NICE National Institute for Health and Care Excellence

# Single Technology Appraisal

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

# Committee Papers

NICE National Institute for Health and Care Excellence

**NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE**

**SINGLE TECHNOLOGY APPRAISAL**

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

**Contents:**



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

# Pre-meeting briefing

## Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

1

# Key issues – clinical effectiveness

- The company have positioned lenvatinib as a potential treatment for people with Child-Pugh Class A liver function, is this appropriate?
- Is it appropriate to exclude BSC as a comparator?
- Is the REFLECT trial generalisable to clinical practice in the NHS?
  - Post progression anti-cancer treatments were received in both arms
  - Trial population had stage B or C HCC, Child-Pugh class A and ECOG PS 0 or 1
- Is it appropriate to use results from the full population (rather than Western subgroup)?
- What censoring rules are most appropriate?
- Have the results of the REFLECT trial showed non-inferiority compared with sorafenib?

2

# Key issues – cost-effectiveness

- Is the company's covariate adjustment of survival models appropriate?
- Is PFS likely to be overestimated when FDA censoring method is applied?
- What parametric curve should be used to extrapolate PFS?
- Should drug wastage be included and how?
  - 7 day wastage or ERG scenario using planned daily number of capsules
- More patients in sorafenib arm received post-progression anti-cancer treatment in REFLECT
  - In UK clinical practice, would post-progression treatments be used?
  - Should adjustments be made to overall survival?
- Is end of life criteria met?
- Most plausible ICER?

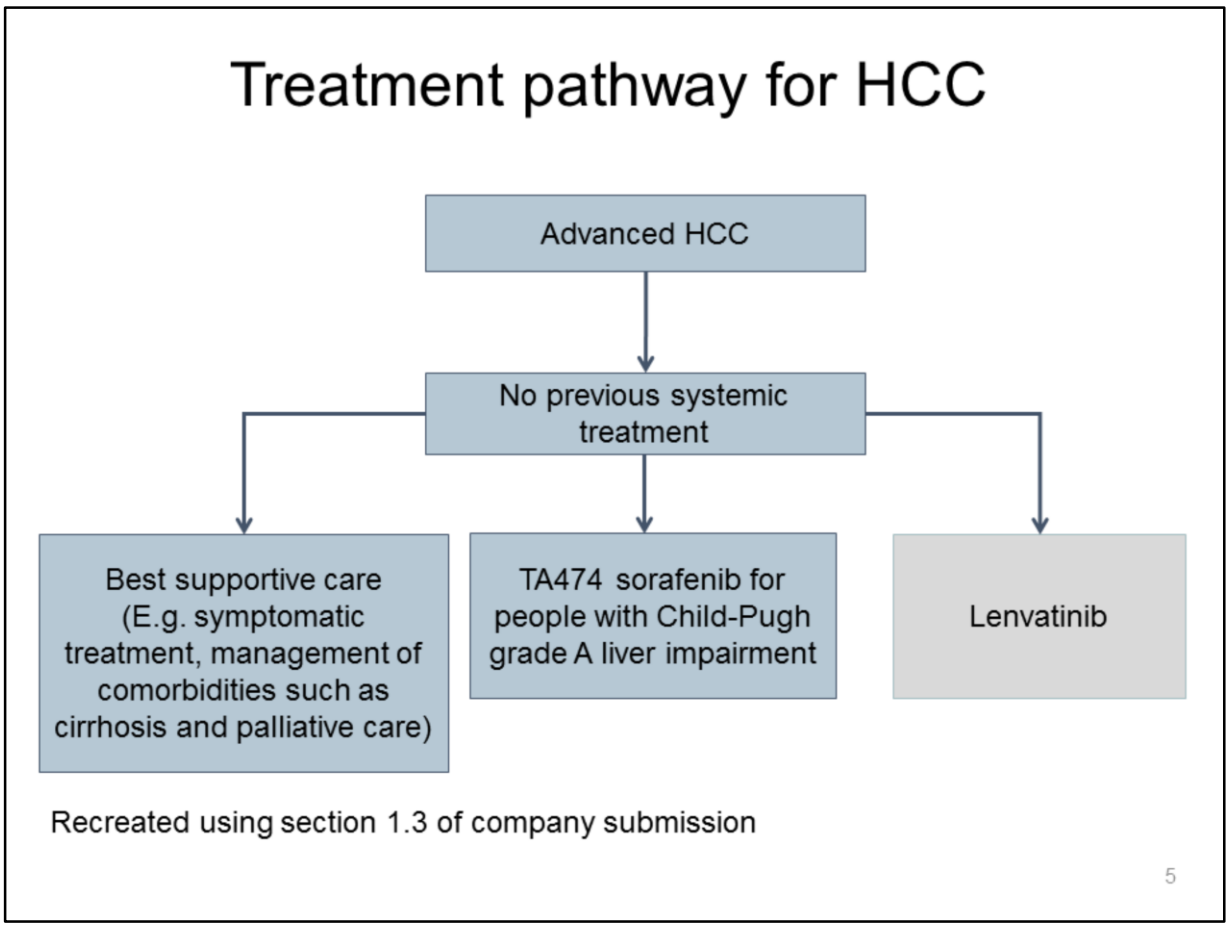
3

# Hepatocellular carcinoma (HCC)

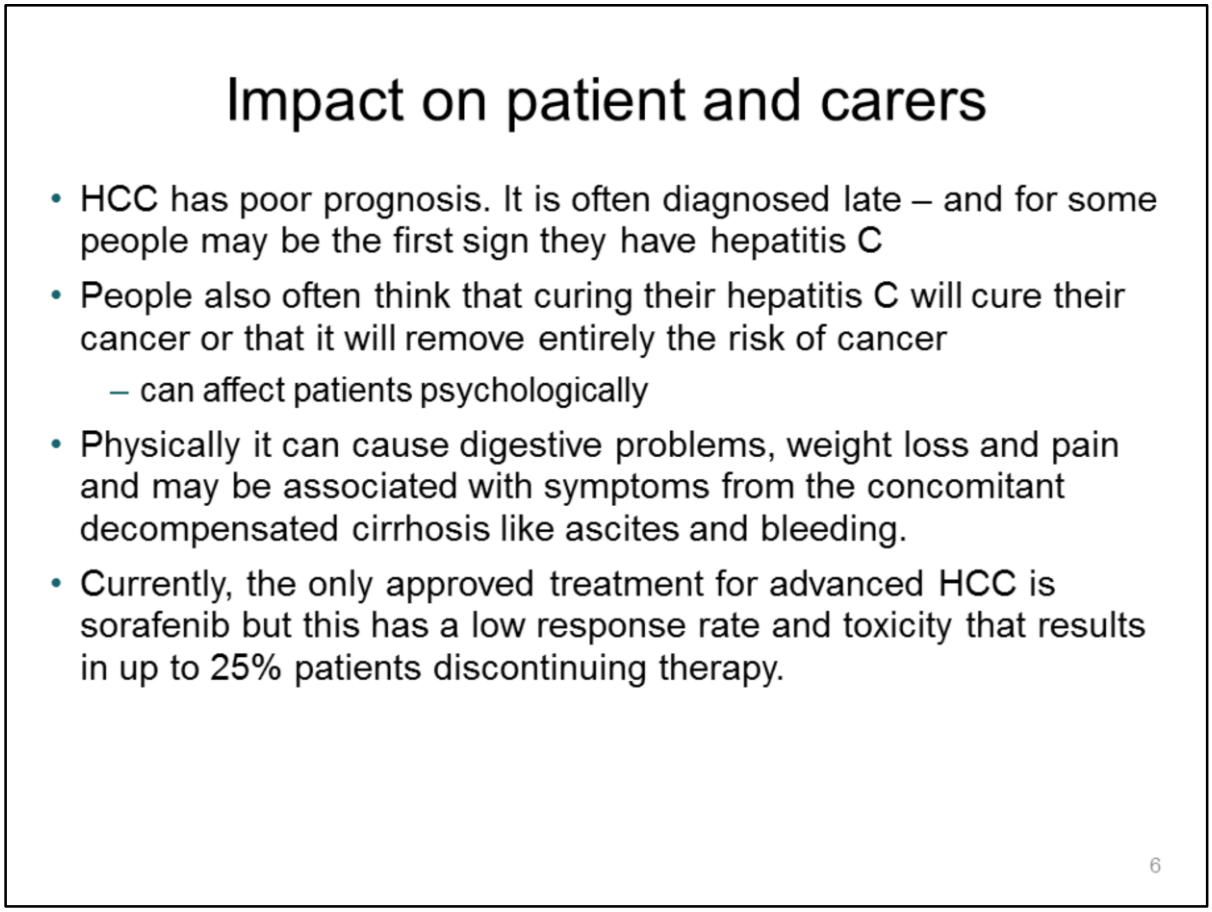
- HCC is the most common type of primary liver cancer in England with 2,456 cases diagnosed in 2015.
- It is commonly associated with cirrhosis (scarring of the liver), which can be caused by excessive alcohol intake, viral infections such as hepatitis B or C, or other conditions.
- Early stage HCC may be treated with potentially curative surgery (hepatic resection), or percutaneous radiofrequency/thermal ablation in patients with well-preserved liver function, or liver transplantation for those with impaired liver function.
- In the UK, only about 30% of patients with HCC are suitable for curative therapy such as liver transplantation, local resection or radiofrequency ablation or palliative chemoembolization.
- Treatment is palliative rather than curative for people with more advanced disease (include interventional procedures such as transarterial chemoembolisation or selective internal radiation therapy, and external beam radiotherapy).
- TA474 recommends sorafenib as an option for treating advanced HCC only for people with Child-Pugh grade A liver impairment.

4

See sections B.1.3 (page 11) of company submission and 2.1 (page 23) of ERG report for more details.




See sections B.1.3 (page 11) of company submission and 2.1 (page 23) of ERG report for more details.




Received submissions from The Hepatitis C Trust, BSG Liver Section, 1 patient expert and 2 clinical experts

# Lenvatinib
## Patient, professional & expert views

- Lenvatinib would improve first line systemic therapy (sorafenib has response rate of around 2% and a toxicity profile that results in up to 25% patients discontinuing therapy).
- Both sorafenib and lenvatinib are administered in the outpatient setting and require outpatient monitoring
  - Clinical review at 2 weekly intervals (first 2 months) and 4 weekly thereafter to manage toxicity and dose adjustment, therefore no difference in terms of clinical resources
- Later onset of cancer symptom (role function, pain, diarrhoea, nutrition and body image) deterioration compared with sorafenib.
- Common adverse events of lenvatinib were hypertension, diarrhoea, fatigue, decreased appetite and weight.
- Lenvatinib has a different side effect profile than sorafenib and that choice may be beneficial for patients.
  - Hypertension more common with lenvatinib (use anti-hypertensives). Diarrhoea and planter palmer erythema more common with sorafenib, (use anti-diarrheal medication and topical creams).

7

Received submissions from The Hepatitis C Trust, BSG Liver Section, 1 patient expert and 2 clinical experts

# Decision problem

| | NICE scope | Company |
|---|---|---|
| **Population** | Adults with unresectable hepatocellular carcinoma who have not previously received systemic treatment | Submission covers **Child-Pugh Class A liver function**, in line with REFLECT trial. Also consistent with TA474 sorafenib |
| **Intervention** | Lenvatinib | |
| **Comparators** | • Sorafenib<br>• Best supportive care (BSC) | Company do not consider BSC to be a relevant comparator because clinical input suggests this is a small proportion of the population (<5%). In the overall population, almost all people will be eligible for systemic therapy. |
| **Outcomes** | Overall survival, progression-free survival, response rate, adverse effects of treatment, health-related quality of life | |

8

See sections B.1.1 (table 1 on page 9) of company submission and 3 (table 3 on page 28) of ERG report for more details.

CONFIDENTIAL

# Lenvatinib (Eisai)

| | | |
|---|---|---|
| **Mechanism of action** | | Multi-kinase inhibitor and selectively inhibits the kinase activities of all vascular endothelial growth factor receptors, in addition to other proangiogenic and oncogenic pathways |
| **Administration and dosage** | | • Oral capsules<br>• Recommended daily dose: 8 mg (2 x 4 mg capsules) if body weight <60 kg and 12 mg (3 x 4 mg capsules) if body weight ≥60 kg. |
| **Marketing authorisation** | **New (subject of this appraisal)** | • Positive CHMP expected July 2018: 'treatment of adult patients who have received no prior systemic therapy for HCC' |
| | **Existing licensed indications** | • For progressive, locally advanced or metastatic, differentiated thyroid carcinoma, refractory to radioactive iodine (ID1059 expected August 2018). |
| **List price** | | £1,437.00 per pack of 30 x 4 mg capsules<br>Cost per cycle: £3,152 (dosing from REFLECT), average cost of a course of treatment (including PAS) is [redacted] |
| **PAS discount** | | Simple discount (magnitude: commercial in confidence) |

9

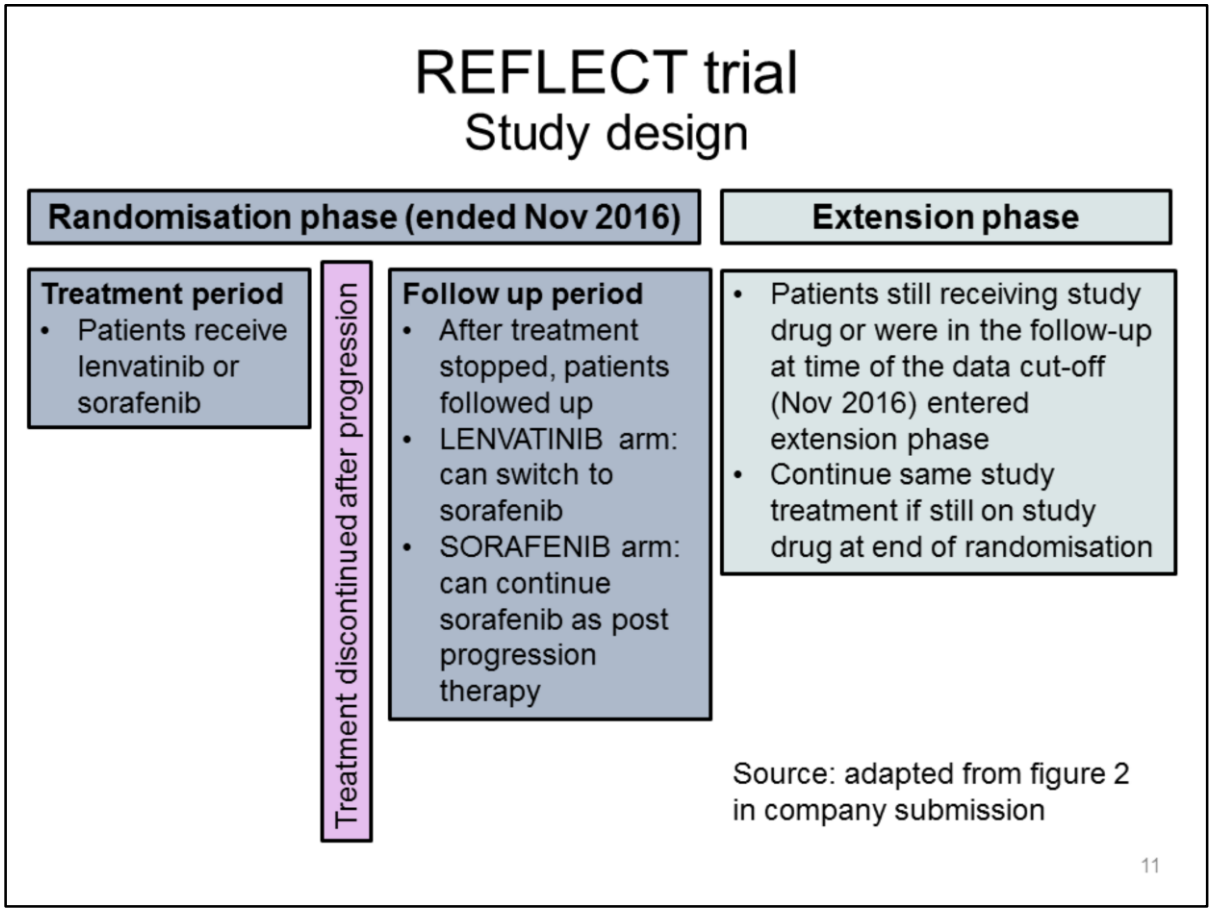
See sections B.1.2 (table 2 on page 10) of company submission and 3.2 (table 5 on page 32) of ERG report for more details.

# REFLECT trial
## Summary

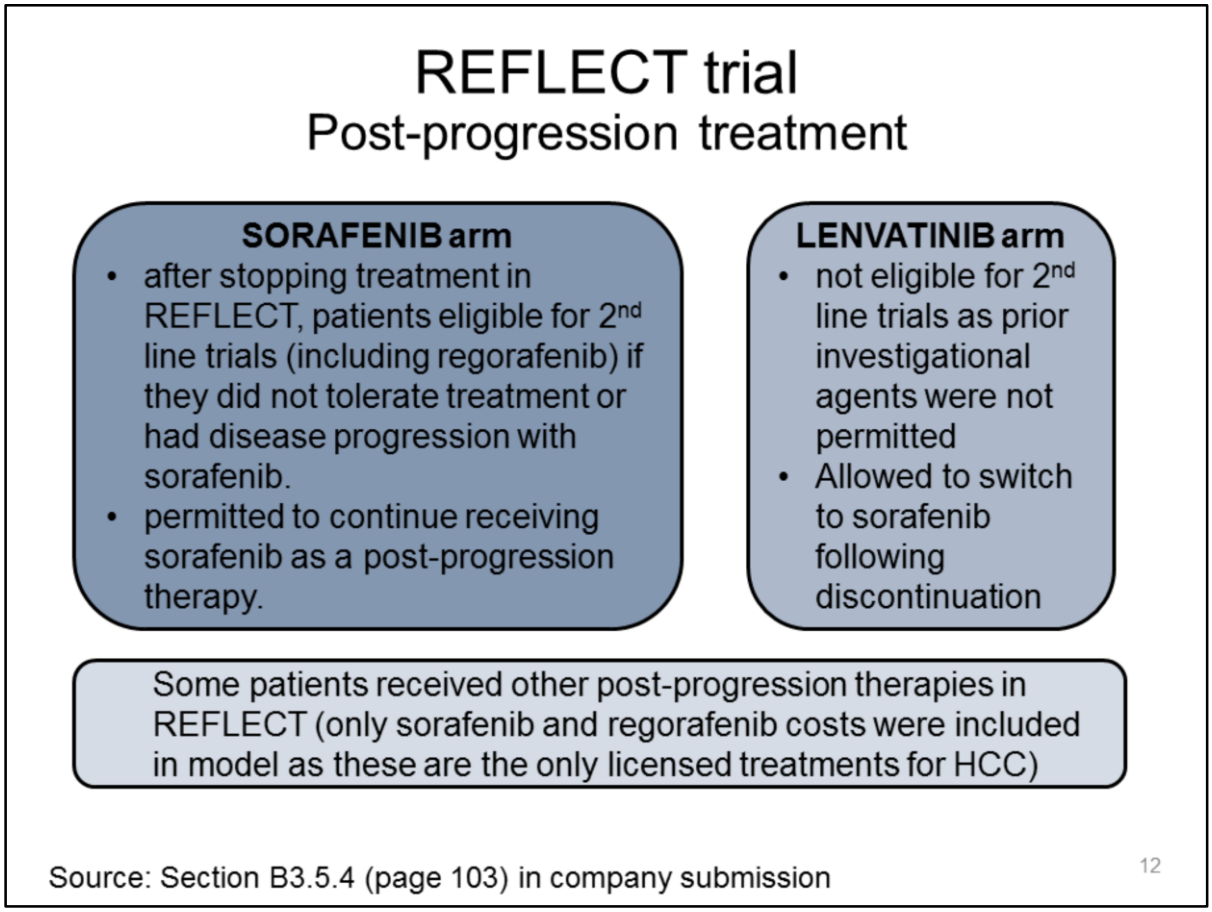
| | REFLECT |
|---|---|
| **Study design** | International, multicentre, randomised, open-label, phase 3 study (non-inferiority) |
| **Population** | 954 adults with histologically or cytologically confirmed diagnosis of unresectable HCC or clinically confirmed diagnosis of HCC, Child-Pugh class A, ECOG PS 0 or 1, stage B or C (based on BCLC staging system) with no previous systemic anti-cancer therapy. |
| **Intervention** | Lenvatinib (12 mg [if baseline body weight ≥60 kg] or 8 mg [if baseline body weight <60 kg] once daily), n=478 |
| **Comparator** | Sorafenib (400 mg twice daily), n=476 |
| **Concomitant medicines** | The following were **not** allowed during trial: surgery or radiotherapy for the treatment of HCC, systemic therapy, antiplatelet agents and anticoagulants that required monitoring (e.g. warfarin) |
| **Subgroups** | Pre-specified analyses based on age (≤65, ≥65 to <75 years and ≥75 years), sex and aetiology (HBV, HCV and alcohol). |
| **Location** | 20 countries including 20 patients from UK |

10

See sections B.2.3 (page 19) of company submission and 4.2 (page 40) of ERG report for more details.




