic technologies such as NGS to identify NTRK fusion positive solid tumours may also provide benefits to patient health and cost efficiencies for health care systems as multiple different actionable targets may be identified, even where NTRK-fusion negative and this could lead to clinical trial availability or treatment with other targeted therapies. Regulatory bodies have formally recognised the innovative nature of entrectinib.
Why this issue is important	Committee can take into account the potential innovative nature of the technology, in particular its potential to make a significant and substantial impact on benefits that are unlikely to be included in the QALY calculation during the appraisal.
Technical team preliminary judgement and rationale	Entrectinib is potentially innovative in that it is a treatment for a newly identified rare gene fusion that occurs in a wide range of tumour types. The technical team recognise that it is one of the first site-agnostic treatments to be appraised by NICE. However, they are aware of other targeted inhibitors that can be used to treat a range of different tumour types and in those cases (for example treatments for the BRAF V600E mutation), larger studies have been done. Entrectinib could represent a step-change in the treatment of cancer however, there is a lack of evidence of demonstrable and distinctive benefits of a substantial nature which may not have been adequately captured in the reference case QALY measure. The technical team consider that a major innovation is already being led by the NHS in developing more sophisticated strategies to improve genomic testing in clinical practice. These advances may facilitate uptake of treatments such as entrectinib if it is to be recommended. However, entrectinib is one of the first site-agnostic treatments to be appraised by NICE and represents potential for a future service redesign based on biological marker rather than histology. The committee will consider the innovative nature of entrectinib when making its recommendations.
Summary of comments	**Comments received from the company:** Maintain that entrectinib is an innovative treatment for reasons outlined in the original submission. **Comments received from NCRI-ACP-RCP-RCR:**

	Entrectinib is innovative. **Comments received from GIST Support UK:** Entrectinib is innovative. **ERG considerations on new evidence received during technical engagement:** No comments.
Technical team judgement after engagement	Committee will take into account the potential innovative nature of entrectinib as part of its decision making.

Issue 23 – Cancer Drugs Fund

Questions for engagement	46. Does entrectinib meet the criteria for inclusion in the Cancer Drugs Fund? 47. What data would be most useful to collect to address the outstanding uncertainties? For example, unrepresented tumour types.
Background/description of issue	**The company** have proactively positioned entrectinib for funding via the Cancer Drugs Fund (CDF) as opposed to by routine commissioning in the NHS. The **technical team is** aware of the arrangements for the Cancer Drugs Fund agreed by NICE and NHS England in 2016, noting NICE's Cancer Drugs Fund methods guide (addendum). The technical team consider that there is clinical uncertainty that could be reduced through data collection via ongoing studies.
Why this issue is important	The CDF is a potential option if there is plausible potential for the drug to satisfy the criteria for routine commissioning, but there is significant remaining clinical uncertainty which needs more investigation, through data collection in the NHS or clinical studies. This means the CDF will fund the drug, to avoid long delays, but would require information on its effectiveness before it can be considered for routine commissioning (when the guidance is reviewed). The company have not provided evidence to demonstrate that entrectinib has plausible potential for cost-effectiveness. The uncertainty around entrectinib's plausible potential to be cost-effective largely rests on the testing costs that will be incurred if entrectinib is recommended in the Cancer Drugs Fund in a period of managed access.
Technical team preliminary judgement and rationale	The technical team considers that entrectinib meets the criteria for inclusion in the Cancer Drugs Fund. However, taking into account its considerations about the end of life criteria, the technical team is currently uncertain about entrectinib's plausible potential to be cost-effective at the offered price given the current uncertainty around testing costs. The company's base case ICER is above the range that NICE would normally consider cost effective when a treatment meets the end-of-life criteria. This ICER does not take into account the ERG and technical team's preferred assumptions or the confidential discounts for other treatments included in the comparator arm. The technical team would like to see scenario analyses around the testing costs included in the analysis as discussed in issue 4. Committee will be interested in the practicalities of data collection within the Cancer Drugs Fund during the course of the appraisal. In particular, they will be interested in NHS England's intentions around data collection being prioritised based on unmet need, tumour types where no data has been collected previously or tumour types with high prevalence of NTRK gene fusions.
Summary of comments	**Comments received from the company:** Entrectinib meets the criteria for inclusion in the CDF. The company's revised base-case ICER (£49,358) that includes 100% of testing costs falls under the maximum end-of-life threshold. The ICER with testing costs excluded, which the company believe is more appropriate, is considerably lower (£35,770) and closely mirrors the ERG's base-case ICER excluding testing (£38,030, pre-technical engagement). Acknowledging the number of clinical and cost-effectiveness uncertainties, the company has proactively proposed entrectinib for a period of data collection in the CDF. The company provide details of their CDF data collection proposals in their response to technical engagement. **Comments received from NCRI-ACP-RCP-RCR:** Entrectinib meets the criteria for inclusion in the CDF. Imperative to collect phase 4 data prospectively on both unrepresented tumour types and those that are represented but with only small numbers. Most useful data to collect to address the outstanding uncertainties include: • Number of patients screened to identify one TRK fusion in each disease type • Prevalence of NTRK fusion across *all* disease types in the UK population • Number of cycles of entrectinib administered • Outcome of patients – RR, real world PFS and OS • Registry for all patients with NTRK fusion (regardless of whether patients receive entrectinib) and analysis of outcomes on each treatment received. **ERG considerations on new evidence received during technical engagement:** The company's revised base case ICER of £49,358 per QALY suggests there is plausible potential for cost-effectiveness if the end-of-life criteria applies.

	The ERG's revised base case ICER including testing costs is £76,322 which is substantially above the £50,000 threshold considered when a treatment meets end-of-life. The ERG maintains that testing costs should be included in the analysis, see issue 4. Where decisions are being made on a single, composite ICER, there is considerable uncertainty because the single ICER does not represent all tumour types known to have NTRK fusions. The plausible potential for cost-effectiveness for these tumour types is unknown. The single ICER also includes some tumour types in which there is evidence that entrectinib is not cost-effective and could not be cost-effective given the cost of identifying patients. By using a single, composite ICER, some tumour types will be included in the CDF despite no evidence indicating the use of entrectinib could plausibly be cost-effective in these tumour types. It may not be appropriate to apply the end-of-life threshold for tumour types in which it would not be met in a single-indication appraisal, and therefore it may not be appropriate to include these tumour types in the CDF. The ERG accepts that data collection in the CDF may help to address some uncertainty but notes that data collected from continued trial follow-up will not necessarily capture survival data in the unrepresented tumour types, which is a notable uncertainty, see issue 7.
Technical team judgement after engagement	The technical team maintains that entrectinib meets the criteria for inclusion in the Cancer Drugs Fund in terms of clinical uncertainty that could be addressed through further data collection. However, entrectinib's plausible potential to be cost-effective remains uncertain until more information is provided by NHS England and the Genomic Medicine Service around testing costs and also on committee's decision about whether the end-of-life criteria apply.

4. Issues for information

Tables 11 to 13 are provided to stakeholders for information only and not included in the technical report comments table provided.

Table 11: Technical team preferred assumptions and impact on the cost-effectiveness estimate

Alteration	Technical team rationale	ICER with testing costs included	Change from base case with testing costs included	ICER with testing costs excluded (from economic model)	Change from base case with testing costs excluded
Company base case	**–**	**£52,609**	**-**	**£36,914**	**-**
Company updated base case (including children and primary CNS tumours)	Issue 8, resolved at technical engagement	**£49,358**	**-**	**£35,770**	**-**
2. Weibull distribution for entrectinib OS and PFS and using alternative distribution of tumour sites	Issue 6 and 15	£62,750	+£13,392	£42,638	+ £6,868
3. Inclusion of incremental testing costs associated with identifying NTRK gene fusions only in comparator arm	Issue 4	£60,234	+£10,876*	-	-
4. Removal of testing costs of NGS for lung cancer and the inclusion of incremental testing costs associated with identifying NTRK gene fusions only	Issue 4	NA – cannot be applied to company base case*	NA – cannot be applied to company base case*	-	-
5. Confirmatory RNA-based NGS test after WGS test and the inclusion of incremental testing costs associated with identifying NTRK gene fusions only	Issue 4	£50,593*	+£1,235*	-	-
6. Testing costs estimated using the number needed to screen based on the whole NTRK population and the inclusion of incremental testing costs associated with identifying NTRK gene fusions only	Issue 4	NA – cannot be applied to company base case*	NA – cannot be applied to company base case*	-	-
7. Duration of treatment with subsequent therapies limited to 6 month duration	Issue 13	£39,890	-£9,469	£26,301	-£9,469
8. eMIT costs as the source for comparator costs	Issue 16	£49,103	-£255	£35,515	-£255
9. Inclusion of drug wastage for entrectinib	Issue 16	£52,103	+£2,745	£38,515	+£2,745