See sections B.2.3.1.3 (page 19) of company submission for more details.




See sections B.2.6.1.1.3 (page 39) of company submission and 4.2.1 (page 44) of ERG report for more details.

CONFIDENTIAL

# REFLECT post-progression treatment
## Regional subgroups and full population

- Post-progression treatment differed by treatment arm (43% in lenvatinib and 51% in sorafenib arm (longer OS for these patients). Imbalance in Western region.

| | Lenvatinib | | | Sorafenib | | |
|---|---|---|---|---|---|---|
| **Anti-cancer treatment during follow-up** | Western (N=157) | Asia-Pacific (N=321) | Total (N=478) | Western (N=157) | Asia-Pacific (N=319) | Total (N=476) |
| **Any anti-cancer therapy*** | 44 (28.0) | 162 (50.5) | 206 (43.1) | 71 (45.2) | 172 (53.9) | 243 (51.1) |
| **Any anti-cancer medication**** | 41 (26.1) | 115 (35.8) | 156 (32.6) | 61 (38.9) | 123 (38.6) | 184 (38.7) |
| **Any anti-cancer procedure** | 11 (7.0) | 111 (34.6) | 122 (25.5) | 18 (11.5) | 112 (35.1) | 130 (27.3) |
| **Targeted therapy†** | [redacted] | NR | [redacted] | [redacted] | NR | [redacted] |
| **Sorafenib‡** | [redacted] | NR | [redacted] | [redacted] | NR | [redacted] |
| **Regorafenib** | [redacted] | NR | [redacted] | [redacted] | NR | [redacted] |

*includes anti-cancer procedures and medications **not given for a procedure †any antineoplastic & immunomodulating agent ‡ In sorafenib arm, patients continued sorafenib after progression.

Source: Table 12 in company submission, tables 4 & 5 in clarification response

13

See sections B.2.6.1.1.3 (page 39) of company submission and 4.2.1 (page 44) of ERG report for more details.

# REFLECT baseline characteristics
## Full population

| Characteristic | Lenvatinib (n=478) | Sorafenib (n=476) |
|---|---|---|
| Age, mean (SD) | 61.3 (11.7) | 61.2 (12.0) |
| ECOG PS 0 | 304 (63.6) | 301 (63.2) |
| Child-Pugh score 5 (Class A) | 368 (77.0) | 357 (75.0) |
| Child-Pugh score 6 (Class A) | 107 (22.4) | 114 (23.9) |
| Concomitant systemic antiviral therapy for hepatitis B or C | 163 (34.1) | 149 (31.3) |
| Any previous radiotherapy | 49 (10.3) | 60 (12.6) |
| Prior anti-cancer procedures* | 327 (68.4) | 344 (72.3) |
| AFP levels ≥200 ng/mL | 222 (46.4) | 187 (39.3) |
| Aetiology of HCC-Hepatitis B | 251 (52.5) | 228 (47.9) |
| Aetiology of HCC-Hepatitis C | 91 (19.0) | 126 (26.5) |
| Aetiology of HCC-Alcohol | 36 (7.5) | 21 (4.4) |

All data reported are n (%) unless otherwise stated, *including radiotherapy
Abbreviations: AFP, alpha-fetoprotein. Red box = company highlighted imbalances

14

Source: Tables 6 & 7 in company submission

See sections B.2.3.2 (page 27) of company submission and 4.2.2 (page 47) of ERG report for more details.

In REFLECT, 99% of patients had Child-Pugh class A liver function. The remainder had class B liver function.

The company highlighted imbalances in baseline characteristics but the ERG considered these to be generally well balanced.

# REFLECT baseline characteristics

## Geographical region

| Characteristic | Lenvatinib (n=478) | | Sorafenib (n=476) | |
|---|---|---|---|---|
| | Asia-Pacific | Western | Asia-Pacific | Western |
| Age, mean (SD) | 60.0 (11.76) | 63.8 (11.15) | 60.2 (11.87) | 63.3 (12.06) |
| ECOG PS 0 | 206 (64.2) | 98 (62.4) | 204 (63.9) | 97 (61.8) |
| AFP levels ≥200 ng/mL | 157 (48.9) | 65 (41.4) | 137 (42.9) | 50 (31.8) |
| Aetiology of HCC-Hepatitis B | 212 (66.0) | 39 (24.8) | 197 (61.8) | 31 (19.7) |
| Aetiology of HCC-Hepatitis C | 50 (15.6) | 41 (26.1) | 70 (21.9) | 56 (35.7) |
| Aetiology of HCC-Alcohol | 17 (5.3) | 19 (12.1) | 8 (2.5) | 13 (8.3) |

All data reported are n (%) unless otherwise stated, *including radiotherapy
Abbreviations: AFP, alpha-fetoprotein. Red box = company highlighted imbalances

Source: Table 8 in ERG report

15

See sections B.2.3.2 (page 27) of company submission and table 8 (page 48) of ERG report for more details.

While differences are noted between the full population of REFLECT and the Western subgroup, evidence from separate Western and Asia–Pacific trials for sorafenib in HCC (SHARP and the Asia–Pacific study) suggest that expected baseline differences based on geographical location but do not necessarily impact relative treatment effects. The ERG does not consider there to be sufficient evidence to justify the inevitable loss of precision on effect estimates by focusing on the Western subgroup rather than the full population of REFLECT, in terms of comparability to a UK population.

CONFIDENTIAL

# REFLECT trial results

## Full population

| Outcome | LEN (n=478) | SOR (n=476) | Result (95% CI) |
|---|---|---|---|
| OS* (months) | Median 13.6 (12.1 to 14.9) | Median 12.3 (10.4 to 13.9) | HR 0.92 (0.79 to 1.06) |
| Adjusted OS** | - | - | [redacted] |
| PFS* (months) investigator-assessed | Median 7.4 (6.9 to 8.8) | Median 3.7 (3.6 to 4.6) | HR 0.66 (0.57 to 0.77)‡ |
| mRECIST & amended censoring | - | - | [redacted] |
| PFS* (months) independent review | Median 7.3 (5.6 to 7.5) | Median 3.6 (3.6 to 3.7) | - |
| mRECIST | - | - | HR 0.64 (0.55 to 0.75) |
| RECIST v1.1 | - | - | HR 0.65 (0.56 to 0.77) |

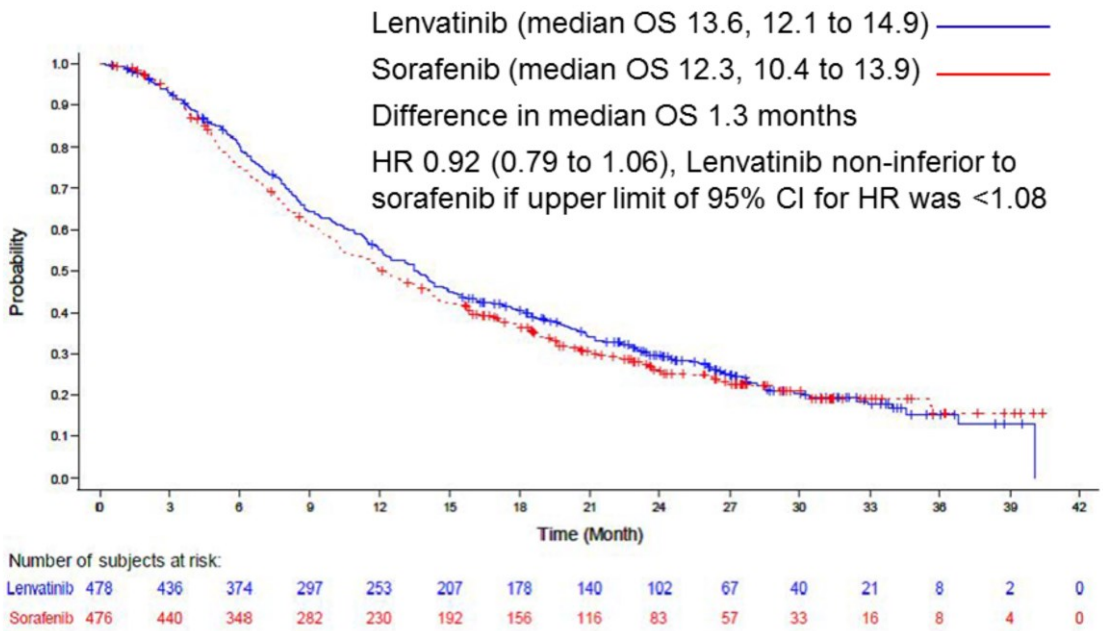
*Stratified by region (Asia–Pacific; Western), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg). **Adjusted by post-treatment anti-cancer treatment (post-hoc analyses). ‡From stratified cox model and censored if discontinued treatment for any reason other than progression. Abbreviations: LEN Lenvatinib; NR not reported; OS overall survival; PFS progression free survival; SOR sorafenib

16

See sections B.2.6. (page 36) of company submission and 4.3 (page 55) of ERG report for more details.

# REFLECT

## OS (not adjusted for post-progression treatment)




Number of subjects at risk:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lenvatinib | 478 | 436 | 374 | 297 | 253 | 207 | 178 | 140 | 102 | 67 | 40 | 21 | 8 | 2 | 0 |
| Sorafenib | 476 | 440 | 348 | 282 | 230 | 192 | 156 | 116 | 83 | 57 | 33 | 16 | 8 | 4 | 0 |

Source: Figure 3 in company submission

17

CONFIDENTIAL

# ERG report
## Censoring

- Censoring in primary analyses (PFS and TTP) based on guidance from FDA (censoring at treatment discontinuation if no disease progression).
  - likely to favour lenvatinib because treatment discontinuation for reasons other than progression (that is, due to TEAEs or patient choice) more common in lenvatinib group than sorafenib group.
  - At clarification company reported sensitivity analyses with censoring in line with EMA guidance
- ERG: consistency in direction of effect provides robust evidence of PFS benefit with lenvatinib, although rules for censoring mean extent of benefit may be [redacted]

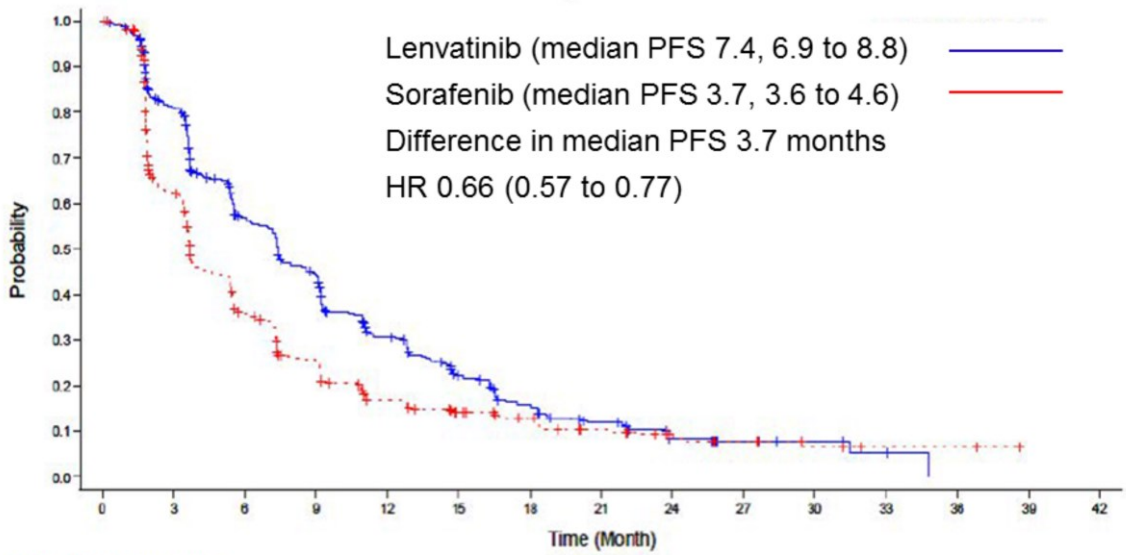
| Outcome | Lenvatinib | Sorafenib | Lenvatinib vs. sorafenib |
|---|---|---|---|
| PFS (months) using FDA | Median 7.4 | Median 3.7 | HR 0.66 (0.57 to 0.77) |
| PFS (months) using EMA* | Median [redacted] | Median [redacted] | [redacted] |

*EMA guidance, included all progressions and deaths as events (not censored at treatment discontinuation if no disease progression).
Abbreviations: HR, hazard ratio

18

See table 13 (page 59) of ERG report for more details.

# REFLECT
## Investigator-assessed PFS (FDA censoring)




Number of subjects at risk:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lenvatinib | 478 | 345 | 223 | 172 | 106 | 69 | 44 | 28 | 14 | 9 | 4 | 2 | 0 | 0 |
| Sorafenib | 476 | 262 | 140 | 94 | 56 | 41 | 33 | 22 | 14 | 9 | 4 | 2 | 2 | 0 |

Source: Figure 4 in company submission

19

CONFIDENTIAL

# Other REFLECT trial results

## Full population

| Outcome | LEN (n=478) | SOR (n=476) | Result (95% CI) |
|---|---|---|---|
| Disease progression | [redacted] | [redacted] | P-value not reported |
| Median time to progression (months) | 8.9 (7.4, 9.2) | 3.7 (3.6, 5.4) | P-value not reported |
| ORR | 115 (24.1) | 44 (9.2) | OR 3.13 (2.15 to 4.56) |
| EQ-5D (HUI or VAS) | Not reported | Not reported | no statistically significant differences in HUI or VAS scores between groups at Cycles 3, 6, 9, 12, 15, or 18 (p>0.05) |

Abbreviations: LEN Lenvatinib; HUI, health utility index; OR odds ratio; ORR objective response rate; SOR sorafenib; VAS, visual analogue scale

- Quality of life (QoL) broadly same across both treatments for majority of function and symptom areas, but there was a clinically meaningful delay in worsening of QoL for lenvatinib compared with sorafenib across several EORTC measures:
  - role functioning (HR 0.83, 0.71 to 0.97), pain (HR 0.82, 0.70 to 0.95), diarrhoea (HR 0.53, 0.45 to 0.63), body image (HR 0.79, 0.68 to 0.93), nutrition (HR0.81, 0.68 to 0.95)

20

See sections B.2.6. (page 36) of company submission and 4.3 (page 55) of ERG report for more details.

CONFIDENTIAL

# REFLECT trial
# Adverse events

Draft SPC states:

- [redacted]
  - ID1059 lenvatinib and sorafenib for DTC *"Side effects such as sore hands and feet more common with sorafenib and hypertension was more common with lenvatinib…The clinical expert explained that additional clinical monitoring visits are needed when starting both treatments and that there is little effect on quality of life when treatment-related symptoms are quickly identified and treated."*
- [redacted]
- [redacted]

21

# REFLECT trial
## Adverse events

| Outcome | Lenvatinib (n=476) | | Sorafenib (n=475) | |
|---|---|---|---|---|
| | All | Western | All | Western |
| Mean treatment duration, months (SD) | 8.2 (7.04) | - | 6.0 (6.47) | - |
| Mean dose intensity, mg/day/patient (SD) | 9.4 (5.71) | 10.2 (9.23) | 663.8 (173.15) | 669.1 (162.45) |
| Any TEAE, n (%) | 470 (98.7) | - | 472 (99.4) | - |
| TEAE ≥Grade 3 | 357 (75.0) | - | 316 (66.5) | - |
| Grade 3 hypertension | 111 (23.3) | 36 (23.2) | 68 (14.3) | 18 (11.5) |
| Grade 3 weight decreased | 36 (7.6) | 18 (11.6) | 14 (2.9) | 11 (7.1) |
| Any serious AE | 205 (43.1) | - | 144 (30.3) | - |
| TEAEs leading to study drug withdrawal | 94 (19.7) | - | 69 (14.5) | - |
| TEAEs leading to study drug dose reduction | 184 (38.7) | - | 185 (38.9) | - |

Abbreviations: TEAE treatment emergent adverse event

22

See sections B.2.10 (page 52) of company submission and 4.3.7 (table 19 on page 70) of ERG report for more details.

CONFIDENTIAL

# REFLECT trial results

## pre-planned subgroups

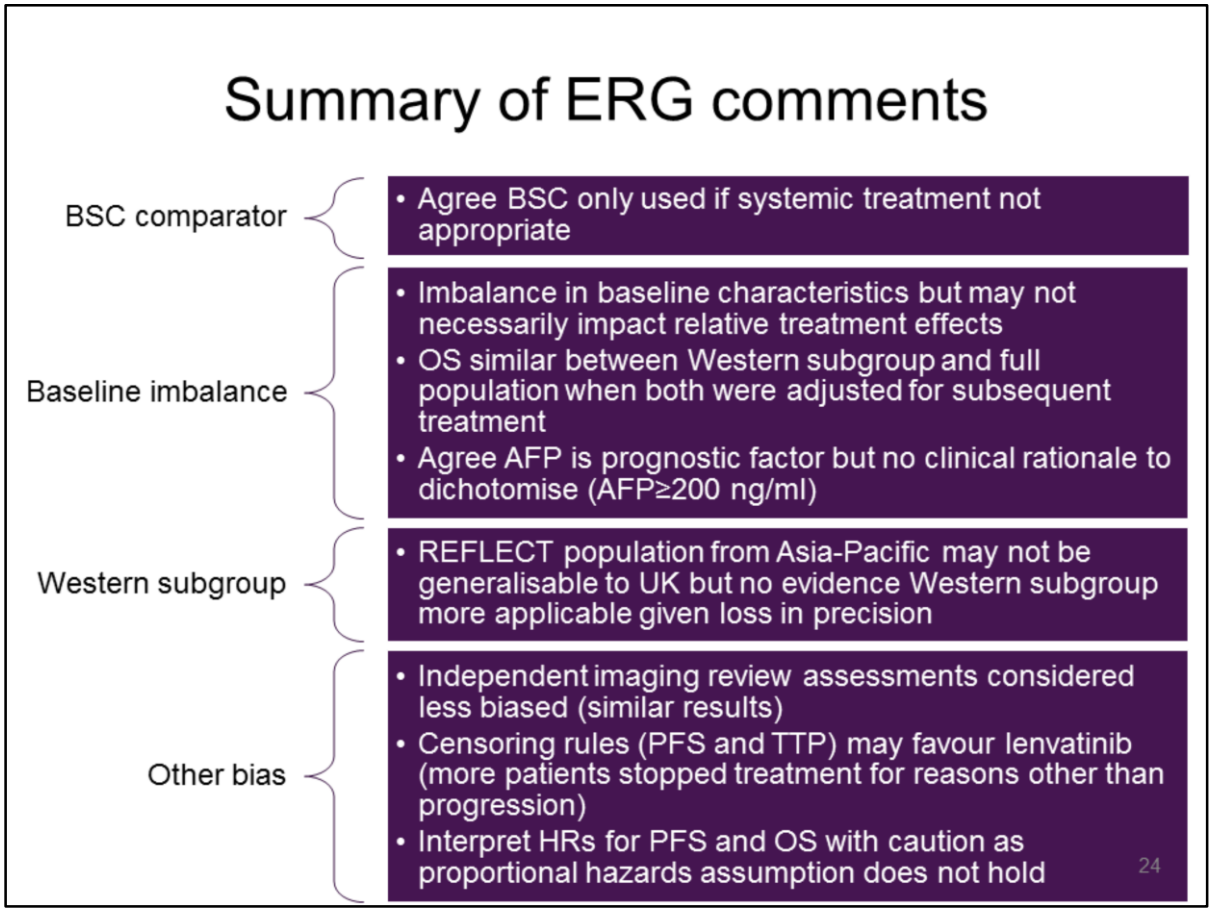
| Outcome | Hazard ratio (95% CI) |
|---|---|
| OS* | Western: 1.08 (0.82 to 1.42)*, [redacted]<br>Asia-Pacific: 0.86 (0.72 to 1.02)*, [redacted] |
| OS* with additional covariates | Baseline AFP (<200 or ≥200 ng/mL): 0.86 (0.74 to 1.00) |
| | Region: 0.92 (0.79 to 1.06)** |
| | Aetiology of HCC (HBV, HBC, alcohol): 0.86 (0.72 to 1.01) |
| Median PFS (months) investigator-assessed | Western: 0.81 (0.61 to 1.08)<br>Asia-Pacific: 0.61 (0.51 to 0.73) |
| | Baseline AFP <200 ng/mL: 0.68 (0.55 to 0.83)<br>Baseline AFP ≥200 ng/mL: 0.59 (0.47 to 0.75) |
| | Aetiology HBV: 0.62 (0.50 to 0.75), HCV: 0.78 (0.56 to 1.09), Alcohol: 0.27 (0.11 to 0.66) |

*Stratified by region (Asia–Pacific; Western), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg). **status of subsequent anti-cancer therapy (yes/no) used as an additional covariate factor. Abbreviations: AFP, alpha-fetoprotein; LEN Lenvatinib; OS overall survival; PFS progression free survival; SOR sorafenib

Source: section B2.6.1.1.2 (page 38) of company submission and table 21 in company submission appendix, tables 11 and 57 in ERG report

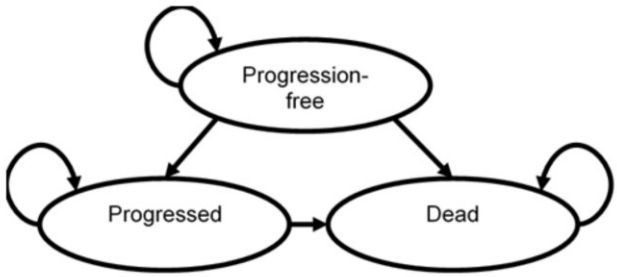
23

See sections B2.6.1.1.2 (page 38) and B.2.7 (page 51) in company's submission and table 11 (page 56) and 57 (page 149) of ERG report for more details.




See sections 1.4, 1.5 and 4.5 of the ERG report for more details.

# Company model
## Model structure




- Partitioned survival model (commonly used in late-stage oncology modelling)
- distribution of patients across all health states at each cycle is modelled, defined by OS and PFS curves.
- REFLECT provides relatively complete observed PFS & OS
  - 64.4% in the lenvatinib arm had experienced disease progression
  - 73.4% died at the end of trial (data cut-off Nov 2016).
- Use patient-level data from RELECT, lifetime horizon, half-cycle correction and 3.5% discount rate

**ERG comments:**
- Model structure and approach is appropriate
- Identified error as there is inconsistency in application of half-cycle correction to costs
- At clarification, company did not consider this an error
- Corrected in ERG's preferred base case

25

See sections B.3.2.2 (page 73) of company submission and 5.4.4 (page 83) of ERG report for more details.

## Summary of base case

| | Company | ERG |
|---|---|---|
| Clinical effectiveness | • Adjust for imbalance in baseline characteristics (but not post-progression treatment & exclude post-progression AE)<br>• Independent models for each arm (PH assumption not supported)<br>• Censor at discontinuation if no progression | Use OS adjusted for post-progression treatment *(scenarios: ↓ PFS treatment effect to assess censoring)* |
| Extrapolation | PFS: log-normal, OS: log-logistic, TTD: none (data almost complete) | PFS: gamma *(scenario: OS gamma)* |
| HRQoL | • No difference between LEN and SOR<br>• Disutilities for AEs not explicitly modelled | No changes |
| Costs | Resource use costs same for LEN and SOR in progression-free and progressed state. Drug wastage not included | Correct half-cycle error *(scenarios: full drug cost & no AE cost)* |
| Post-progression treatment | One-off cost and only sorafenib & regorafenib costs included | Remove post-progression costs *(scenarios: include post progression procedures & mortality adjustment)* |

26

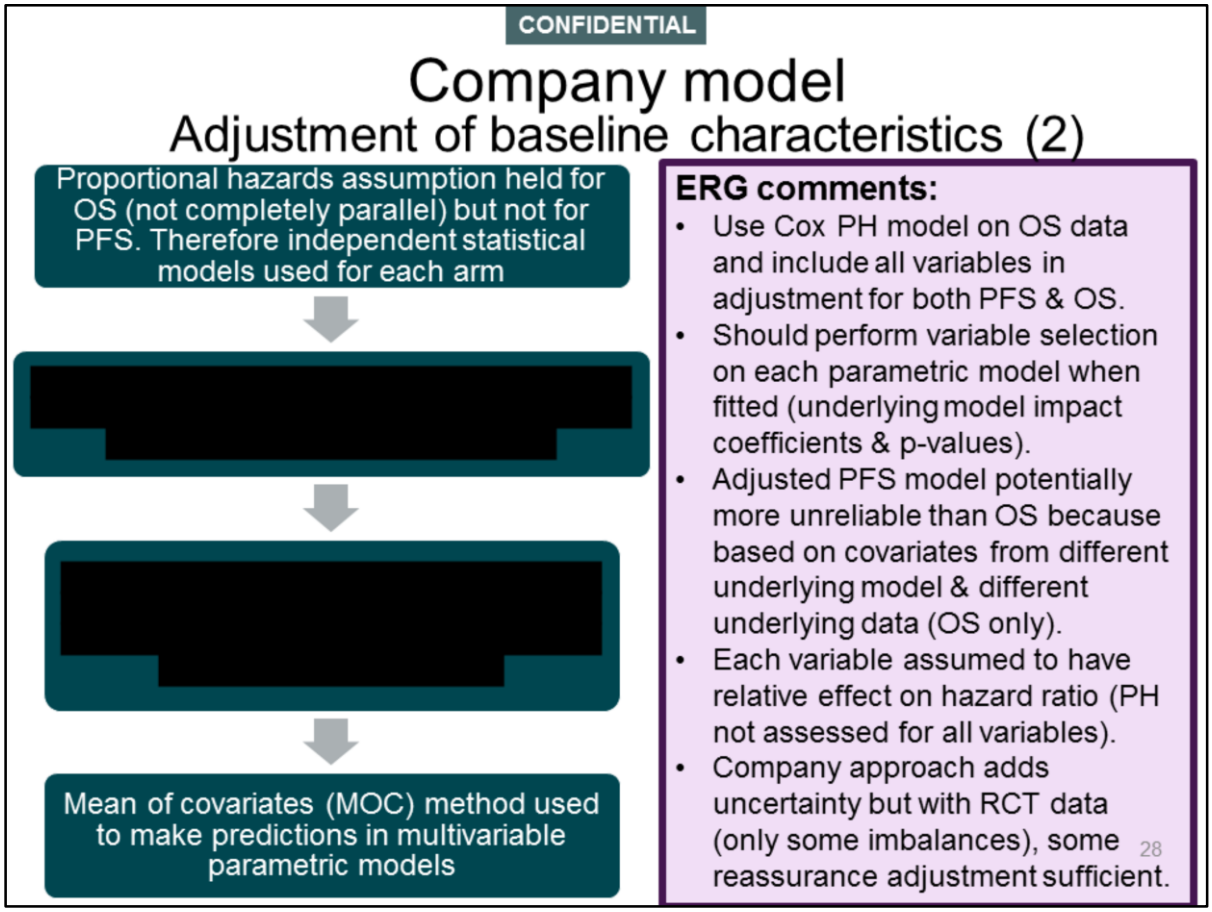
See sections B.3.6.1 (page 104) of company submission and 6.3 (table 48 on page 125) of ERG report for more details.

# Company model
## Adjustment of baseline characteristics (1)

- Imbalance in baseline characteristics:
  - AFP levels ≥200 ng/mL (strong independent predictor of outcomes regardless of treatment type, higher proportion in lenvatinib arm 46% compared with 39% in sorafenib arm)
  - aetiology of HCC (HBV, HCV, alcohol)
- Company base case adjusts for imbalance using multivariable adjustments to the PFS and OS curves
  - scenario 1: unadjusted parametric models
  - scenario 2: adjusts for AFP and stratification factors only
- Adjustment for covariates generally improves the precision and efficiency of the analysis and avoids conditional bias from chance covariate imbalance (guidance from European Medicines Agency)

27

See sections B.3.3.1.2 (page 81) of company submission and 5.4.5.2 (page 86) of ERG report for more details.




See sections B3.3.1 in company submission and section 5.4.5.3.1 in ERG report for more details.

An additional minor point regarding the covariate adjustment is that the company also provided an analysis using the corrected group prognosis (CGP) approach to estimate adjusted model outputs, as opposed to the MOC method used in the company's base case analysis. The results provided by the company were similar using the two approaches, so the ERG considers the methodological differences in the two approaches to have negligible impact on the final ICER.

# Company model
## Extrapolation

| Outcome | Extrapolation | Distribution in base case |
|---|---|---|
| PFS | • Extrapolation needed<br>• Survival for sorafenib was 6% at last observed data point (KM curve) | log-normal** |
| OS | • Extrapolation needed<br>• 73% in lenvatinib and 74% in sorafenib arm died at last observed data point | log-logistic* |
| TTD | • KM curve used<br>• data almost complete 0% lenvatinib and 4% sorafenib and assume 4% stopped treatment at end of follow up | N/A |

*Note: Weibull, exponential, log-logistic, log-normal, gamma and Gompertz distributions investigated for all outcomes (scenario analyses for all listed distributions).*

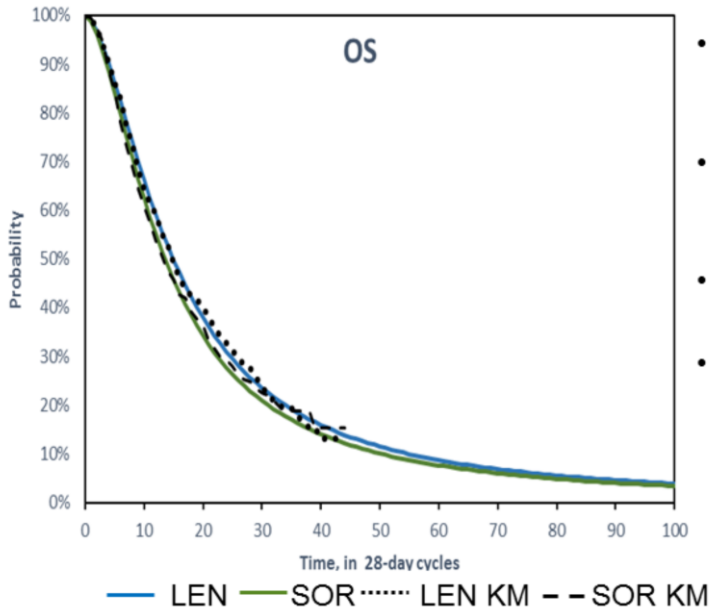
*company: consistent with the 2016 sorafenib reconsideration (2 real-world data sources showed small proportion survive for extended period of time, indicating log-normal curve is better fit than Weibull), as for certain ranges of the parameters, the shape of the log-normal and log-logistic hazard functions can be very similar.

**gamma distribution preferred for sorafenib arm but this led to extrapolations in which PFS for sorafenib exceeded that of lenvatinib (not clinically plausible, so not considered further).

29

See sections B.3.3.1.3 (page 84) of company submission for more details.

# ERG report
## Extrapolation of overall survival




- Clinical expert advice to ERG: 10 year plausible max OS for untreated advanced HCC
- Extrapolated curved in line with this (approx. 1-2% alive in each group at 10 years)
- No changes to company OS extrapolation in ERG base case
- ERG scenario: Gamma curve because it was a similarly well fitted curve but with a slightly more conservative extrapolation

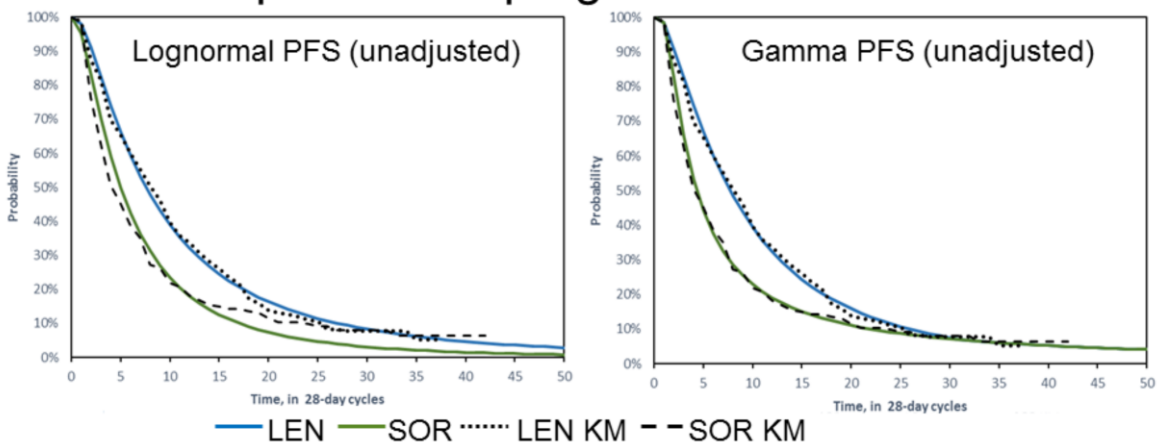
**ERG base case:** use log-logistic distribution to extrapolate OS (no change)
**ERG scenario:** use gamma distribution to extrapolate OS

30

See section 5.4.5.3.5 (page 94, figure 14 on page 95) of ERG report for more details.

# ERG report

## Extrapolation of progression-free survival




- Lognormal not a good fit for SOR. Company should have assessed unadjusted models against KM data.
- ERG's agrees with using same functional form across treatment groups but reasonable fit should be considered in both groups.
- Gamma model the best fit in terms of AIC and BIC & also good fit for lenvatinib.

**ERG base case:** use gamma model for PFS (not log-normal) and prevent curves from crossing over

31

See section 5.4.5.3.5 (page 94, figures 15 and 16 on page 96) of ERG report for more details.

# Company model
## Health related quality of life (HRQoL)

- EQ-5D-3L collected in REFLECT (UK value set) & AE disutility not applied
- Adjusted (by covariates that influence QoL) mean utility values for progression-free and progressed health states similar between the lenvatinib and sorafenib arms.
  - **Company base case**: utility values in full REFLECT population for both arms
  - **Company scenario**: committee preferred in TA474 sorafenib and TA514 regorafenib because base case may not capture reduced HRQoL after disease progression, particularly near end of life

| Health state | Utility value: mean (standard error) |
|---|---|
| **Progression-free** | 0.745 (0.0079) |
| **Progressed** | 0.678 (0.0118) |

Source: Table 33 in company submission

Small difference between progression-free and progressed utility values may be due to close post-progression measurement after disease progression

Clinical experts: post-progression estimate higher than expected, as significant impact on well-being

32

See sections B3.4.1 (table 32 on page 92) in company submission for more details.

# ERG report

## Adverse events (AEs)

| Company approach | ERG comments |
|---|---|
| Company model all grade 3 or 4 TEAEs with incidence >5%:<br>1. Hypertension (LEN 23% vs. SOR 14%)<br>2. Weight ↓ (8% vs. 3%)<br>3. Blood bilirubin ↑ (7% vs. 5%)<br>4. Proteinuria (6% vs. 2%)<br>5. Gamma-glutamyltransferase ↑ (6% vs. 4%)<br>6. Platelets ↓ (6% vs. 3%)<br>7. Aspartate aminotransferase ↑ (5% vs. 4%)<br>8. Diarrhoea* (4% both arms)<br>9. Fatigue* (4% both arms)<br>10. Hand-foot syndrome (3% vs. 11%)<br>11. Asthenia* (3% vs. 2%) | • Company's approach reasonable but excluded grade 5 AE. ERG's clinical experts: all relevant AEs included.<br>• Company did not include AEs associated with post-progression anti-cancer interventions but this is not model driver so unlikely to have large impact on ICER.<br>• At clarification company explain grade 5 AEs not included to avoid double counting mortality cost.<br>• **Company scenario**: apply cost of 1 hospitalisation to 12.8% in LEN and 7.6% in SOR arm with grade 5 AE (minimal impact on ICER).<br>• **ERG scenario:** Western AE data correctly used (minimal impact on ICER). |

*occurred in <5% of patients, but were also included as the company's clinical experts expected those AEs to have significant clinical or economic impacts

33

See section 5.4.7.1 (page 98, table 20 on page 72) in ERG report for more details.

CONFIDENTIAL

# Company model
## Treatment costs

- Drug costs based on mean observed dose in trials but drug wastage not included (scenario: 7 days wastage in line with TA474 sorafenib)
  - TA514 regorafenib: assuming additional days of drug wastage to model drug cost is arbitrary and associated with significant uncertainty
- At end of trial follow-up, all non-censored patients in LEN arm discontinued, 4% of non-censored patients in SOR arm remained on treatment (company assume no patients received SOR after follow-up)
- No administration or monitoring costs for LEN or SOR:
  - Outpatient administration (costs captured in background medical management)
  - Clinical expert: in practice, monitoring requirements same for both treatments (also in line with AG model for ID1059 for thyroid cancer)

| Use of post-progression therapy | Proportion leaving progression-free state | | Mean duration of post-progression therapy (days) | |
|---|---|---|---|---|
| | LEN | SOR | LEN | SOR |
| Sorafenib† | [redacted] | [redacted] | [redacted] | [redacted] |
| Regorafenib‡ | [redacted] | [redacted] | [redacted] | [redacted] |

† mean dose 663.8mg assumed to be the same as first line; ‡mean dose 144 mg taken from the RESORCE trial. Abbreviations: LEN lenvatinib; SOR sorafenib

34

Source: Table 42 in company submission

See sections B3.5.1 (page 94 and table 42 on page 104) in company submission for more details.

# Company model
## Health state costs

- Resource use from survey (source: TA189 & updated in TA474 sorafenib for HCC). Also used in TA514 (regorafenib) but new survey used for hospital admission rate.
- In TA474 committee preferred pooling the original and updated results of the resource use survey (but company do not have access to full data):
  - For original survey results: new unit costs applied
  - For updated survey results: 2015/16 costs were uplifted to 2016/17 (resource use not provided directly). Weighted average cost calculated based on the number of clinicians responding to each survey

| | Weighted average cost per cycle | |
|---|---|---|
| | Progression-free | Progressed |
| **Physician visits** | £159.63 | £384.40 |
| **Laboratory tests** | £161.78 | £135.56 |
| **Radiological tests** | £30.04 | £27.25 |
| **Hospitalisation** | £91.52 | £196.78 |
| **Hospital follow-up*** | £168.50 | £726.26 |
| **Social care*** | £21.19 | £1,066.07 |
| **Total** | £632.67 | £2,536.32 |

*Not reported in the 2016 reconsideration of sorafenib by NICE; these values are therefore based only on the survey results presented in the original sorafenib submission to NICE. 35

See section B3.5.2 (page 97 and table 38 on page 100) of the company submission for more details.

# ERG report
## Resource use and costs

- ERG disagrees with the company's decision to exclude drug wastage costs
  - TA474 agreed most plausible ICER should account for drug wastage for up to 7 days but ERG considered 7 days arbitrary
- Issues relating to implementation of post-progression anti-cancer treatment
  - some aspects of post-progression management costs, potentially double counted (minimal impact on ICER)
  - Clinical expert to ERG: patients in UK would not receive post-progression treatment, (regorafenib not recommended by NICE & counterintuitive to offer regorafenib/sorafenib after failure on same drug class)
- Post-progression drug costs applied to 'newly progressed' patients but company overestimate this.

**ERG scenario:** drug costs based on the planned number of daily capsules
**ERG scenario:** mortality adjustment for newly progressed patients

36

See section 5.4.9.6 (page 111) in ERG report for more details.