Technical report template 2 – AFTER technical engagement

Alteration	Technical team rationale	ICER with testing costs included	Change from base case with testing costs included	ICER with testing costs excluded (from economic model)	Change from base case with testing costs excluded
10. ERG's revised estimation of number requiring confirmatory testing and updated incidence of thyroid tumour	-	£49,539*	+£181*	-	-
11. Removal of inappropriate comparators as suggested by NHSE and NHSI	Issue 11	£49,294*	-£65*	£35,933	+£163
12. Inclusion of oral chemotherapy tariff cost	Issue 17	£51,491*	+£2,133*	£37,903	+£2,133*
13. Revised cost of progressed disease health state	Issue 17	£49,647*	+£288*	£36,058	+£288*
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate	**–**	**£79,330**	**+£29,972**	**£40,717**	**+£4,947**

* corrected for a minor error in the ERG model. Committee presented with these correct values in the committee presentation.

Table 12: Outstanding uncertainties in the evidence base

Area of uncertainty	Why this issue is important	Likely impact on the cost-effectiveness estimate
Clinical evidence for entrectinib	The clinical evidence base for entrectinib is small, n=66 including people with primary NCS tumours and paediatric patients. The largest group of patients is 13 for a single tumour type (sarcoma) and only 1 patient for some tumour sites (breast, pancreas, appendix, congenital mesoblastic nephroma) from 3 separate trials. There is also considerable uncertainty regarding the extent to which the high response rates seen in the results from the larotrectinib clinical trials results translate into clinically meaningful survival benefits. It is unclear if However, they note that the overall survival data are immature, see below.	The effect of the limitations of the evidence base is unknown. However, the limitations increase parameter uncertainty in the economic model, and increase uncertainty in the cost-effectiveness estimates.
Immature evidence base	The analysis from the entrectinib trials is of short duration. Median overall survival in the trial has not yet been reached ([redacted] had an event). Progression-free survival data are also immature ([redacted] had an event).	Unknown.