CONFIDENTIAL

# Company's base case results
## After clarification stage (PAS for lenvatinib only)

- Company's base case amended at clarification:
  - categorised Child–Pugh class as categorical variable (A vs B) rather than continuous (multivariate analyses were re-estimated)
  - corrected end of life care costs

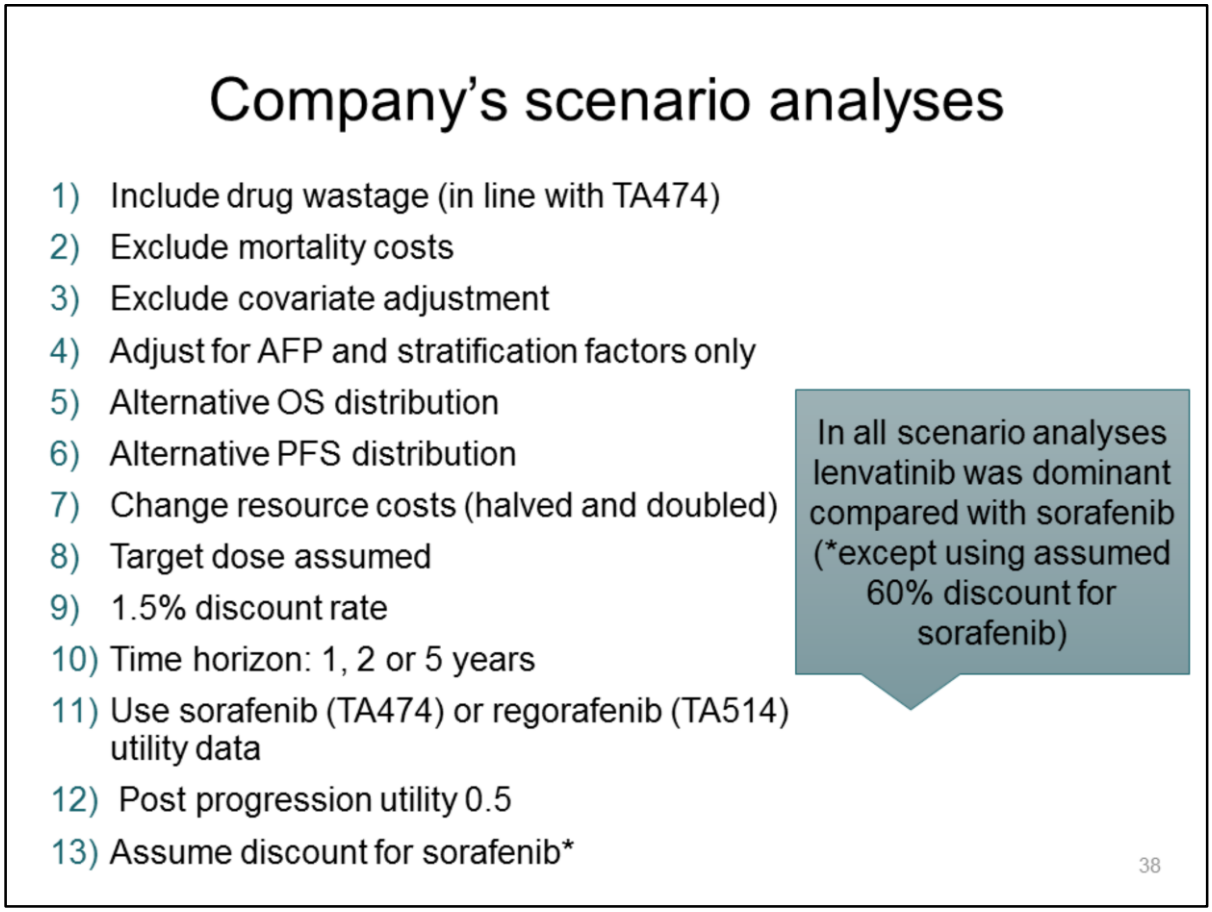
| | Total | | | Incremental | | | |
|---|---|---|---|---|---|---|---|
| **Drug** | Costs (£) | LYG | QALYs | Costs | LYG | QALYs | **ICER** |
| Deterministic | | | | | | | |
| **SOR** | £65,592 | 1.46 | 1.03 | - | - | - | - |
| **LEN** | [redacted] | 1.69 | [redacted] | [redacted] | 0.23 | [redacted] | Dominated |
| Probabilistic sensitivity analysis | | | | | | | |
| **SOR** | £65,688 | - | 1.03 | - | - | - | - |
| **LEN** | [redacted] | - | [redacted] | [redacted] | - | [redacted] | Dominated |

- Most influential parameters:
  - Baseline hazard for PFS and OS
  - Proportion of patients randomised to lenvatinib but used sorafenib as a post-progression treatment
  - Duration of post-progression sorafenib use after lenvatinib

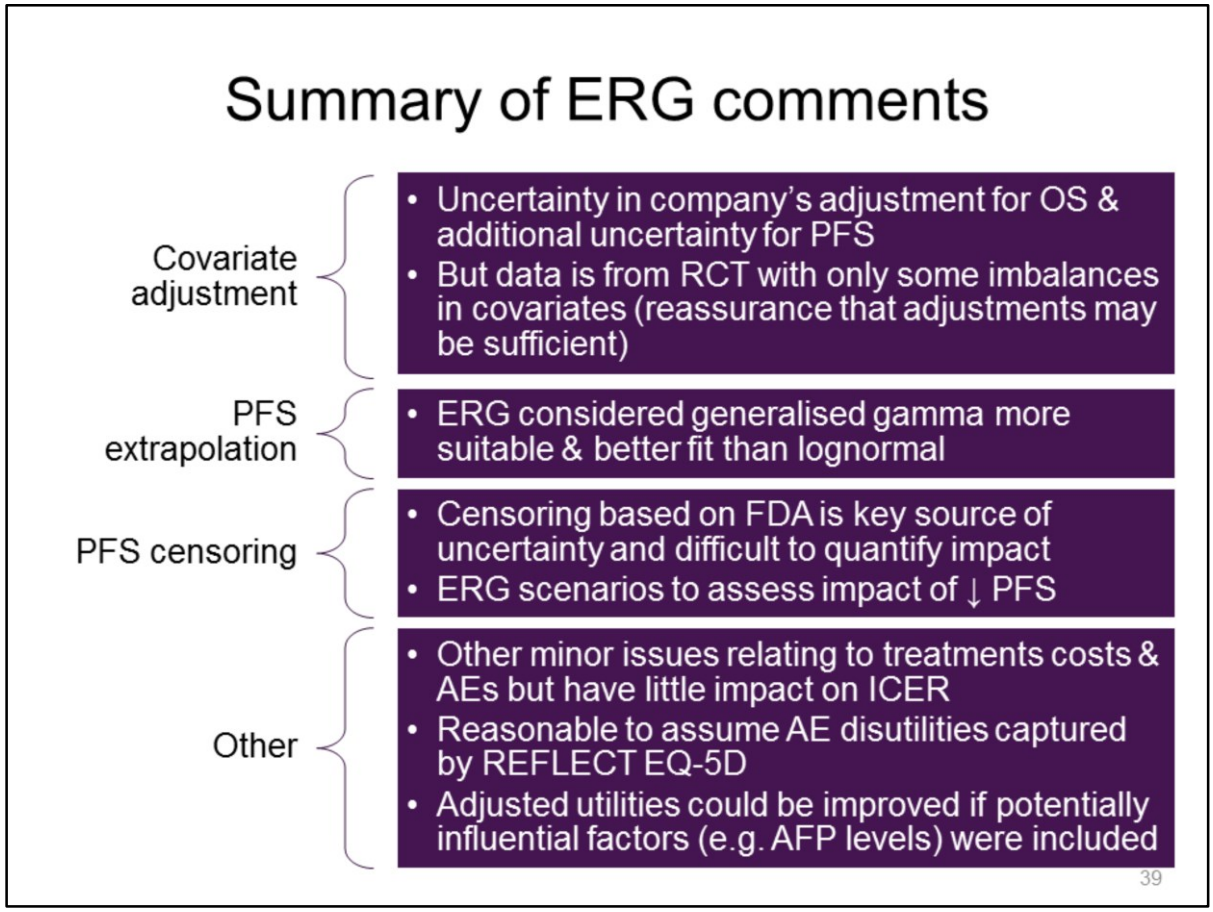
Source: Tables 43 and 45 in ERG report

37

See section B3.7.1 and B.3.8.1 (table 43 on page 107 and page 108) of the company submission for more details.




See section B3.8.3 (table 44 on page 112) of the company submission for more details.




See section 1.4 and 1.5 (page 16) in ERG report for more details.

# Other issues in ERG report
## Post-progression treatment

- Imbalance in post-progression treatment in REFLECT trial:
  - 43.1% in the lenvatinib arm and 51.1% in sorafenib arm had post-treatment anti-cancer therapy
  - likely to lead to longer OS in sorafenib group compared with lenvatinib
  - imbalances more extreme in Western subgroup so difference in treatment benefit may have been underestimated further
- Company use crude adjustment (post-progression treatment as binary variable) but adjusted analysis preferable to company's base case
  - company base case adjusts for imbalances in some of the covariates, but it does not adjust for post-progression anti-cancer interventions.
  - Using adjusted OS avoids the need to offset potential benefits incurred by these treatments by applying the respective costs (important given that no subsequent treatments are recommended beyond second-line in the UK)

**ERG base case:** use OS adjusted for post-progression anti-cancer treatment and remove post progression treatment costs

40

See section 5.4.5.3.2 (page 93) in ERG report for more details.

CONFIDENTIAL

# Other issues in ERG report

## Censoring and Western subgroup

| Company | ERG comments |
|---|---|
| Base case uses full population (not Western subgroup) | • At clarification, requested additional analyses on Western subgroup but these were based on reduced sample size (from 954 to 314) so less robust<br>• Results showed OS in lenvatinib group worse than sorafenib group (HR 1.08, 0.82 to 1.42) but result was not significant. After adjusting for imbalances in post-progression anti-cancer therapies results similar to full population<br>• Analysis based on full population appropriate |
| In base case, data censored if no disease progression at treatment discontinuation | • Using this approach PFS modelling potentially overestimated (company's scenario including all progressed disease events: HR ↑ from 0.66 to [redacted]<br>• Both analyses unreliable as PH assumption not met<br>• ERG scenarios: diminish the treatment effect for PFS to assess the potential impact of this censoring |

**ERG base case:** use full population (no change from company base case)
**ERG scenario:** lower PFS treatment effect

41

See sections 5.4.5.3.4 and 5.4.5.3.3 (pages 93 and 94) in ERG report for more details.

CONFIDENTIAL

# ERG scenario analyses
## Lenvatinib vs. sorafenib (PAS for lenvatinib only)

| Scenario | Incremental | | | ICER |
|---|---|---|---|---|
| | Cost | LY | QALY | |
| Company's corrected base case | [redacted] | 0.23 | [redacted] | Dominant |
| 1. Adjusted OS for post-progression anti-cancer intervention & no post-progression treatment cost | [redacted] | 0.30 | [redacted] | Dominant |
| 2. PFS: gamma (prevention of curves crossing) | [redacted] | 0.23 | [redacted] | Dominant |
| 3. OS: gamma (prevention of curves crossing) | [redacted] | 0.18 | [redacted] | Dominant |
| 4. Mortality adjustment for newly progressed patients | [redacted] | 0.23 | [redacted] | Dominant |
| 5. All post-progression intervention costs (including procedures) | [redacted] | 0.23 | [redacted] | Dominant |
| 6. Scenario 4 and 5 | [redacted] | 0.23 | [redacted] | Dominant |
| 7. Scenario 1 with Western subgroup with Western subgroup AEs | [redacted] | 0.11 | [redacted] | Dominant |
| 8. Full costs of drugs (no dose reductions) | [redacted] | 0.23 | [redacted] | Dominant |
| 9. Removal of AEs | [redacted] | 0.23 | [redacted] | Dominant |

42

See sections 6.2 (page 122 and table 47 on page 124) in ERG report for more details.

CONFIDENTIAL

# ERG's base case results

## Lenvatinib vs. sorafenib (PAS for lenvatinib only)

| | Total | | Incremental | | ICER | |
|---|---|---|---|---|---|---|
| **Drug** | Costs (£) | QALYs | Costs | QALYs | vs. base case | All changes |
| **1. Company's corrected base case** | | | | | | |
| **SOR** | £65,574 | 1.03 | - | - | - | - |
| **LEN** | [redacted] | [redacted] | [redacted] | [redacted] | Dominant | Dominant |
| **2. Post-progression adjustment to OS and no post progression therapy costs** | | | | | | |
| **SOR** | £60,243 | 0.95 | - | - | - | - |
| **LEN** | [redacted] | [redacted] | [redacted] | [redacted] | Dominant | Dominant |
| **3. Gamma distribution for PFS (with prevention of curves crossing)** | | | | | | |
| **SOR** | £56,237 | 0.96 | - | - | - | - |
| **LEN** | [redacted] | [redacted] | [redacted] | [redacted] | Dominant | Dominant |
| **ERG base case (1 to 3)** | | | | | | **Dominant** |

- ERG unable to incorporate uncertainty around its preferred survival models in PSA (covariance matrices not provided).
- No PSA conducted as ERG considered results without this uncertainty to be unreliable and potentially misleading.

*Note: ICERs with comparator CAA discount to be presented in part 2 slides*

Source: Table 48 in ERG report

43

See sections 6.3 (page 125 and table 48) in ERG report for more details.

CONFIDENTIAL

# ERG's base case scenario analyses

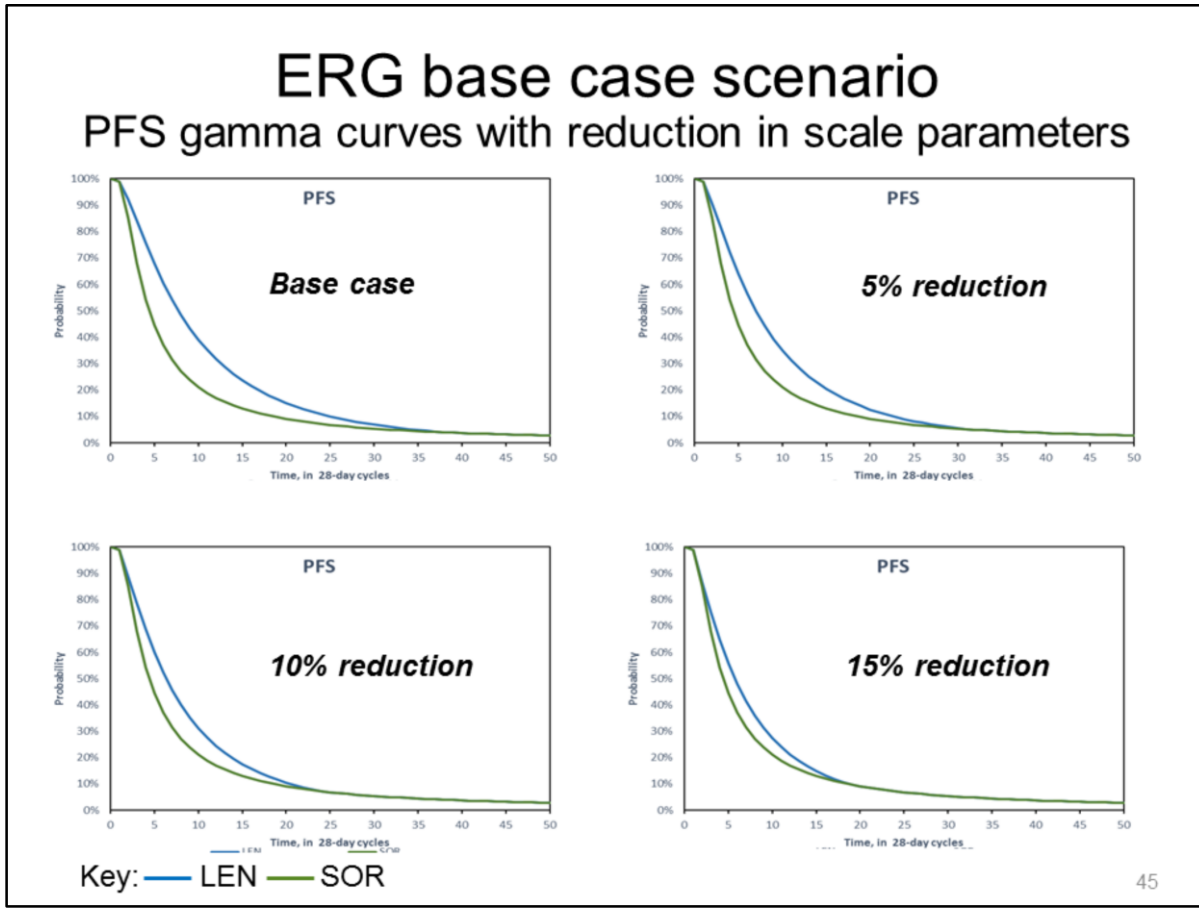
## Lenvatinib vs. sorafenib (PAS for lenvatinib only)

| Scenario | Incremental | | | ICER vs. base case |
|---|---|---|---|---|
| | Cost | LY | QALY | |
| ERG base case | [redacted] | 0.30 | [redacted] | Dominant |
| 1. Reduce scale of gamma PFS function by 5% | [redacted] | 0.30 | [redacted] | £2,085 |
| 2. Reduce scale of gamma PFS function by 10% | [redacted] | 0.30 | [redacted] | £8,490 |
| 3. Reduce scale of gamma PFS function by 15% | [redacted] | 0.30 | [redacted] | £14,024 |
| 4. Gamma distribution for OS (with prevention of curves crossing) | [redacted] | 0.24 | [redacted] | Dominant |
| Abbreviations: ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALYs, quality-adjusted life years. | | | | |

Source: Table 49 in ERG report

44

See sections 6.4 (page 126 and table 49 on page 127) in ERG report for more details.




See figure 21 (page 127) of ERG report for more details

# End of life
## Short life expectancy

| Company | ERG |
|---|---|
| Yes, median survival for patients with advanced HCC is <1 year; 4 to 8 months if untreated and 6 to 11 months with sorafenib. | ESMO-ESDO guidelines give median OS based on natural history as 4 to 8 months, and 6 to 11 months with sorafenib, for people with BCLC Stage C HCC.<br>From the company submission:<br>• Sorafenib median OS in REFLECT (primary analysis) = 12.3 months (upper quartile 25.4 months).<br>• Sorafenib mean OS = 17.5 months (company base case), 16.2 months (ERG's base case). |

Source: Table 50 in ERG report

46

See sections B.2.13.3 (table 24 on page 67) in company submission and section 7 (table 50 on page 129) in ERG report for more details.

CONFIDENTIAL

# End of life
## Extension to life

| Criterion | Company | ERG |
|---|---|---|
| **Met criterion?** | Yes, mean OS benefit for lenvatinib of 3.1 months compared with sorafenib (page 67 in company submission) | Incremental mean OS benefit from the company's economic model:<br>• Company base case at clarification = 2.7 months<br>• ERG base case = 3.6 months. |
| **LY gain** | 0.22 (company submission)<br>0.23 (after clarification stage) | 0.33 (ERG preferred base case) |
| **Treatment effect** | **Lenvatinib**: Median 13.6 (12.1 to 14.9 months)<br>**Sorafenib**: Median 12.3 (10.4 to 13.9 months)<br>Unadjusted HR 0.92 (0.79 to 1.06)<br>Adjusted for post-progression treatment HR [redacted] | |

Source: Table 50 in ERG report

47

See sections B.2.13.3 (table 24 on page 67) in company submission and section 7 (table 50 on page 129) in ERG report for more details.

# Innovation and equality

**Innovation**

- Company consider lenvatinib innovative because it is a multiple receptor tyrosine kinase (RTK) inhibitor with a novel binding mode that inhibits the kinase activities of VEGF and FGF receptors in addition to other proangiogenic and oncogenic pathway-related RTKs involved in tumour proliferation.
- Sorafenib currently the only available systemic treatment option for patients with advanced HCC and there is a clear unmet need for new treatments which delay progression and improve survival without negatively impacting patients' quality of life.

**Equality**

- Company state no known equality issues
  - Clinical expert statement: "HCC is much more common in men and this was reflect in the trial population in which 85% were male."
  - Hep C trust: "Liver cancer disproportionately affects men (though not women) living in deprived areas in England. It also disproportionately affects Asian and Black people."

48

See section B.2.12 (page 60) in company submission for more details.

# Authors

- **Abi Senthinathan**
  Technical Lead
- **Alex Filby**
  Technical Adviser
- with input from the Lead Team (Gail Coster, David Chandler and Natalie Hallas)

49

# NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

## Single technology appraisal

# Lenvatinib for untreated advanced or unresectable hepatocellular carcinoma [ID1089]

## Document B

## Company evidence submission

**February 2018**

| File name | Version | Contains confidential information | Date |
|---|---|---|---|
| | | Yes/no | |

# Contents



# Tables and figures



# Abbreviations

| | |
|---|---|
| AASLD | American Association for the Study of Liver Diseases |
| ACS | American Cancer Society |
| AE | Adverse event |
| AFP | Alpha fetoprotein |
| AIC | Akaike Information Criteria |
| ASCO | American Society of Clinical Oncology |
| BCLC | Barcelona Clinic Liver Cancer |
| BIC | Bayesian Information Criteria |
| BID | Twice daily |
| BOR | Best overall response |
| CADTH | Canadian Agency for Drugs and Technologies in Health |
| CEA | Cost-effectiveness analysis |
| CEAC | Cost-effectiveness acceptability curve |
| CEP | Cost-effectiveness plane |
| CHMP | Committee for Medicinal Products for Human Use |
| CI | Confidence interval |
| CR | Complete response |
| CRF | Case report form |
| CTCAE | Common Terminology Criteria for Adverse Events |
| DCR | Disease control rate |
| DSU | Decision Support Unit |
| EASL | European Association for the Study of the Liver |
| ECOG PS | Eastern Cooperative Oncology Group Performance Status |
| EMA | European Medicines Agency |
| EORTC | European Organisation for Research and Treatment of Cancer |
| ESDO | European Society of Digestive Oncology |
| ESMO | European Society for Medical Oncology |
| EQ-5D-(3L) | EuroQol-5 Dimensions (3 level version) |
| FACT-G | Functional Assessment of Cancer Therapy-General |
| FAS | Full Analysis Set |
| FDA | Food and Drug Administration |
| FGF (R) | Fibroblast growth factor (receptor) |
| HBV | Hepatitis B virus |
| HCC | Hepatocellular carcinoma |
| HCHS | Hospital and Community Health Services |
| HCV | Hepatitis C virus |
| HR | Hazard ratio |
| HRQoL | Health-related quality of life |
| ICER | Incremental cost-effectiveness ratio |
| IIR | Independent imaging review |

| | |
|---|---|
| INAHTA | International Network of Agencies for Health Technology Assessment |
| INR | International normalised ratio |
| ISPOR | International Society for Pharmacoeconomics and Outcomes Research |
| IVRS | Interactive Voice Response System |
| MedDRA | Medical Dictionary for Regulatory Activities |
| mRECIST | Modified Response Evaluation Criteria in Solid Tumours |
| MTA | Multiple technology appraisal |
| MTD | Maximum tolerated dose |
| NCCN | National Comprehensive Cancer Network |
| NICE | National Institute for Health and Care Excellence |
| ORR | Objective response rate |
| OS | Overall survival |
| PAS | Patient Access Scheme |
| PBS | Pharmaceutical Benefits Scheme (Australia) |
| PD | Progressive disease |
| PDGFR | Platelet-derived growth factor receptor |
| PFS | Progression-free survival |
| PH | Proportional hazard |
| PHARMAC | Pharmaceutical Management Agency (New Zealand) |
| PPS | Per Protocol Analysis Set |
| PR | Partial response |
| PSM | Partitioned survival model |
| PSSRU | Personal and Social Services Research Unit |
| PY | Patient year |
| Q | Quartile |
| QALY | Quality-adjusted life year |
| QD | Once daily |
| QLQ | Quality of Life Questionnaire |
| QoL | Quality of life |
| RD | Risk difference |
| RECIST | Response Evaluation Criteria in Solid Tumours |
| RR | Relative risk |
| RTK | Receptor tyrosine kinase |
| SAE | Serious adverse event |
| SAP | Statistical Analysis Plan |
| SD | Standard deviation |
| SMC | Scottish Medicines Consortium |
| STA | Single technology appraisal |
| StD | Stable disease |
| TACE | Transcatheter arterial chemoembolisation |
| TEAE | Treatment-emergent adverse event |
| TTD | Time to discontinuation |

| | |
|---|---|
| TTP | Time to progression |
| TSD | Technical Support Document |
| VAS | Visual analogue scale |
| VEGF (R) | Vascular endothelial growth factor (receptor) |

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

## *B.1.1 Decision problem*

The submission covers the technology's full marketing authorisation for this indication.

**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 | Adults with unresectable hepatocellular carcinoma who have not previously received systemic treatment | Adults with untreated advanced or unresectable HCC and Child-Pugh Class A liver function. This is consistent with the pivotal lenvatinib RCT (REFLECT) population which included some patients with BCLC stage B (those who were considered ineligible for TACE), with most patients having BCLC stage C disease. | This population is consistent with that of REFLECT, and the lenvatinib licensed indication.*<br>The population is also consistent with the NICE recommended population for the use of sorafenib in HCC, the SHARP study which was the evidence base for this recommendation, and with UK clinical practice as specified in the sorafenib NICE recommendation (i.e. predominantly BCLC stage C (advanced) disease, predominantly good liver function (Child-Pugh Class A) and good ECOG performance status (0–2)). |
| Intervention | Lenvatinib | As per scope | NA |
| Comparator(s) | • Sorafenib<br>• Best supportive care | • Sorafenib | BSC is not considered to be a relevant comparator due to the small numbers of patients in the population defined above (<5% according to a UK clinical expert [see section B.3.3.3 for details]) that would receive this as an alternative to sorafenib. Feedback from UK clinical experts is that in the defined patient population (adults with untreated advanced or unresectable HCC and Child-Pugh Class A liver function), almost all patients would be eligible to receive systemic therapy. |
| Outcomes | • Overall survival<br>• Progression-free survival<br>• Time to progression<br>• Response rates<br>• Adverse effects of treatment<br>• Health-related quality of life | As per scope | |

* The current draft SPC is presented in Appendix C. The population addressed in this submission and detailed in this table is based on anticipated changes to the licensed indication requested by the EMA rapporteur that have not yet been incorporated into the draft SPC.

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; BSC, best supportive care; ECOG, Eastern Cooperative Oncology Group; HCC, hepatocellular carcinoma; NA, not applicable; NICE, National Institute for Health and Care Excellence; RCT, randomised controlled trial; TACE, transcatheter arterial chemoembolisation.



## *B.1.2 Description of the technology being appraised*

**Table 2: Technology being appraised**

| | |
|---|---|
| **UK approved name and brand name** | UK approved name: Lenvatinib<br>Brand name: Lenvima® * |
| **Mechanism of action** | Lenvatinib is a RTK inhibitor that inhibits the activity of the VEGF receptors VEGFR1 (FLT1), VEGFR2 (KDR), and VEGFR3 (FLT4). Lenvatinib also inhibits other RTKs that have been implicated in angiogenesis, tumour growth, and cancer progression, including the FGF receptors FGFR1, 2, 3, and 4, PDGFRα, KIT, and RET. |
| **Marketing authorisation/CE mark status** | A regulatory submission was made to the EMA on 24th July 2017.<br>CHMP positive opinion is expected in April 2018 with marketing authorisation expected to be granted by the European Commission by June 2018. |
| **Indications and any restriction(s) as described in the summary of product characteristics (SmPC)** | The current draft indication† for lenvatinib is for the treatment of adult patients who have received no prior systemic therapy for HCC. |
| **Method of administration and dosage** | The recommended daily dose of lenvatinib is 8 mg (two 4 mg capsules) given orally QD for patients with a body weight of <60 kg and 12 mg (three 4 mg capsules) orally QD for patients with a body weight of ≥60 kg. The daily dose is to be modified, as needed, according to the dose/toxicity management plan. |
| **Additional tests or investigations** | None. |
| **List price and average cost of a course of treatment** | £1,437.00 per pack of 30 x 4 mg capsules.<br>The average cost of a course of treatment (including PAS) is [redacted] |
| **Patient access scheme (if applicable)** | There is a simple PAS agreed with the Department of Health and the PAS price is incorporated in the submission. |

[redacted] †The current draft SPC is presented in Appendix C. The population addressed in this submission and detailed in this table is based on anticipated changes to the licensed indication requested by the EMA rapporteur that have not yet been incorporated into the draft SPC.

Abbreviations: CHMP, Committee for Medicinal Products for Human Use; EMA, European Medicines Agency; FGF(R), fibroblast growth factor (receptor); HCC, hepatocellular carcinoma; PAS, Patient Access Scheme; PDGFRα, platelet derived growth factor receptor alpha; QD, once daily; RTK, receptor tyrosine kinase; VEGF(R), vascular endothelial growth factor (receptor).

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

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer in England, with 2,456 cases diagnosed in 2015 (1). It is characterised by a well-developed vascular network coupled with high levels of vascular endothelial growth factor (VEGF) and fibroblast growth factor (FGF) expression which play a role in tumour growth and angiogenesis (2-4). HCC prevalence increases with age, and is more common in men than women, accounting for 55% of male and 28% of female primary liver cancer diagnoses (5). Since the early 1990s, liver cancer incidence rates in the UK have increased by 142% and the rate is predicted to increase by a further 38% between 2014 and 2035 from 9.5 to 15 cases per 100,000 people (6).

Approximately 70–90% of HCC cases occur in the presence of liver cirrhosis (7, 8). A UK clinical expert indicated that this is reflective of UK clinical practice. In the UK, the most common causes of liver cirrhosis are non-alcoholic fatty liver disease (41% of cases), alcoholic liver disease (30% of cases) and hepatitis C (HCV)/ hepatitis B (HBV) infection (12-16% of cases) (9-11). Recent evidence suggests that the increasing incidence of HCC in the UK is being driven by alcohol consumption and obesity (12) and the incidence of non-alcoholic fatty liver disease is predicted to continue rising (13).

The prognosis of HCC is poor; 5-year survival rates for liver cancer in England are less than 15% (14). The main predictors of survival in patients with HCC are liver function, tumour burden (size and number of HCC nodules, vascular invasion), performance status and serum alpha fetoprotein (AFP) level (15-17). Approximately 50% of HCC patients have advanced disease (18). In addition, up to 70% of patients who initially undergo potentially curative procedures will have recurrent, advanced-stage disease within 5 years (19).

Treatment options for HCC depend on disease stage and liver function. The Barcelona Clinic Liver Cancer (BCLC) classification divides HCC into five stages (0 [Very early], A [Early], B [Intermediate], C [Advanced] and D [Terminal]) based on the following prognostic variables:

- number and size of tumours
- performance status
- liver function as measured by the Child-Pugh score (class A = well compensated disease, class B = significant functional compromise, class C = decompensated disease).

The BCLC classification system also recommends treatment options for each stage (Figure 1). This classification system has been adopted by the European Association for the Study of the Liver-European Organisation for Research and Treatment of Cancer (EASL–EORTC) and European Society for Medical Oncology-European Society of Digestive Oncology (ESMO-ESDO) guidelines for HCC (20, 21) and is widely used in the UK to predict prognosis and determine treatment (22). An update of the BCLC staging system was published in January 2018 (23), however this update has not yet been incorporated into the above mentioned treatment guidelines. One key difference is the substitution of 'sorafenib' for 'systemic therapy' as the treatment option for patients with advanced stage disease (BCLC stage C) (23), in order to reflect the availability of new treatment options in this area.

Treatment options for patients with advanced or unresectable disease are very limited. Sorafenib is currently the only NICE-recommended targeted systemic therapy for the treatment of HCC; it is recommended for advanced HCC with Child-Pugh grade A liver impairment (24). In the UK, it is expected that lenvatinib will be used in the same population as sorafenib (see Table 1).

For the small minority of patients who are not suitable for systemic treatments (<5% according to UK clinical expert opinion [see Section B.3.3.3]), the only option currently available is BSC, which comprises symptomatic treatment (such as sleep disturbances, depression, fatigue, malnutrition, anorexia, pain and psychological issues), management of comorbidities such as cirrhosis and palliative care (25, 26). Due to the complex interaction between the tumour and the underlying disease, patients typically require management by a multidisciplinary clinical team in order to maximise outcomes (26).

Median survival for patients with advanced HCC is still less than one year; 4–8 months if untreated and 6–11 months with sorafenib treatment (21).

Since the approval of sorafenib in 2006, there have been no new first-line therapies approved for advanced HCC. Several other investigational therapies have failed to meet the endpoints of non-inferiority or superiority for OS compared with sorafenib (sunitinib (27), brivanib (28), linifanib (29), sorafenib + erlotinib (30), sorafenib + doxorubicin (31)). There is therefore a need for additional treatments with improved clinical benefits for this patient population.

Lenvatinib is a multiple RTK inhibitor, which is currently under assessment for regulatory approval by the European Medicines Agency (EMA). The current draft licensed indication for lenvatinib is for the treatment of adult patients who have received no prior systemic therapy for HCC. It is expected that lenvatinib will be used as an alternative to sorafenib in patients with advanced or unresectable HCC and Child-Pugh class A liver function who have not previously received systemic treatment (see dotted green box, Figure 1). As indicated in Figure 1, the majority of patients expected to be treated with lenvatinib would have BCLC Stage C disease. In addition, as highlighted in Table 1, in UK clinical practice, this will also include some patients with BCLC Stage B disease.

**Figure 1: BCLC staging and treatment recommendations for HCC with proposed lenvatinib place in therapy**

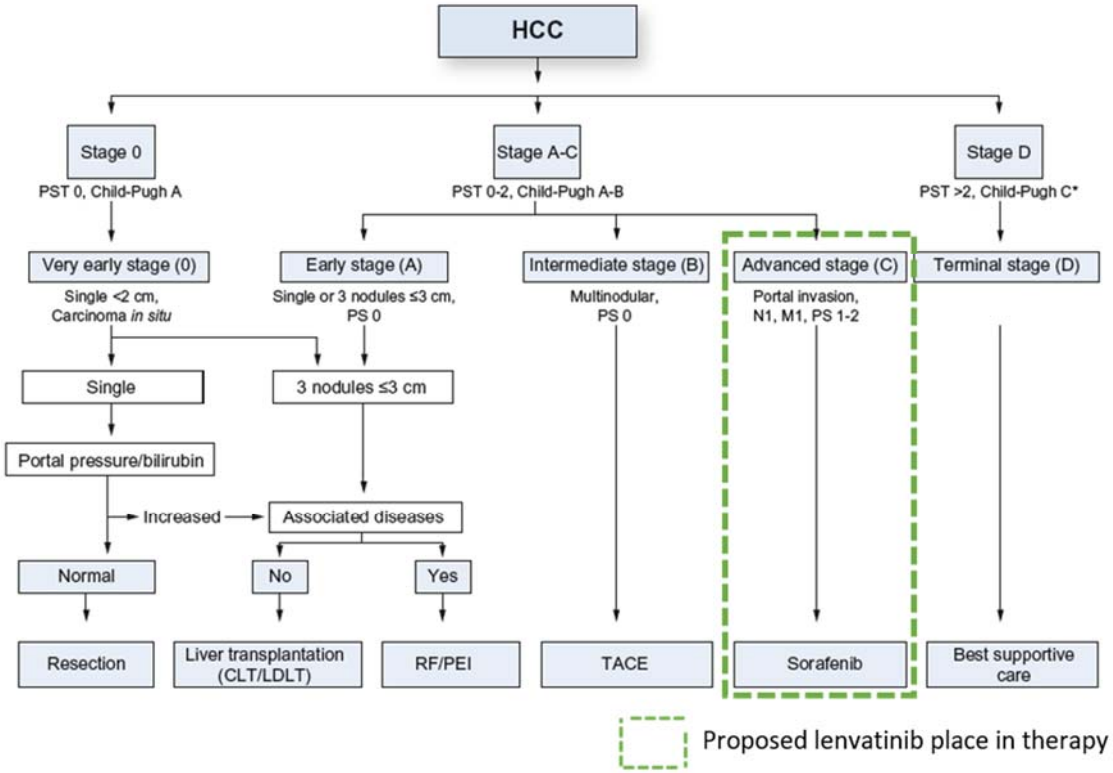



Diagram adapted from EASL-EORTC, 2012 (20).

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; CLT, cadaveric liver transplantation; EASL-EORTC; European Association for the Study of the Liver- European Organisation for Research and Treatment of Cancer; HCC, hepatocellular carcinoma; LDLT, living donor liver transplantation; PEI, percutaneous ethanol injection; RF, radiofrequency ablation; TACE, transcatheter arterial chemoembolisation.

## *B.1.4 Equality considerations*

Use of lenvatinib is not expected to raise any equality issues.

## B.2 Clinical effectiveness

### *B.2.1 Identification and selection of relevant studies*

See Appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to the technology being appraised.

### *B.2.2 List of relevant clinical effectiveness evidence*

A summary of the clinical effectiveness evidence for lenvatinib is provided in Table 3 and Table 4.

**Table 3: Clinical effectiveness evidence – pivotal trial**

| | | | | | |
|---|---|---|---|---|---|
| **Study** | **E7080-G000-304 (REFLECT)**<br>**Data sources†: CSR (32), SAP (33), Data on file (34, 35)**<br>**Kudo et al, 2018 (36)** | | | | |
| **Study design** | Multicentre, randomised, open-label, Phase III trial | | | | |
| **Population** | Patients with unresectable BCLC stage B (those who were ineligible for TACE) or C HCC and Child Pugh Class A liver disease | | | | |
| **Intervention(s)** | Lenvatinib QD | | | | |
| **Comparator(s)** | Sorafenib BID | | | | |
| **Indicate if trial supports application for marketing authorisation** | Yes | ✓ | **Indicate if trial used in the economic model** | Yes | ✓ |
| | No | | | No | |
| **Rationale for use/non-use in the model** | Used in CE model: Pivotal trial comparing the efficacy and safety of lenvatinib against sorafenib, which is the only available intervention for patients with advanced HCC and a key comparator in the NICE scope. | | | | |
| **Reported outcomes specified in the decision problem** | **Overall survival**<br>**Progression-free survival**<br>Time to progression<br>Objective response rate<br>**Adverse effects**<br>**Quality of life** (EORTC QLQ-30, EORTC HCC-18**, EQ-5D-3L)** | | | | |
| **All other reported outcomes** | Pharmacokinetics, pharmacodynamics, exploratory endpoints including disease control rate, clinical benefit rate, tumour biomarkers | | | | |

†Kudo et al, 2018 was published after the date on which the SLR searches were conducted. Five abstracts reporting on the REFLECT trial were identified in the clinical SLR (reported in Appendix D), however these were not used as data sources due to the level of detail contained within the documents listed here.

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; BID, twice daily; CSR, clinical study report; EMA, European Medicines Agency; EORTC, European Organisation for Research and Treatment of Cancer; EQ-5D-3L, EuroQol-5 Dimensions (3 level version); HCC, hepatocellular carcinoma; QD, once daily; QLQ, Quality of Life Questionnaire; SAP, statistical analysis plan; TACE, transcatheter arterial chemoembolisation.



**Table 4: Clinical effectiveness evidence – supporting Phase I/II dose finding study**

| | | | | | |
|---|---|---|---|---|---|
| **Study** | **E7080-J081-202**<br>**Data sources: CSR (37), Tamai et al, 2017, (38), Ikeda et al, 2017 (39)** | | | | |
| **Study design** | Phase I/II dose finding study | | | | |
| **Population** | Patients with advanced HCC | | | | |
| **Intervention(s)** | Lenvatinib QD | | | | |
| **Comparator(s)** | None | | | | |
| **Indicate if trial supports application for marketing authorisation** | Yes | ✓ | **Indicate if trial used in the economic model** | Yes | |
| | No | | | No | ✓ |
| **Rationale for use/non-use in the model** | Not used in CE model: Dose finding study which informed the Phase III clinical trial design. | | | | |
| **Reported outcomes specified in the decision problem** | Overall survival, progression-free survival, objective response rate, adverse effects | | | | |
| **All other reported outcomes** | Time to progression | | | | |

Abbreviations: CSR, clinical study report; HCC, hepatocellular carcinoma; QD, once daily.

Study E7080-J081-202 was not used to populate the economic model, but a brief summary of the methods and results is included in this section. The results of this study support the rationale for the dosing strategy of lenvatinib used in the Phase III study. This study was not included in the economic model because it was a Phase I/II dose finding study, provided no comparative evidence and did not provide any additional evidence to that already available from the pivotal Phase III study.

### B.2.2.1 Supporting dose finding study – E7080-J081-202

Study E7080-J081-202 (37-39) was a multicentre, open-label Phase I/II study consisting of two phases – a dose escalation and dose determination phase, and an expansion phase. The study was conducted in Japan (Phase I and II) and South Korea (Phase II).

***Dose Escalation Component (Phase I):*** designed primarily to determine the maximum tolerated dose (MTD) of lenvatinib based on dose-limiting toxicity. Lenvatinib doses of 8, 12, and 16 mg once daily (QD) were tested in 20 patients. The

12-mg QD dose was determined to be the MTD in 9 patients with a Child-Pugh score of 5 or 6 (Class A) and was the recommended dose for the Phase 2 portion of the study.

***Expansion phase (Phase II):*** Designed primarily to evaluate the efficacy (time to progression [TTP]) by Modified Response Evaluation Criteria in Solid Tumours (mRECIST) and safety of lenvatinib at the recommended dose of 12 mg QD in 46 patients with advanced HCC and Child-Pugh score 5–6. Secondary efficacy endpoints included objective response rate (ORR), disease control rate (DCR), overall survival (OS), progression-free survival (PFS) and adverse events (AEs).

Median PFS and TTP determined by independent imaging review using mRECIST, were both 7.4 months (95% confidence interval [CI]: 5.5, 9.4). The frequency of study drug dose reduction due to any toxicity in Cycle 1 was 32.6%. In total, 80% of patients whose body weight was <60 kg had a dose reduction in Cycle 1.

A population pharmacokinetic (PK) analysis showed that both clearance and volume increased with increasing body weight. Consequently, lenvatinib AUC increased as body weight decreased in patients with HCC. Patients with a low body weight (<60 kg) had a higher lenvatinib AUC, which appears to have led to the high rate of dose reductions and discontinuations in these patients.

Population PK and population PK/pharmacodynamic analyses of lenvatinib safety and efficacy for subjects in Study E7080-J081-202 were conducted to identify the optimal lenvatinib starting dose(s) for further clinical development for HCC. No relationship between the efficacy endpoints and lenvatinib AUC based on starting dose was detected. Therefore, the use of a lower starting dose (8 mg) in patients weighing less than 60 mg was not expected to impact efficacy.

Based on these results, a 2-tier dosing strategy based on body weight was proposed to achieve comparable lenvatinib exposures and to manage toxicity: 12 mg QD for patients weighing ≥60 kg, and 8 mg QD for patients weighing <60 kg. These were the lenvatinib starting doses in the Phase III study E7080-G000-304 (REFLECT).

## *B.2.3 Summary of methodology of the relevant clinical effectiveness evidence*

### B.2.3.1 Summary of trial methodology – REFLECT (Study E7080-G000-304)

#### *B.2.3.1.1 Study objectives*

The primary objective of the REFLECT study was to compare OS in patients treated with lenvatinib versus sorafenib as a first-line treatment for unresectable stage B or C (according to BCLC staging system) HCC and Child Pugh Class A liver disease.