Table 13: Other issues for information

Issue	Comments
Primary CNS tumours and paediatric tumours (Issue 8)	In its original submission, the company did not include the 5 adults with primary CNS tumours or the 7 paediatric patients in its base case population despite their eligibility for inclusion based on the anticipated marketing authorisation. At clarification the company provided an updated model including a scenario analysis that included these additional tumour types. The company highlighted concerns that inclusion of the paediatric patients in the economic model is challenging because of the absence of a counterfactual or any robust comparator data and that CNS primary tumours were scored using a different response measure to solid tumours in the entrectinib clinical trials. The technical team consider it appropriate to include primary CNS metastases and paediatric tumours in the base case as it increases the generalisability of the evidence base to the population likely to be seen in clinical practice. The absence of robust comparator data does not add as much uncertainty to the analysis as the structural issues of a pooled analysis. **Following technical engagement, the company updated its base-case to include people with primary CNS tumours and paediatric patients.**
Company's design of its basket trial is not as per the conventional design	The ERG highlight that the company's basket trial was not designed as a basket trial in the sense intended by the FDA, EMA and EUnetHTA. The baskets in the company's trial were based on molecular targets (ALK, ROS1 and NTRK) rather than tumour type for each molecular target. The ERG state that assumptions underpinning the analysis of a basket trial may not hold for the analysis of post hoc subgroups within the NTRK fusion positive basket.
Oncology genomic testing is not yet fully operational in clinical practice	The NHS England and NHS Improvement National Clinical Lead for the Cancer Drugs Fund states that WGS will be fully operational by quarter 2 in the year of 2020/21 and NGS panel testing will be available by quarter 1 in 2020/21. Uptake of molecular testing across the 7 genomic hubs will increase during 2020/21 as genomic pathways are embedded and links are made with the clinical teams. Given the complexity of implementation, it may take a further 12 months for molecular testing to become fully embedded in practice.
Using a single ICER to represent the cost-effectiveness of entrectinib conceals the potential for significant variation in tumour specific ICERs	Entrectinib may not be cost-effective across all tumour types that are NTRK fusion positive. The ERG utilised its response based model to integrate the results of the Bayesian hierarchical analysis and generated tumour type specific ICERs. This exploratory analysis showed that the tumour type specific ICER's varied significantly from £57,467 per QALY in sarcoma to £128,679 per QALY in thyroid cancer. The ICER for all tumour types (including the ERG's preferred assumptions) is £95,723 per QALY.
Company's estimate of the eligible population may be overestimated	Company's estimate of population size includes people with any type of cancer, rather than just those with locally advanced or metastatic solid tumours harbouring an NTRK gene fusion. The ERG's have re-estimated the number of people eligible to receive entrectinib each year in England as 196. The ERG used tumour specific rates of NTRK gene fusions and disease incidence and limited the population to people at the relevant stage of the treatment pathway for each tumour type. Thirty different tumour types were included in the ERG's calculation. The ERG's estimate is conservative as it does not account for cancers in which an NTRK fusion has not yet been identified.
NTRK2 gene fusion positive tumours are under-represented in the entrectinib clinical trials	Only one person (2%) with an NTRK2 gene fusion positive solid tumour was included in the entrectinib clinical trials. It is uncertain whether people with NTRK2 gene fusion solid tumours have different prognoses to people with NTRK1- and NTRK3- gene fusions or whether there is the potential for different responses to entrectinib based on NTRK fusion type. At clarification, **the company** suggested that the low representation was because of a lower absolute prevalence of NTRK2 gene fusions, which comprise only [redacted] of *NTRK* fusions. The company stated that response to entrectinib were independent of the NTRK fusion gene. **The ERG** received clinical advice that suggested that it is plausible that the different gene fusions have different prognoses and different responses to entrectinib. The ERG highlight that data presented to the US Food and Drug Administration (FDA) as part of an new drug application (NDA) Multidisciplinary Review and Evaluation of larotrectinib in patients with NTRK-gene fusions, suggests that patients with NTRK2 gene fusions had a lower overall response rate than those with NTRK1 and -3 gene fusions, which may suggest differential response to TRK inhibitors in this population.
Company applied the discount rate twice to the costs of post-progression second line treatment costs and omitted the drug administration costs	Company provided a corrected model at the clarification stage.
Company's model in original submission did not include data from the most recent data cut from the entrectinib clinical trials	Company provided a corrected model at the clarification stage.

Equality considerations	No equality issues were identified by the company. The technical team are concerned that entrectinib is administered orally and therefore only people who have the ability to swallow will be able to use entrectinib. At clarification the company noted that paediatric patients in the STARTRK-NG study who were unable to swallow capsules were administered with an experimental formulation which could be sprinkled over food. The company stated that they are currently testing GI tube administration of the commercial formulation and is developing a new age appropriate formulation. This product does not yet have a UK market authorisation. The ERG highlight that the application of the higher willingness to pay threshold for a technology that meets the end-of-life criteria may potentially raise issue about the equity of access to treatment. In a case where the end-of-life criteria is considered to be met for the whole population, despite some of the sub-populations not meeting the criteria for end-of-life, this implies that people are able to access therapy that would otherwise be considered cost-ineffective based on conventional thresholds. In this situation, a QALY generated in NTRK positive and NTRK negative patients are being valued differently.

Authors

Gary McVeigh

Appraisal committee chair

Emily Eaton Turner

Technical lead

Christian Griffiths

Technical adviser

Linda Landells

Associate director

With input from the lead team:

Guy Makin

Lead team member

Peter Hall

Lead team member

Malcolm Oswald

Lead team member

NHS England submission for entrectinib (ID1512) and larotrectinib (ID1299) appraisals for patients with locally advanced/metastatic solid cancers which bear the NTRK gene fusion

1. NHS England regards that there is a unique set of exceptional circumstances as regards genomic testing which all meet at the current time and at the same time as NHS England has to inform NICE as to the incremental diagnostic costs for the detection of patients with solid tumours bearing NTRK gene fusions.