Secondary objectives included a comparison of the following in patients treated with lenvatinib vs sorafenib:

- PFS, TTP, and ORR using mRECIST
- Health-related quality of life (HRQoL) using the generic cancer HRQoL instrument EORTC Quality of Life Questionnaire (QLQ)-C30, the HCC-specific module (EORTC QLQ-HCC18) and the generic EuroQol-5 Dimensions 3 level version (EQ-5D-3L) instrument
- Safety and tolerability

#### *B.2.3.1.2 Location*

Patients were enrolled at 183 sites across 20 countries (China, Hong Kong, Japan, Republic of Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand, Belgium, Canada, France, Germany, Italy, Israel, Poland, Russia, Spain, USA, UK [6 specialist cancer centres which enrolled 20 patients). All tumour assessments were performed at the study site by appropriately qualified personnel.

#### *B.2.3.1.3 Trial design*

REFLECT was an international, multicentre, randomised, open-label Phase III study. An overview of the study design is presented in Figure 2. The study was conducted in 3 phases:

***1. Pre-randomisation Phase:*** Included a screening period to establish eligibility and a baseline period to confirm eligibility and establish baseline characteristics, prior to randomisation.

***2. Randomisation Phase:*** Consisted of two sub-phases – the Treatment Period and the Follow-Up Period. The Randomisation Phase began when the first patient was randomised and ended at the data cut-off date for the primary analysis (13th November 2016). Patients received study drug in 28-day treatment cycles until completion of the Off-Treatment visit, which occurred within 30 days after the final administration of study drug.

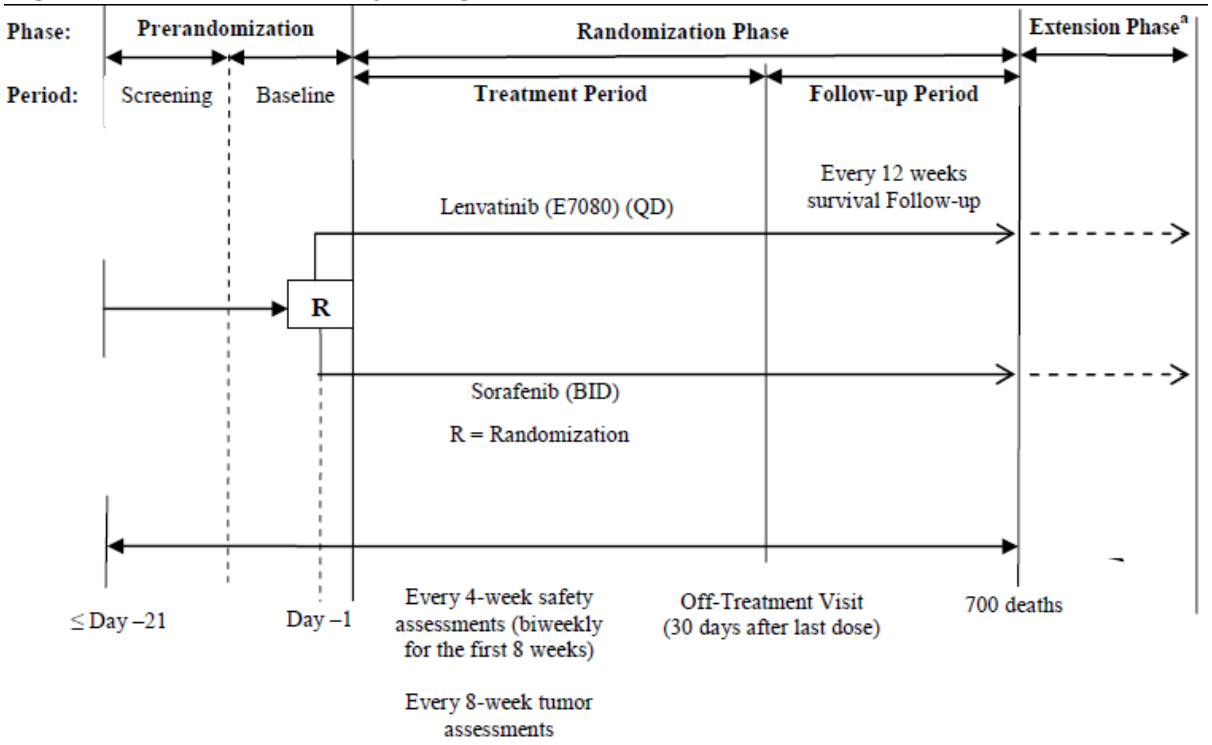
Patients discontinued treatment at the time of objectively documented disease progression, development of unacceptable toxicity, patient request, or withdrawal of consent. Following treatment discontinuation, patients entered the Follow-Up Period which continued as long as the patient was alive, unless they withdrew consent.

Patients who discontinued study treatment for any reason other than disease progression were followed in the Randomisation Phase until disease progression or start of another anti-cancer therapy, then entered the Extension Phase for survival follow-up. Patients who were still receiving study drug or were in the Follow-Up Period at the time of the data cut-off entered the Extension Phase.

***3. Extension Phase:*** Patients who were still on study drug at the end of the Randomisation Phase could continue on the same study treatment in the Extension Phase. Patients who had disease progression during the Randomisation Phase and all patients who discontinued study treatment entered the Follow-up Period. Patients were followed for survival and all anti-cancer treatments were recorded until the time of death. The Follow-up Period continued as long as study patients were alive or until discontinuation of survival follow-up by the sponsor.

**Figure 2: Overview of study design - REFLECT**




[a] Extension Phase also included a Treatment Period and Follow-up Period. All patients still on treatment at the end of the Randomisation Phase entered the Extension Phase and continued on the same study treatment they received in the Randomisation Phase.

Abbreviations: QD once daily; BID, twice daily.

#### *B.2.3.1.4 Eligibility criteria*

Details of key inclusion and exclusion criteria for REFLECT are presented in Table 5.



**Table 5: Eligibility criteria - REFLECT**

| Inclusion | Exclusion |
|---|---|
| • Adult patients (≥18 years of age) with histologically or cytologically confirmed diagnosis of unresectable HCC or a clinically confirmed diagnosis of HCC according to the American Association for the Study of Liver Diseases criteria, including cirrhosis of any aetiology, or with chronic hepatitis B or C infection<br>• ≥1 measurable target hepatic or non-hepatic lesion according to mRECIST, and adequate liver, bone marrow, blood coagulation, renal, and pancreatic function<br>  • Measurable disease was defined as follows:<br>  • Hepatic lesion (i) the lesion could be accurately measured in ≥1 dimension as ≥1.0 cm (viable tumour for typical; and longest diameter for atypical), and (ii) lesion was suitable for repeat measurement<br>  • Non-hepatic lesion (i) lymph node lesion that measured in ≥1 dimension as ≥1.5 cm in the short axis, except for porta hepatis lymph node that measured ≥2.0 cm in the short axis (ii) non-nodal lesion that measured ≥1.0 cm in the longest diameter<br>• Lesions previously treated with radiotherapy or locoregional therapy must have shown radiographic evidence of disease progression to be deemed a target lesion<br>• Patients categorised to stage B (not applicable for TACE) or stage C based on the BCLC staging system<br>• Adequately controlled blood BP with up to 3 antihypertensive agents, defined as BP ≤150/90 mm Hg at Screening and no change in antihypertensive agents within 1 week prior to Cycle 1/Day 1<br>• Child-Pugh class A<br>• ECOG PS 0 or 1<br>• Survival expectation of 12 weeks or longer after starting study drug | • Patients who had any previous systemic anti-cancer therapy or any systemic investigational anti-cancer agents, including lenvatinib, for advanced/unresectable HCC<br>• Imaging findings that indicated HCC with ≥50% liver occupation, clear invasion into the bile duct or main portal branch invasion (Vp4), or receipt of any blood-enhancing treatment within 28 days of randomisation<br>• Patients who had not recovered (recovery defined as severity Grade <2 per CTCAE) from toxicities as a result of prior anti-cancer therapy, except alopecia and infertility.<br>• Significant CV impairment: history of congestive heart failure greater than NYHA Class II, unstable angina, myocardial infarction or stroke within 6 months of the first dose of study drug, or cardiac arrhythmia requiring medical treatment at Screening<br>• Prolongation of QTc interval to >480 ms<br>• Bleeding or thrombotic disorders or use of anticoagulants requiring therapeutic INR monitoring. Treatment with low molecular weight heparin and factor X inhibitors which did not require INR monitoring was permitted. Antiplatelet agents were prohibited throughout the study.<br>• GI malabsorption or any other condition that might have affected the absorption of lenvatinib in the opinion of the investigator<br>• GI bleeding event or active haemoptysis (bright red blood of at least 0.5 teaspoon) within 28 days prior to randomisation<br>• Gastric or oesophageal varices that required active interventional treatment within 28 days prior to randomisation. Prophylaxis with pharmacologic therapy (e.g. non-selective beta-blocker) was permitted.<br>• Patients whose only target lesion was in bone<br>• Meningeal carcinomatosis<br>• History of or current brain or subdural metastases |

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; BP, blood pressure; CTCAE, Common Terminology Criteria for Adverse Events; CV, cardiovascular; ECOG PS, Eastern Cooperative Oncology Group Performance Status; GI, gastrointestinal; HCC, hepatocellular carcinoma; INR, International Normalised Ratio; mRECIST, modified Response Evaluation Criteria In Solid Tumours; NYHA, New York Heart Association; TACE, transcatheter arterial chemoembolisation.



#### *B.2.3.1.5 Method of blinding and randomisation*

REFLECT was an open-label study. An open-label design was chosen in the interests of patient safety (for further details, please see Section B.2.13.2). Patients were randomised to receive either lenvatinib (12 mg [if baseline body weight ≥60 kg] or 8 mg [if baseline body weight <60 kg] QD) or sorafenib (400 mg twice daily [BID]) in a 1:1 ratio based on a computer-generated randomisation scheme that was reviewed and approved by an independent statistician. Allocation of randomisation numbers was performed using an Interactive Voice Response System (IVRS) based on the following stratification factors:

- Region: Region 1 (Asia-Pacific); Region 2 (Western regions, such as EU, North America, other)
- Macroscopic portal vein invasion or extrahepatic spread or both: Yes; No
- Eastern Cooperative Oncology Group Performance Status (ECOG PS): PS = 0; PS = 1
- Body weight: <60 kg; ≥60 kg

In order not to compromise the robustness of the data, it was not possible to include AFP as an additional stratification factor (see Section B.2.3.1.11).

#### *B.2.3.1.6 Trial drugs*

***Intervention – lenvatinib (N=478):*** Lenvatinib was provided as 4 mg capsules. Each patient's dose was based on body weight at Baseline; patients weighing ≥60 kg were treated with 12 mg QD administered as three 4 mg capsules at the same time and patients weighing <60 kg were treated with 8 mg QD administered as two 4 mg capsules at the same time. Lenvatinib was orally self-administered by the patients at the same time each day (consistently either with or without food) in continuous 28-day cycles. Dose interruption, dose reduction, or treatment discontinuation were allowed for patients who experienced lenvatinib-related toxicity. Full details are provided in Appendix L, Section L1. Once the dose was reduced, it could not be increased.

***Comparator – sorafenib (N=476):*** Sorafenib (NEXAVAR®) was provided as 200 mg tablets. Patients were treated with 400 mg of sorafenib BID administered as two

200 mg tablets. Sorafenib was orally self-administered by the patients, at the same time each day (consistently either with or without food) in continuous 28-day cycles. Patients who experienced sorafenib-related toxicity had dose adjustments in accordance with the prescribing information in each country or region. Once the dose was reduced, it could not be increased.

#### *B.2.3.1.7 Permitted and disallowed concomitant medications*

***Permitted concomitant medications:*** Drugs used to treat complications or AEs, or drugs used to ameliorate symptoms (including blood products, blood transfusions, fluid transfusions, antibiotics, steroids, antidiarrheal drugs, and tranquilisers, etc). These drugs and treatments were concomitantly used, based on the judgment of the investigator or sub-investigator.

***Disallowed concomitant medications:*** The following were disallowed from the time of patient enrolment until discontinuation of study drug.

1. Surgery or radiotherapy for the treatment of HCC; palliative radiotherapy of up to two painful pre-existing nontarget bone metastases was permitted without being considered progressive disease.

2. Systemic therapy, hepatic intra-arterial chemotherapy, immunological therapy (e.g. interferon, interferon-type drugs, etc), hormonal therapy, or local therapy of any kind (e.g. percutaneous ethanol injection, radiofrequency ablation, transcatheter arterial chemoembolisation [TACE]) for the treatment of HCC. Patients who were receiving antiviral therapy for hepatitis B virus were allowed to continue to receive this therapy at the discretion of the investigator.

3. Other investigational drugs.

4. Antiplatelet agents and anticoagulants that required international normalised ratio (INR) monitoring, such as warfarin (treatments that did not require INR monitoring, such as low molecular weight heparin and certain factor X inhibitors were permitted).

#### *B.2.3.1.8 Primary outcome (including scoring methods and timings of assessments)*

The primary outcome was OS, which was measured from the date of randomisation to the date of death from any cause. Patients who were lost to follow-up were censored at the last date that they were known to be alive and patients who remained alive were censored at the time of data cut-off. Patients were followed every 12 weeks for survival information (date and cause, and post-treatment cancer therapies).

#### *B.2.3.1.9 Other outcomes used in the economic model/specified in the scope*

The following pre-specified secondary efficacy outcomes were specified in the scope and/or utilised in the economic model:

***PFS:*** the time from the date of randomisation to the date of first documentation of disease progression, or death, whichever occurred first. This endpoint was based on tumour response evaluations as determined by the investigator according to mRECIST for HCC for hepatic lesions (40). Tumour assessments were performed every 8 weeks. A retrospective, blinded independent imaging review (IIR) of PFS was also conducted using both mRECIST and RECIST 1.1 criteria (41).

***TTP:*** the time from the date of randomisation to the date of first documentation of disease progression. Time-to-progression censoring rules were defined in the statistical analysis plan (SAP).

***ORR:*** the proportion of patients with a best overall response (BOR) of complete response (CR) or partial response (PR). This endpoint was based on tumour response evaluations as determined by the investigator according to mRECIST for HCC for hepatic lesions (40). Tumour assessments were performed every 8 weeks.

***Quality of life:*** Assessments of HRQoL scores were performed using the generic cancer HRQoL instrument (EORTC QLQ-C30), the HCC-specific module (EORTC QLQ-HCC18), and the generic HRQoL instrument, EQ 5D-3L. Patients were asked to complete each questionnaire at the Baseline visit, Day 1 of each treatment cycle and at the Off-Treatment visit, which occurred within 30 days after the final dose of

study treatment. Differences between the treatments arms (lenvatinib versus sorafenib) were investigated cross-sectionally and longitudinally.

***Adverse events:***

- An AE was defined as any untoward medical occurrence in a patient administered an investigational product, regardless of causal relationship to study drug. All AEs were graded on a 5-point severity scale according to Common Terminology Criteria for Adverse Events (CTCAE) and the assessment of the relationship to study drug was recorded.
- Serious adverse events (SAEs) were defined as any AE that led to death, was life threatening, required hospitalisation (or prolonged hospitalisation), resulted in persistent or significant disability/incapacity or led to a congenital abnormality/birth defect.
- AEs were recorded for 30 days after the last dose of study treatment.

#### *B.2.3.1.10 Adjustments for covariates*

The SAP specified randomisation stratification factors (region [Asia-Pacific; Western], macroscopic portal vein invasion or extrahepatic spread or both [yes; no], ECOG PS [PS=0; PS=1] and body weight [<60 kg; ≥60 kg]) for use in stratified log-rank tests and stratified Cox proportional hazard models for the primary endpoint (OS) and key secondary endpoints (PFS, TTP and ORR). The SAP allowed for other baseline factors to be used in the model as covariates in supportive analyses for the endpoints.

#### *B.2.3.1.11 Pre-planned subgroups*

For the primary efficacy endpoint, OS, and for the key secondary endpoints, PFS, TTP, and ORR, summaries and test statistics (i.e. HR and 95% CI) were provided for subgroups based on randomisation stratification factors, age group (≤65, ≥65 to <75 years and ≥75 years), sex and aetiology (HBV, HCV and alcohol). Additional subgroup analyses that were deemed necessary were also conducted, including AFP at baseline (<200 ng/mL, ≥200 ng/mL) and post-treatment anti-cancer therapy (yes, no). Due to the existing four stratification factors listed in previous section, in

order to not compromise the robustness of the data, it was not possible to include AFP as an additional stratification factor.

### B.2.3.2 Baseline characteristics and demographics – REFLECT (Study E7080-G000-304)

Patient characteristics by treatment group at baseline are summarised in Table 6 and disease characteristics by treatment group at baseline are reported in Table 7. A greater proportion of patients enrolled in the study were from the Asia Pacific region (68.7%) than from the Western region (32.9%). In total, 68.7% of patients (68% of lenvatinib-treated and 69.3% of sorafenib-treated patients) weighed ≥60 kg.

Demographics and baseline characteristics were generally well balanced between the lenvatinib and sorafenib treatment arms, however there were some notable differences in disease characteristics:

- The proportion of patients with AFP levels ≥200 ng/mL, a marker of poor HCC prognosis (17), was higher in the lenvatinib arm (46.4%) than the sorafenib arm (39.3%).
- The proportion of patients with an aetiology of HCV was lower in the lenvatinib arm (19%) than in the sorafenib arm (26.5%). Evidence suggests that patients with HCV aetiology may derive more clinical benefit from sorafenib than patients with other aetiologies, particularly HBV (42).

**Table 6: Demographic and baseline characteristics – FAS**

| | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg† (N = 151) | 12 mg† (N = 327) | Total (N = 478) | (N = 476) |
| Age (years) | | | | |
| Mean (SD) | 63.1 (12.30) | 60.4 (11.32) | 61.3 (11.69) | 61.2 (12.01) |
| Median | 65.0 | 62.0 | 63.0 | 62.0 |
| Q1, Q3 | 56.0, 72.0 | 53.0, 68.0 | 54.0, 70.0 | 54.0, 70.0 |
| Min, max | 20, 86 | 24, 88 | 20, 88 | 22, 88 |
| Age group (years), n (%) | | | | |
| <65 | 69 (45.7) | 201 (61.5) | 270 (56.5) | 283 (59.5) |
| ≥65 to <75 | 56 (37.1) | 94 (28.7) | 150 (31.4) | 126 (26.5) |
| ≥75 | 26 (17.2) | 32 (9.8) | 58 (12.1) | 67 (14.1) |

| | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg† (N = 151) | 12 mg† (N = 327) | Total (N = 478) | (N = 476) |
| Sex, n (%) | | | | |
| Male | 106 (70.2) | 299 (91.4) | 405 (84.7) | 401 (84.2) |
| Female | 45 (29.8) | 28 (8.6) | 73 (15.3) | 75 (15.8) |
| Region, n (%) | | | | |
| Western‡ | 21 (13.9) | 136 (41.6) | 157 (32.8) | 157 (33.0) |
| Asia-Pacific‡ | 130 (86.1) | 191 (58.4) | 321 (67.2) | 319 (67.0) |
| Race, n (%) | | | | |
| White | 17 (11.3) | 118 (36.1) | 135 (28.2) | 141 (29.6) |
| Black/African American | 0 (0.0) | 7 (2.1) | 7 (1.5) | 6 (1.3) |
| Asian | 134 (88.7) | 200 (61.2) | 334 (69.9) | 326 (68.5) |
| American Indian/Alaskan native | 0 (0.0) | 1 (0.3) | 1 (0.2) | 0 (0.0) |
| Hawaiian/Pacific Islander | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (0.2) |
| Other | 0 (0.0) | 1 (0.3) | 1 (0.2) | 2 (0.4) |
| Weight (kg) | | | | |
| Mean (SD) | 52.7 (4.90) | 75.9 (14.40) | 68.6 (16.32) | 68.1 (13.90) |
| Median | 53.0 | 72.0 | 66.2 | 67.0 |
| Q1, Q3 | 50.0, 56.5 | 65.3, 82.0 | 57.0, 76.2 | 57.6, 77.0 |
| Min, max | 39, 60 | 60, 142 | 39, 142 | 39, 123 |
| Body weight group | | | | |
| <60 kg | 151 (100.0) | 2 (0.6) | 153 (32.0) | 146 (30.7) |
| ≥60 kg | 0 (0.0) | 325 (99.4) | 325 (68.0) | 330 (69.3) |
| ECOG PS, n (%) | | | | |
| 0 | 93 (61.6) | 211 (64.5) | 304 (63.6) | 301 (63.2) |
| 1 | 58 (38.4) | 116 (35.5) | 174 (36.4) | 175 (36.8) |
| NYHA classification, n (%) | | | | |
| I | 4 (2.6) | 33 (10.1) | 37 (7.7) | 44 (9.2) |
| II | 1 (0.7) | 7 (2.1) | 8 (1.7) | 6 (1.3) |
| Not applicable | 145 (96.0) | 287 (87.8) | 432 (90.4) | 426 (89.5) |
| Missing | 1 (0.7) | 0 (0.0) | 1 (0.2) | 0 (0.0) |
| Child-Pugh score, n (%) | | | | |
| 5 | 111 (73.5) | 257 (78.6) | 368 (77.0) | 357 (75.0) |
| 6 | 40 (26.5) | 67 (20.5) | 107 (22.4) | 114 (23.9) |
| 7 | 0 (0.0) | 3 (0.9) | 3 (0.6) | 4 (0.8) |
| 8 | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (0.2) |

| | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg† (N = 151) | 12 mg† (N = 327) | Total (N = 478) | (N = 476) |
| Macroscopic portal vein invasion, n (%) | | | | |
| Yes | 38 (25.2) | 71 (21.7) | 109 (22.8) | 90 (18.9) |
| No | 113 (74.8) | 256 (78.3) | 369 (77.2) | 386 (81.1) |
| Extrahepatic spread, n (%) | | | | |
| Yes | 91 (60.3) | 200 (61.2) | 291 (60.9) | 295 (62.0) |
| No | 60 (39.7) | 127 (38.8) | 187 (39.1) | 181 (38.0) |
| Macroscopic portal vein invasion, extrahepatic spread, or both, n (%) | | | | |
| Yes | 105 (69.5) | 224 (68.5) | 329 (68.8) | 336 (70.6) |
| No | 46 (30.5) | 103 (31.5) | 149 (31.2) | 140 (29.4) |
| Underlying cirrhosis, n (%) | | | | |
| Yes§ | 75 (49.7) | 168 (51.4) | 243 (50.8) | 231 (48.5) |
| No | 76 (50.3) | 159 (48.6) | 235 (49.2) | 245 (51.5) |

† 8mg and 12 mg were the lenvatinib starting doses based on patients' body weight (<60 kg, ≥60 kg) at Baseline; ‡ Western region consists of North America and Europe including Russia and Israel; Asia-Pacific region consists of China, Hong Kong, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand; § The proportion of patients with underlying cirrhosis at baseline (49.7%) was likely underestimated as this information was collected on the CRF under medical history, and the presence or absence of cirrhosis was verified only when needed to confirm the clinical diagnosis of HCC.