2. NHS England is already committed via its Long Term Plan to roll out next generation sequencing (NGS) gene panel testing for patients with cancer.

3. NHS England's Genomic Medicine Service is at a crucial time in its creation and implementation of the first ever national service for genomic testing for both cancer and non-cancer conditions.

4. NHS England recognises the very substantial pipeline in the development of cancer drugs which require future genomic testing whether these drugs have marketing authorisations which are histology independent or targeted to specific tumours. The costs of instituting gene panel testing will therefore detect many current and future genomic alterations which are potentially actionable in terms of systemic therapy.

5. The Accelerated Access Collaborative (AAC) has prioritised histology independent (tumour agnostic) products for support because of the opportunity for the NHS to pioneer the introduction of a new class of cancer therapy, both in terms of diagnostic provision and the robust translation of genomic provision into clinical practice.

6. NHS England recognises that the first 2 products in the histology independent class are/likely to be licensed at a critical time in the implementation of the Genomics Medicine Service and therefore wishes to (at least initially) bear a large proportion of the testing costs, thereby supporting the AAC's position.

7. NHS England reserves the right to reconsider how it accounts for genomic diagnostic costs when NICE re-appraises entrectinib and larotrectinib at CDF exit (at the end of any managed access period) in the context of a routine commissioning recommendation.

8. NHS England will use the principle of an individual company's contribution to the NICE economic modelling of cost effectiveness being based on the scale of the testing requirements, the incidence of the genomic alteration and any relevant considerations as outlined in the above paragraphs.

Specific NHS England submission for the incremental diagnostic costs in the entrectinib and larotrectinib appraisals

NHS England recommends that NICE includes a cost of £6,800 per NTRK gene fusion positive patient for the purposes of NICE's value assessment in the NICE technology appraisals of entrectinib and larotrectinib, this recommendation assuming that gene panel testing is in place for all locally advanced/metastatic cancers.

NHS England recognises the challenge of instituting gene panel testing for such a large population of cancer patients at such a crucial time in the development of the Genomic Medicine Service and thus NTRK gene fusion testing will be phased into practice over the next 2 years.

Prof Peter Clark NHS England CDF National Clinical Lead November 2019

Options for phasing the introduction of the detection of NTRK gene fusions for solid tumours

Immunohistochemistry (IHC) screening and then Next Generation Sequencing (NGS) on 10% of positive IHC cases (Roche base case)

NHS E does not consider that it is practical to do IHC screening for a short time as an interim measure before NGS panel testing is fully implemented for patients with advanced/metastatic solid tumours. This is because the histopathology service is already under significant pressure consequent to current workforce issues as well as escalating workload. There would be a need for 100,000 additional IHC screens per year and this would add to the workload of pathologists. A further issue is that the IHC test is known to have a significant false negative rate. Even if IHC screening were to be introduced in practice, the facility of panel testing with NGS would still be necessary for 10% of patients. By the time any IHC training had been completed and IHC screening was fully functional, it would be time to dismantle the IHC screening service as NGS capacity would then be sufficient to test all 100,000 samples.

NHS England also does not consider it to be desirable to rely on IHC screening (a 20th century technology) when the future 21st century technologies of NGS and whole genome sequencing (WGS) are beckoning. Both NGS and WGS can readily deliver the many new genomic test results necessary for the future drug pipeline. NHS England, the Genomic Laboratory Hubs (GLHs) and pathology services therefore need to spend their energies driving forward the genomic testing pathways necessary for the realisation of a genomic medicine service equipped for the future.

Options for the introduction of NGS detection of NTRK gene fusions

The optimal way of testing solid tumour patients with locally advanced/metastatic solid tumours is to do this at the time that patients embark on their locally advanced/metastatic disease treatment pathways. This would require 100,000 tests in England/year and would allow standard therapies to proceed but knowledge as to whether TRK inhibitors would be indicated or not would be known early in the systemic therapy care pathway. Standard systemic therapies are defined as ones commissioned by NHS England either as NICE-recommended treatments or ones common to UK/European treatment guidelines. Most cancers have 1-3 standard therapies in their care pathways and these would be used before TRK inhibitors although there are some (uncommon) cancers which do not have any effective standard systemic therapies. In these cancers, TRK inhibitors would be used as upfront therapy.