Abbreviations: CRF, case report form; ECOG PS, Eastern Cooperative Oncology Group Performance Status; FAS, Full Analysis Set; HCC, hepatocellular carcinoma; kg, kilograms; NYHA, New York Heart Association; Q, quartile; SD, standard deviation.



**Table 7: Disease history and characteristics – FAS**

| | Lenvatinib N=478 n (%) | Sorafenib N=476 n (%) |
|---|---|---|
| Time since first diagnosis (months) | | |
| Mean (SD) | 21.1 (30.17) | 23.3 (34.66) |
| Median | 8.2 | 9.0 |
| Q1, Q3 | 1.6, 27.3 | 2.0, 27.2 |
| Min, max | 0, 180 | 0, 250 |
| Age at first diagnosis (years) | | |
| Mean (SD) | 59.6 (11.57) | 59.3 (11.54) |
| Median | 61.0 | 60.0 |
| Q1, Q3 | 52.0, 68.0 | 52.0, 67.0 |
| Min, max | 15, 87 | 20, 85 |
| BCLC stage, n (%) | | |
| B: Intermediate stage | 104 (21.8) | 92 (19.3) |
| C: Advanced stage | 374 (78.2) | 384 (80.7) |
| Involved disease sites[†], n (%) | | |
| Liver | 441 (92.3) | 430 (90.3) |
| Lung | 163 (34.1) | 144 (30.3) |
| Lymph nodes | 127 (26.6) | 141 (29.6) |
| Bone | 51 (10.7) | 43 (9.0) |
| Other | 82 (17.2) | 97 (20.4) |
| Number of involved disease sites per patient, n (%) | | |
| 1 | 207 (43.3) | 207 (43.5) |
| 2 | 167 (34.9) | 183 (38.4) |
| ≥3 | 103 (21.5) | 86 (18.1) |
| Factor of carcinogenesis[‡], n (%) | | |
| Hepatitis B | 251 (52.5) | 228 (47.9) |
| Hepatitis C | 91 (19.0) | 126 (26.5) |
| Alcohol | 36 (7.5) | 21 (4.4) |
| Other | 38 (7.9) | 32 (6.7) |
| Unknown | 62 (13.0) | 69 (14.5) |
| Baseline alpha-fetoprotein level (ng/mL) | | |
| Mean (SD) | 17507.5 (105137.39) | 16678.5 (94789.46) |
| Median | 133.1 | 71.2 |
| Q1, Q3 | 8.0, 3730.6 | 5.2, 1081.8 |
| Min, max | 0, 1567470 | 0, 1446396 |

| | Lenvatinib N=478 n (%) | Sorafenib N=476 n (%) |
|---|---|---|
| Baseline alpha-fetoprotein group, n (%) | | |
| <200 ng/mL | 255 (53.3) | 286 (60.1) |
| ≥ 200 ng/mL | 222 (46.4) | 187 (39.3) |
| Missing | 1 (0.2) | 3 (0.6) |
| Ammonia level (μg/dL) | | |
| Mean (SD) | 38.2 (29.98) | 36.7 (32.90) |
| Median | 31.8 | 30.0 |
| Q1, Q3 | 22.0, 45.0 | 21.0, 42.3 |
| Min, max | 4, 246 | 4, 473 |
| Concomitant systemic antiviral therapy for Hepatitis B or Hepatitis C, n (%) | 163 (34.1) | 149 (31.3) |

† Patients may be counted in more than 1 disease site; ‡ Based on the combined data from HCC diagnosis and medical history. Patients may be counted in more than 1 factor.

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; FAS, Full Analysis Set; HCC, hepatocellular carcinoma; Q, quartile; SD, standard deviation.

Prior anti-cancer procedures (including radiotherapy) were performed in 68.4% and 72.3% of patients in the lenvatinib and sorafenib arms, respectively (**Error! Not a valid bookmark self-reference.**). Most patients in both treatment arms had one (30.1%, lenvatinib; 31.3% sorafenib) or two (24.7%, lenvatinib; 25.0%, sorafenib) previous procedures. The most common previous anti-cancer procedures were transarterial chemoembolisation (51.5% of patients in both treatment arms); hepatectomy (25.9% of the lenvatinib arm and 30.3% of the sorafenib arm), and radiofrequency ablation (18.8% of the lenvatinib arm and 23.1% of the sorafenib arm). Overall, 10.3% of lenvatinib-treated and 12.6% of sorafenib-treated patients received previous radiotherapy.

**Table 8: Previous anti-cancer procedures and radiotherapy – FAS**

| | Lenvatinib N=478 n (%) | Sorafenib N=476 n (%) |
|---|---|---|
| Patients with any previous anti-cancer procedure, n (%) | 327 (68.4) | 344 (72.3) |
| Number of previous procedures, n (%) | | |
| 1 | 144 (30.1) | 149 (31.3) |
| 2 | 118 (24.7) | 119 (25.0) |
| 3 | 44 (9.2) | 56 (11.8) |
| 4 | 14 (2.9) | 15 (3.2) |
| ≥5 | 7 (1.5) | 5 (1.1) |
| Previous procedure name* | | |
| Hepatic intra-arterial chemotherapy | 22 (4.6) | 35 (7.4) |
| Transarterial chemoembolisation | 246 (51.5) | 245 (51.5) |
| Radiofrequency ablation | 90 (18.8) | 110 (23.1) |
| Cryoablation | 1 (0.2) | 1 (0.2) |
| Percutaneous ethanol injection | 15 (3.1) | 19 (4.0) |
| Hepatectomy | 124 (25.9) | 144 (30.3) |
| Other† | 85 (17.8) | 72 (15.1) |
| Patients with any previous radiotherapy treatment, n (%) | 49 (10.3) | 60 (12.6) |

*A patient may be counted in multiple categories; †Previous anti-cancer procedures that were reported on the case report form by the investigator in the "other" category were varied, but were primarily hepatectomy, microwave therapy, biopsies, or pulmonary resections.

Abbreviations: FAS, Full Analysis Set.

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

### B.2.4.1 Populations analysed

The following populations were considered in the study:

- **Full Analysis Set (FAS; also called the Intent to Treat Analysis Set)** included all patients who were randomised. This was the primary analysis set for all efficacy evaluations.
- **Per Protocol Analysis Set (PPS)** included patients who were randomised and received at least 1 dose of the assigned study drug and had no major protocol deviations. This was the secondary analysis set for all efficacy evaluations.

- **Safety Analysis Set** included patients who received at least one dose of the study treatment. This was the analysis set for all safety evaluations.

### B.2.4.2 Statistical information

A summary of the statistical methods employed in the REFLECT trial is presented in Table 9.

**Table 9: Summary of statistical analyses**

| **REFLECT (E7080-G000-304)** | |
|---|---|
| Hypothesis objective | OS was compared between lenvatinib and sorafenib testing the null hypothesis:<br>H0: log HR (lenvatinib: sorafenib) ≥ (1 – δ) log HR (placebo/sorafenib) against the alternative:<br>H1: log HR (lenvatinib: sorafenib) < (1 – δ) log HR (placebo/sorafenib)<br>where log HR (lenvatinib: sorafenib) was the HR for lenvatinib versus sorafenib on a log-scale and log HR (placebo/sorafenib) was the estimated HR for the placebo versus sorafenib on a log-scale by a meta-analysis. δ was the proportion of sorafenib treatment effect retained by lenvatinib. |
| Statistical analysis | ***OS:*** A non-inferiority test of OS between lenvatinib and sorafenib was performed using a 2-sided 95% CI of HR (lenvatinib: sorafenib). The HR and the corresponding 2-sided 95% CI were estimated using a Cox proportional hazard model with treatment group as a factor and stratified by the randomisation stratification factors.<br>Superiority hypotheses were tested for OS using a stratified log-rank test with the randomisation stratification factors. No multiplicity adjustments were needed for testing of the non-inferiority and superiority of OS due to the closed testing principle. Two interim analyses were performed. The early stopping rule for both interim analyses was set for futility based on non-inferiority only. Since the study was not stopped at the first or second interim analysis, non-inferiority for OS was tested first at the final analysis with a non-inferiority margin of 1.08, which indicated that lenvatinib preserved at least 60% (corresponding to δ = 0.60) of the sorafenib treatment effect versus placebo as observed in the sorafenib SHARP and Asia-Pacific trials (43, 44). Non-inferiority was declared if the upper limit of the 2-sided 95% CI for HR was <1.08 at the final analysis. If non-inferiority was declared for OS, then superiority (corresponding to δ = 1) was to be tested for OS. Superiority would be declared if the 2-sided p-value was <0.05 using the stratified log-rank test at the final analysis. The primary analysis was performed when the target number of 700 events (deaths) occurred.<br>Treatment group comparison for the secondary efficacy endpoints was performed if non-inferiority for the primary efficacy endpoint, OS, was declared. The fixed sequence procedure was used to control the overall type I error rate of analyses for the secondary endpoints at α = 0.05 (2-sided).<br>***PFS:*** The difference between lenvatinib and sorafenib was evaluated using a stratified log-rank test with the randomisation stratification factors, tested at an alpha level of 0.05 (2-sided). The corresponding estimate of the HR, calculated from the Cox proportional hazard model with treatment group as a factor and stratified by the randomisation stratification factors, was presented with a 2-sided 95% CI. Median, Q1 and Q3 of PFS, and the cumulative probability of PFS at 6-month intervals were calculated for each treatment group and presented with corresponding 2-sided 95% CIs. Kaplan-Meier estimates of PFS for each |

| REFLECT (E7080-G000-304) | |
|---|---|
| | group were plotted over time.<br>***TTP:*** The difference in TTP between lenvatinib and sorafenib was evaluated using the same procedure as for PFS, with the exception that death was censored.<br>***ORR***: The statistical significance of the difference in ORR between treatment groups was evaluated using the Cochran-Mantel-Haenszel chi-square test with the stratification factors as strata, tested at an alpha level of 0.05 (2-sided). The 2-sided 95% CIs for the odds ratio and the difference in ORR were calculated as well as 2-sided 95% CIs for the rate within treatment group. |
| Sample size, power calculation | Sample size determination was based primarily on the required number of target events to detect the non-inferiority and superiority of lenvatinib to sorafenib in the comparison of OS. The required number of target events was estimated based on the following assumptions:<br>• Exponential distribution was assumed for OS. The estimated median OS of sorafenib was approximately 10 months, and an improvement of 2.5 months was derived from the underlying objective of achieving a HR of 0.8, which would be of marked clinical benefit.<br>• Using a non-inferiority test by the 95% CI lower limit method on the log HR for OS with an assumed true HR of 0.80 and non-inferiority margin of 1.08 (corresponding to 60% retention of sorafenib effect versus placebo), the power of the study to declare non-inferiority was approximately 97%.<br>• The power of the study to declare superiority of lenvatinib to sorafenib was approximately 82% using the superiority test, with an assumed true HR of 0.80. The overall false-positive rate was set at a 2-sided alpha of 0.05<br>Based on these assumptions, the required number of events was estimated to be 666 events (deaths) based on the PPS. Assuming that 1) approximately 5% of patients with major protocol deviations and 2) patients assigned to treatment but who did not actually receive it would be excluded from the PPS, approximately 700 events (deaths) based on the FAS would be required at the time of primary analysis. Two interim analyses for futility (1 at approximately 30%, and a second at approximately 70% of the target number of events) were taken into account for the estimation. It was estimated that approximately 940 patients (470 patients per treatment group), needed to be randomised to observe 700 events. The primary OS analysis was performed when the target number of events was observed. |
| Data management, patient withdrawals | At the time of the primary analysis if a patient was alive or did not have disease progression, they were censored. OS patients lost to follow-up were censored at the last date the patient was known to be alive, and patients who remained alive were censored at the time of data cut-off. For PFS the censoring rules followed FDA guidance (45) and are described in detail in Section 8 of the SAP (33). |

Abbreviations: CI, confidence interval; FAS Full Analysis Set; FDA, Food and Drug Administration; HR, hazard ratio; OS, overall survival; ORR, objective response rate; PFS, progression-free survival; PPS, Per Protocol Set; SAP, Statistical Analysis Plan, TTP, time to progression.

### B.2.4.3 Participant flow in the relevant randomised controlled trials

In total, 1,492 patients were screened for entry. Of these, 954 patients (63.9%) were randomly assigned in a 1:1 ratio to receive either lenvatinib (478 patients) or sorafenib (476 patients). For, further details, please refer to Appendix D, Section D.2.

### *B.2.5 Quality assessment of the relevant clinical effectiveness evidence*

A complete quality assessment for the REFLECT trial is presented in Table 10.

**Table 10: Quality assessment results for parallel group RCTs**

| Trial number (acronym) | REFLECT |
|---|---|
| Was randomisation carried out appropriately? | Yes. Patients were assigned to treatment based on a computer-generated randomisation scheme. Allocation of randomisation numbers was performed using an interactive voice/web response system based on stratification factors. |
| Was the concealment of treatment allocation adequate? | Yes. Patients were assigned to treatment based on a computer-generated randomisation scheme. |
| Were the groups similar at the outset of the study in terms of prognostic factors? | Yes. Stratified randomisation was performed which generates separate schedules for subsets of participants defined by potentially important prognostic factors (region; macroscopic portal vein invasion or extrahepatic spread of both; ECOG; and body weight). Demographic and other baseline characteristics were generally well balanced between treatment arms with the exception of some notable differences in disease characteristics as described in Section B.2.3.2. |
| Were the care providers, participants and outcome assessors blind to treatment allocation? | No. This was an open label trial. This open-label design was chosen in the interests of patient safety (see Section B.2.13.2). OS was the primary outcome and assessment bias was therefore unlikely. For the secondary endpoints of PFS, TTP and ORR, tumour assessments were performed by the investigator and there was therefore a risk of bias. However, PFS results from a post-hoc, blinded, retrospective IIR using both mRECIST and RECIST 1.1 criteria were consistent with the investigator assessments of PFS (see Section B.2.13.1). |
| Were there any unexpected imbalances in drop-outs between groups? | No. The overall drop-outs were generally well-balanced between treatment arms and the primary reasons for treatment discontinuation were also well-balanced between treatment arms (as per the detailed CONSORT diagram presented in Appendix D, Section D.2. |
| Is there any evidence to suggest that the authors measured more outcomes than they reported? | No there are no reported changes to the planned analysis in the CSR and the outcomes listed in the study protocol and CSR are consistent. |
| Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data? | Yes. The primary analysis was based on a full analysis set (intent to treat analysis set) including all patients who were randomised. At the time of the primary analysis if a patient was alive or did not have disease progression, they were censored. OS patients lost to follow-up were censored at the last date the patient was known to be alive, and patients who remained alive will be censored at the time of data cut-off. For PFS the censoring rules followed FDA guidance (45). |

Abbreviations: CSR, Clinical Study Report; ECOG, Eastern Cooperative Oncology Group; FDA, Food and Drug Administration; (m)RECIST, (modified) Response Evaluation Criteria In Solid Tumours; NA, not applicable; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; TTP, time to progression.

## *B.2.6 Clinical effectiveness results of the relevant trials*

### B.2.6.1 REFLECT (Study E7080-G000-304)

#### *B.2.6.1.1 Primary efficacy outcome: overall survival*

At the data cut-off of 13th November 2016, 73.4% of patients in the lenvatinib arm and 73.5% of patients in the sorafenib arm had died. Median OS was 13.6 months for lenvatinib and 12.3 months for sorafenib (hazard ratio [HR]: 0.92; 95% CI: 0.79, 1.06), thus meeting the primary endpoint of non-inferiority of lenvatinib vs sorafenib (Table 11 and Figure 3). Although numerical improvements in median OS were seen in the lenvatinib arm, the result did not meet the criteria for statistical superiority. Median duration of survival follow-up was 27.7 months in the lenvatinib arm and 27.2 months in the sorafenib arm.

**Table 11: Overall survival based on randomisation stratification factors – FAS**

| | Lenvatinib N=478 | Sorafenib N=476 |
|---|---|---|
| Deaths, n (%) | 351 (73.4) | 350 (73.5) |
| Censored patients, n (%) | 127 (26.6) | 126 (26.5) |
| Lost to follow-up | 5 (1.0) | 11 (2.3) |
| Withdrawal of consent | 13 (2.7) | 8 (1.7) |
| Alive | 109 (22.8) | 107 (22.5) |
| Overall survival (months)* | | |
| Median (95% CI) | 13.6 (12.1, 14.9) | 12.3 (10.4, 13.9) |
| Q1 (95% CI) | [redacted] | [redacted] |
| Q3 (95% CI) | [redacted] | [redacted] |
| Overall survival rate, % (95% CI)† at: | | |
| 6 Months | [redacted] | [redacted] |
| 12 Months | [redacted] | [redacted] |
| 18 Months | [redacted] | [redacted] |
| 24 Months | [redacted] | [redacted] |
| Stratified cox model hazard ratio (95% CI)‡,§ | 0.92 (0.79, 1.06) | |
| Duration of survival follow-up (months)*,** | | |
| Median (95% CI) | 27.7 (26.4, 29.4) | 27.2 (25.9, 28.4) |
| Q1 (95% CI) | 23.3 (22.3, 24.5) | 22.6 (20.9, 23.7) |
| Q3 (95% CI) | 32.8 (31.4, 34.2) | 31.3 (29.6, 33.1) |

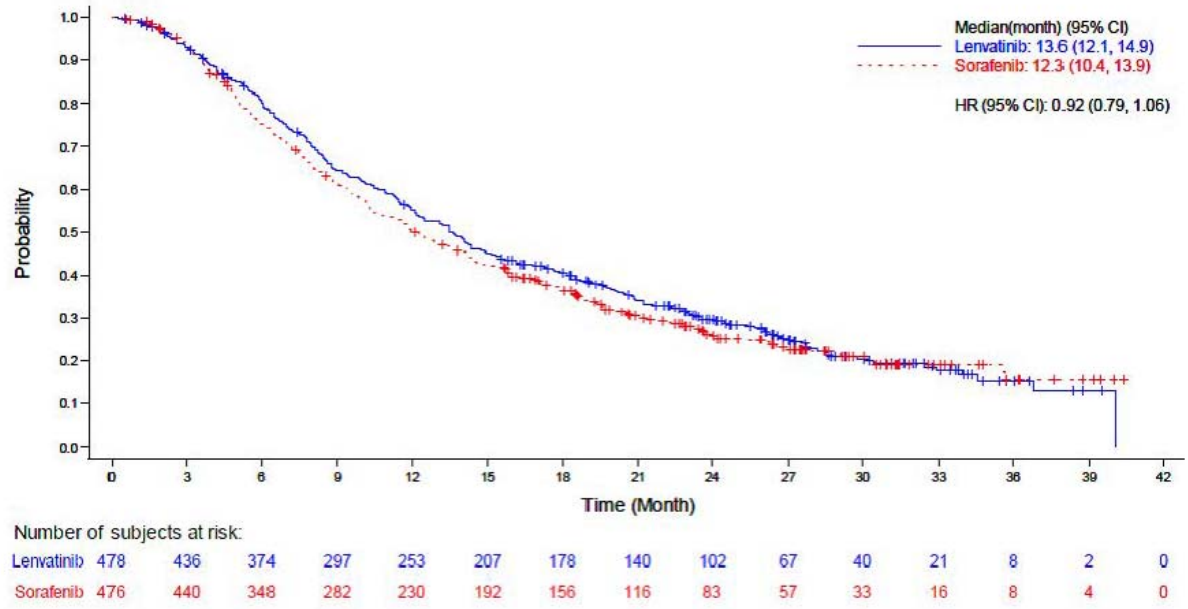
*Quartiles are estimated by the Kaplan-Meier method, and the 95% CIs are estimated with a generalised Brookmeyer and Crowley method; †OS rate and 95% CI were calculated using Kaplan-Meier product-limit method and Greenwood Formula; ‡Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor. Efron method was used for ties; §Stratified by region (Region 1: Asia-Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg); **Duration of survival follow-up is the duration from randomisation to patient's last OS follow-up time, and it has the same numeric value but opposite censoring indicator as compared to OS.