The seven GLHs currently vary in their ability and capacity to institute NGS panel testing in early 2020. It will take time for the tissue collection processes to be fully operational in view of the new genomic diagnostic pathways which need to develop for such widespread testing and in such numbers. If there was no planned phasing of testing for the TRK inhibitors in the next 1-2 years, there would be significant geographical variability across England in access to NGS panel testing.

It is thus inevitable given the revolution currently taking place within the Genomic Medicine Service and the establishment of the GLHs currently taking place that there needs to be a phased introduction over the next 1-2 years of NGS panel testing for patients with advanced solid tumours. The options of phasing are these:

1. **Test patients after failure of standard therapies** This option would reduce the number tested as there is always a significant rate of fall off from one line of therapy to another, this varying from disease to disease. It is estimated that this would reduce the number of patients requiring testing to about one third of the 100,000 figure. The advantages of this approach are that the timing of testing then fits in to the relevant treatment pathway and there is equity between cancers. The disadvantage is that if testing is slow, patients who are fit for treatment at the time of testing may not be fit enough to start treatment at the time of the positive NTRK gene fusion result.
2. **Test patients on the likelihood of finding NTRK gene fusions** There are some very rare cancers with very high incidences of NTRK gene fusions (>75%), some rare cancers with high incidences (20-40%), some uncommon cancers with modest NTRK gene fusion expression (3-20%) and all the commoner cancers with low incidences (3% or less). The number of patients for testing with anything other than low incidences levels is small (probably <2000/year). The difficulty in ranking the incidences and thus the order of phasing in NGS panel testing is that the literature varies as to the incidence of NTRK gene fusions in individual diseases. If there was robust data as to incidence levels and for a practical phased implementation of NGS panel testing, the low incidence group (3% or less) could be split into low and very low incidence groups. However, the data is insufficient at present to be able to do this and would be contestable.
3. **Test first those patients who are currently having single gene testing and are in any case planned to move onto panel testing in the NHS Long Term Plan** This applies to 30,000 patients with NSCLC, colorectal cancer, melanoma, ovarian cancer and some patients with breast cancer who have single gene tests, the results of which directly change patient management. All of these cancers have low incidences of NTRK gene fusions. The equity issue arises as to the exclusion of types of cancer which currently do not have actionable genomic changes.
4. **Test those patients most likely to benefit from TRK inhibitors** Whilst this is the standard approach used in Health Technology Assessment processes if subgroup data is robust, the difficulty of this approach for TRK inhibitors is that there is insufficient data at present on which to rank chance of benefit and duration of benefit.
5. **Test those patients with the most 'unmet need' eg those patients in whom there are no standard therapies or only those patients in whom there is only 1 standard therapy** Such an approach would be complicated, contestable and have equity issues.
6. **Combination of any of above**

Conclusion

NHS England has examined all of the above options and has concluded that the only feasible way of phasing the introduction of NGS panel testing, until such time that testing of all the relevant patients is possible at the time of commencing on the locally advanced/metastatic disease systemic therapy care pathway, is to test at the time of failure of systemic therapies which are considered to be standard and thus commissioned by NHS England. NHS England recognises the need for speed of NGS panel testing at this point in the systemic therapy care pathway.

NHS England also recognises that it needs to set an implementation timetable by which phased testing has been completed. It also needs to set out the working arrangements by which patients in the catchment area of one GLH which is still in the set-up phase of NGS panel testing can still access testing via their regional GLH working with another GLH that has its NGS panel testing processes up and running.

It is envisaged that by the end of 2020/21, NGS panel testing for NTRK gene fusions will be working at the rate of testing 30,000 patients/year (for those who have failed standard systemic therapies for locally advanced/metastatic disease) and by the end of 2021/22, such testing will be operational at the rate of 100,000/year (for those embarking on the locally advanced/metastatic systemic therapies care pathways).

Prof Peter Clark

Clinical Lead for the Cancer Drugs Fund

November 2019