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; FAS, Full Analysis Set; OS, overall survival; Q, quartile.



**Figure 3: Kaplan-Meier curve for overall survival – FAS**




Abbreviations: CI, confidence interval; FAS, Full Analysis Set; HR, hazard ratio.

#### B.2.6.1.1.1 Secondary analysis of primary outcome: per protocol set

The results of the primary efficacy analysis of OS based on the FAS were supported by the results of the analysis based on the PPS, for which median OS was 13.7 months for lenvatinib and 12.3 months for sorafenib (HR: 0.91; 95% CI: 0.78, 1.06).

#### B.2.6.1.1.2 Pre-planned supportive analyses: overall survival adjusted by imbalanced baseline characteristics

As described in Section B.2.3.2, there were baseline imbalances between the lenvatinib and sorafenib treatment arms regarding the proportion of patients with AFP levels ≥200 ng/mL, and in the aetiology of HCC (HBV, HCV, alcohol). Covariate analyses were performed to evaluate baseline factors that may have impacted OS in the overall study population, including AFP and HCC aetiology. Full details of the results are presented in Appendix L, Section L.2. In summary:

- For the FAS, the results adjusted by the individual baseline characteristics generally were consistent with those of the primary OS analysis (HR <1).
- For baseline AFP, lenvatinib was nominally superior to sorafenib (median OS 13.6 months vs 12.3 months, respectively; HR 0.856 with upper limit of the 95% CI <1 [95% CI: 0.736, 0.995], p=0.0342).

- For aetiology of HCC (HBV, HCV, alcohol), the HR for lenvatinib vs sorafenib was 0.855 (95% CI: 0.721, 1.013).

#### B.2.6.1.1.3 Post-hoc analysis: overall survival adjusted by use of post-treatment anti-cancer therapy

There was an imbalance between the treatment arms regarding the proportion of patients who received post-treatment anti-cancer therapy (including procedures and medications) during survival follow-up (Table 12). Fewer patients in the lenvatinib arm (43.1%) had post-treatment anti-cancer therapy than patients in the sorafenib arm (51.1%).

Importantly, OS was shown to be almost twice as long for patients who received anti-cancer therapy compared with those who did not in both the lenvatinib and sorafenib treatment arms (full details are presented in the Subgroup analysis section; Appendix E, Section E.2).

**Table 12: Use of post-treatment anti-cancer therapy by region – FAS**

| | Lenvatinib | | | Sorafenib | | |
|---|---|---|---|---|---|---|
| | **Western (N=157)** | **Asia-Pacific (N=321)** | **Total (N=478)** | **Western (N-157)** | **Asia-Pacific (N=319)** | **Total (N=476)** |
| Received any anti-cancer therapy* during survival follow-up, n (%) | 44 (28.0) | 162 (50.5) | 206 (43.1) | 71 (45.2) | 172 (53.9) | 243 (51.1) |
| Received any anti-cancer medication (not given for a procedure) during survival follow-up, n (%) | 41 (26.1) | 115 (35.8) | 156 (32.6) | 61 (38.9) | 123 (38.6) | 184 (38.7) |
| Underwent any anti-cancer procedure during survival follow-up, n (%) | 11 (7.0) | 111 (34.6) | 122 (25.5) | 18 (11.5) | 112 (35.1) | 130 (27.3) |

*Posttreatment anti-cancer therapy includes both posttreatment anti-cancer procedures and posttreatment anti-cancer medications received during survival follow-up.

Abbreviations: FAS, Full Analysis Set.

Table 13 presents a post-hoc analysis of OS results adjusted by use of post-treatment anti-cancer therapy.

[redacted]

[redacted]

**Table 13: Overall survival adjusted by use of post-treatment anti-cancer treatment, overall and by region – FAS**

| | Stratified Cox Model Hazard Ratio (95% CI)* | |
|---|---|---|
| | **Without Adjustment** | **With Adjustment†** |
| Overall | 0.92 (0.79, 1.06) | [redacted] |
| Region | | |
| Asia-Pacific | 0.86 (0.72, 1.02) | [redacted] |
| Western | 1.08 (0.82, 1.42) | [redacted] |

*Hazard ratio is for lenvatinib:sorafenib, based on a Cox model including treatment group as a factor. The Efron method used for correction of tied events. Stratified by region (Region 1: Asia-Pacific; Region 2: Western), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1), and body weight (<60 kg, ≥60 kg); †Status of post-treatment anti-cancer therapy (yes/no) was used as an additional covariate factor.

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; FAS, Full Analysis Set.

#### *B.2.6.1.2 Secondary efficacy outcome relevant to HE model and/or scope: progression-free survival*

Median PFS was 7.4 months in the lenvatinib arm compared with 3.7 months in the sorafenib arm; a statistically significant and clinically meaningful difference (HR: 0.66; 95% CI: 0.57, 0.77; p<0.00001) (Table 14 and Figure 4). The p-value is for the superiority test of lenvatinib versus sorafenib. The results were consistent in the PPS; median PFS was 7.4 months for lenvatinib versus 3.7 months for sorafenib (HR = 0.66; 95% CI: 0.57, 0.77; p<0.00001).

**Table 14: Progression-free survival assessed by clinical trial investigators based on randomisation stratification factors – FAS**

| | Lenvatinib N=478 | Sorafenib N=476 |
|---|---|---|
| Patients with events, n (%) | [redacted] | [redacted] |
| Progressive disease | [redacted] | [redacted] |
| Death | [redacted] | [redacted] |
| Censored patients, n (%) | [redacted] | [redacted] |
| No post-baseline tumour assessment | [redacted] | [redacted] |
| Death or progression after ≥1 missing assessment | [redacted] | [redacted] |
| New anti-cancer treatment started | [redacted] | [redacted] |
| No progression at the time of data cut-off | [redacted] | [redacted] |
| No progression at the time of treatment discontinuation | [redacted] | [redacted] |
| Progression-free survival (months)* | | |
| Median (95% CI) | 7.4 (6.9, 8.8) | 3.7 (3.6, 4.6) |
| Q1 (95% CI) | [redacted] | [redacted] |
| Q3 (95% CI) | [redacted] | [redacted] |
| Progression-free survival rate (%) (95% CI)† at | | |
| 6 Months | [redacted] | [redacted] |
| 12 Months | [redacted] | [redacted] |
| 18 Months | [redacted] | [redacted] |
| 24 Months | [redacted] | [redacted] |
| Stratified cox model hazard ratio (95% CI)‡,§ | 0.66 (0.57, 0.77) | |
| Stratified log-rank test p-value§ | <0.00001 | |
| Follow-up time for progression-free survival (months)*,†† | | |
| Median (95% CI) | [redacted] | [redacted] |
| Q1 (95% CI) | [redacted] | [redacted] |
| Q3 (95% CI) | [redacted] | [redacted] |

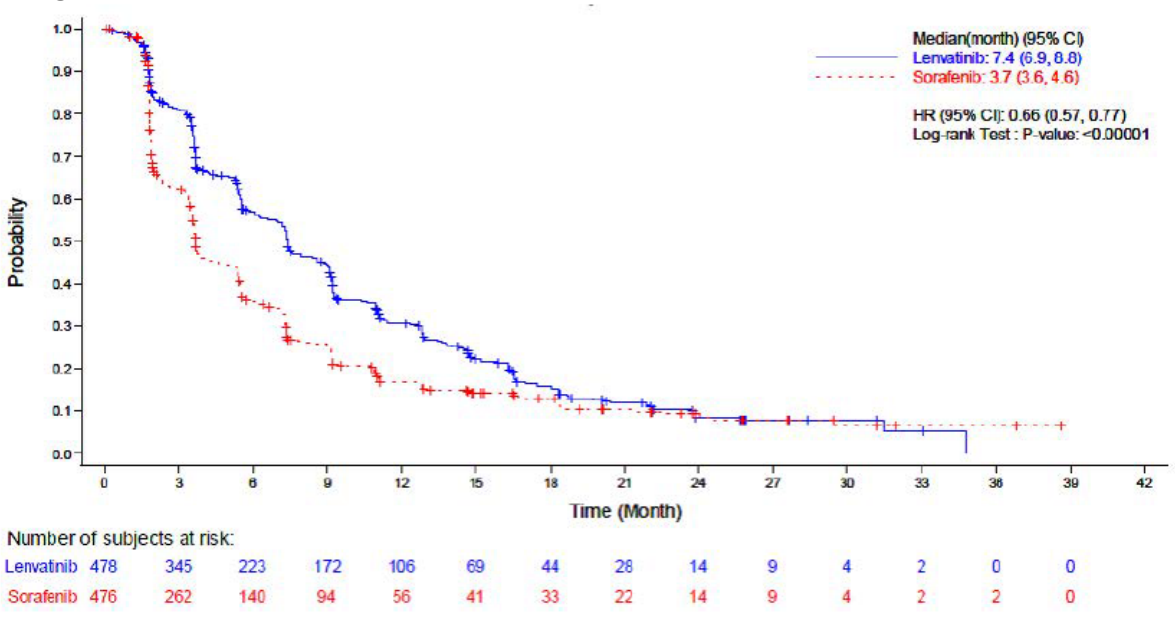
*Quartiles are estimated by Kaplan-Meier method, and the 95% confidence intervals are estimated with a generalised Brookmeyer and Crowley method; †PFS rate and 95% CI were calculated using the Kaplan-Meier product-limit method and the Greenwood Formula; ‡Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor; §Stratified by region (Region 1: Asia-Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg); **Follow-up time for PFS was measured from the date of randomisation to the date of the patient's last PFS follow-up, and it has same numeric value but opposite censoring indicator as compared to PFS.

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; FAS, Full Analysis Set; PFS, progression-free survival; Q, quartile.



**Figure 4: Kaplan-Meier curve for progression-free survival as assessed by the clinical trial investigators – FAS**




Abbreviations: CI, confidence interval; FAS, Full Analysis Set; HR, hazard ratio.

#### B.2.6.1.2.1 Post-hoc analysis: independent review of progression-free survival

A retrospective, blinded independent imaging review (IIR) of PFS was conducted using both mRECIST and RECIST 1.1 criteria. The results of this analysis are not presented within the REFLECT CSR. They were presented at the American Society of Clinical Oncology Gastrointestinal Cancers Symposium in January 2018 (41) and are detailed in the draft SPC (see Appendix C). Assessments were performed using a two-reviewer-plus-adjudicator paradigm.

IIR-assessed PFS per mRECIST and RECIST v1.1 were almost identical (Table 15) and were similar to the PFS findings of the clinical trial investigators (presented in Section B.2.6.1.2). Kaplan-Meier curves for IIR-assessed PFS are presented in Appendix L, Section L.3).

**Table 15: Progression-free survival – independent imaging review**

| | Lenvatinib N=478 | Sorafenib N=476 |
|---|---|---|
| ***mRECIST*** | | |
| Median PFS in months (95% CI)[*] | 7.3 (5.6, 7.5) | 3.6 (3.6, 3.7) |
| Hazard ratio (95% CI)[†,‡] | 0.64 (0.55, 0.75) | |
| P-value[‡,§] | <0.00001 | |
| ***RECIST 1.1*** | | |
| Median PFS in months (95% CI)[*] | 7.3 (5.6, 7.5) | 3.6 (3.6, 3.9) |
| Hazard ratio (95% CI)[†,‡] | 0.65 (0.56, 0.77) | |
| P-value[‡,§] | <0.00001 | |

[*]Quartiles are estimated by the Kaplan-Meier method, and the 95% CIs are estimated with a generalized Brookmeyer and Crowley method; [†] Hazard ratio is for lenvatinib vs. sorafenib, based on a Cox model including treatment group as a factor; [‡] Stratified by region (Region 1: Asia-Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg); [§]p-value is for the superiority test of lenvatinib versus sorafenib.

Abbreviations: CI, confidence interval; PFS, progression-free survival.

### *B.2.6.1.3 Secondary efficacy outcome relevant to HE model and/or scope: time to progression*

Lenvatinib treatment significantly prolonged TTP compared with sorafenib, with a median TTP that was more than twice as long as that of sorafenib: 8.9 months for lenvatinib vs 3.7 months for sorafenib (HR: 0.63; 95% CI: 0.53, 0.73; p<0.00001) (Table 16 and Figure 5).

**Table 16: TTP based on randomisation stratification factors – FAS**

| | Lenvatinib (N=478) n (%) | Sorafenib (N=476) n (%) |
|---|---|---|
| Patients with disease progression, n (%) | [redacted] | [redacted] |
| Censored patients, n (%) | [redacted] | [redacted] |
| No post-baseline tumour assessment* | [redacted] | [redacted] |
| Progression after more than 1 missing assessment | [redacted] | [redacted] |
| New anti-cancer treatment started | [redacted] | [redacted] |
| No progression at the time of data cut-off | [redacted] | [redacted] |
| No progression at the time of treatment discontinuation | [redacted] | [redacted] |
| Time to progression (months)† | | |
| Median (95% CI) | 8.9 (7.4, 9.2) | 3.7 (3.6, 5.4) |

| | Lenvatinib (N=478) n (%) | Sorafenib (N=476) n (%) |
|---|---|---|
| Q1 (95% CI) | [redacted] | [redacted] |
| Q3 (95% CI) | [redacted] | [redacted] |
| Cumulative progression rate (%) (95% CI)‡ at | | |
| 6 Months | [redacted] | [redacted] |
| 12 Months | [redacted] | [redacted] |
| 18 Months | [redacted] | [redacted] |
| 24 Months | [redacted] | [redacted] |
| Stratified Cox model hazard ratio (95% CI)§,** | 0.63 (0.53, 0.73) | |
| Stratified log-rank test p-value § | <0.00001 | |
| Follow-up time for time to progression (months)†, †† | | |
| Median (95% CI) | [redacted] | [redacted] |
| Q1 (95% CI) | [redacted] | [redacted] |
| Q3 (95% CI) | [redacted] | [redacted] |

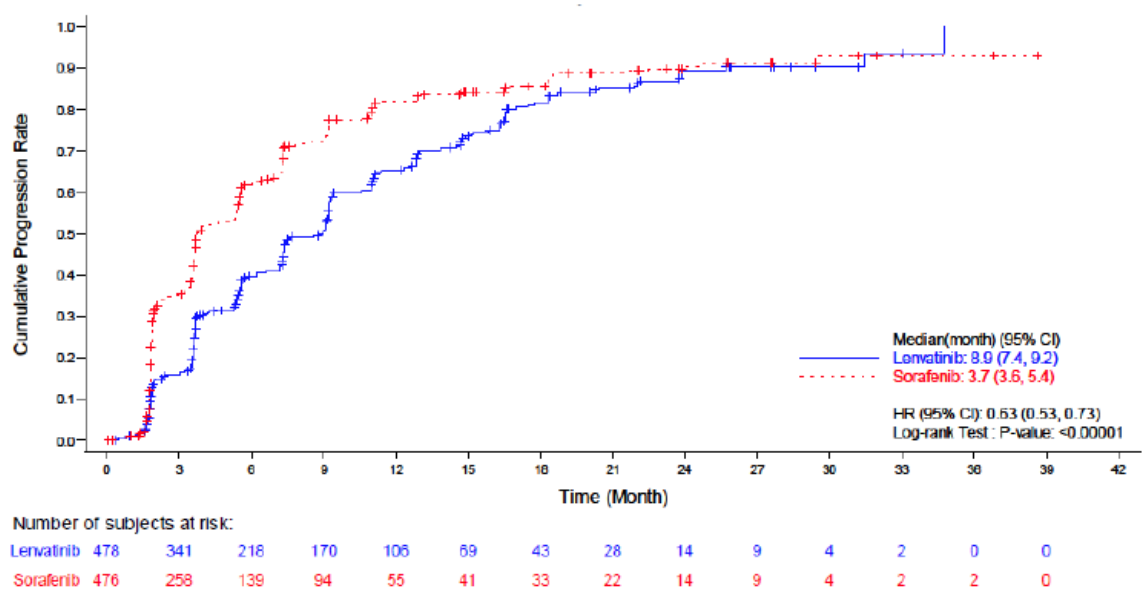
*Deaths were not counted as progression events in this analysis; †Quartiles are estimated by Kaplan-Meier method, and the 95% confidence intervals are estimated with a generalised Brookmeyer and Crowley method; ‡Cumulative progression rate was calculated using the Kaplan-Meier product-limit method and Greenwood Formula; §Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor. Efron method was used for ties; **Stratified by region (Region 1: Asia-Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg); ††Follow-up time for TTP was measured from the date of randomisation to the date of the patient's last follow-up, and it has same numeric value but opposite censoring indicator as compared to TTP.

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; FAS, Full Analysis Set; TTP, time to progression.



**Figure 5: Kaplan-Meier curve for TTP based on randomisation stratification factors**




Abbreviations: CI, confidence interval; HR, hazard ratio.

### *B.2.6.1.4 Secondary efficacy outcome relevant to HE model and/or scope: objective response rate*

The ORR was 24.1% in the lenvatinib arm compared with 9.2% in the sorafenib arm; a statistically significant difference of 14.8% (95% CI: 10.2, 19.4); OR 3.13 (95% CI: 2.15, 4.56; p<0.00001). The proportion of patients with CR, PR and stable disease (StD) are detailed in Table 17. The results were consistent in the PPS; the ORR was [redacted]

**Table 17: Objective response – FAS**

| | Lenvatinib N=478 | Sorafenib N=476 |
|---|---|---|
| BOR, n (%) | | |
| CR | 6 (1.3) | 2 (0.4) |
| PR | 109 (22.8) | 42 (8.8) |
| StD | 246 (51.5) | 244 (51.3) |
| Durable StD (duration of StD ≥23 weeks after randomisation) | 167 (34.9) | 139 (29.2) |
| PD | 71 (14.9) | 147 (30.9) |
| Unknown/not evaluable | 46 (9.6) | 41 (8.6) |
| No baseline tumour assessment | [redacted] | [redacted] |
| No post-baseline tumour assessment | [redacted] | [redacted] |
| 1 or more lesions not evaluable | [redacted] | [redacted] |
| Early StD (StD <7 weeks) | [redacted] | [redacted] |
| ORR (CR + PR), n % | 115 (24.1) | 44 (9.2) |
| 95% CI* | (20.2, 27.9) | (6.6, 11.8) |
| Difference (%) (95% CI)* | 14.8 (10.2, 19.4) | |
| Odds ratio (95% CI)† with stratification factors in IVRS | 3.13 (2.15, 4.56) | |
| P-value† | [redacted] | |
| Odds ratio (95% CI)† with stratification factors in CRF | [redacted] | |
| P-value† | [redacted] | |

*95% CI was calculated using asymptotic normal approximation; †Odds ratio and p-value (for superiority test) were calculated using the Cochran-Mantel-Haenszel method, stratified by IVRS or CRF stratification factors; Abbreviations: BOR, best overall response; CI, confidence interval; CR, complete response; CRF, case report form; ORR, objective response rate; PD, progressive disease; PR, partial response; SD, standard deviation; StD, stable disease.

#### *B.2.6.1.5 Secondary efficacy outcome relevant to HE model and/or scope: HRQoL*

Overall, compliance for the completion of HRQoL questionnaires was high (>90%) and consistent for all measures through the Randomisation Phase.

[redacted]

**Table 18: Available and missing HRQoL data at the specified timepoints among patients in the cross-sectional population**

| | Patients with available data | Patients missing data |
|---|---|---|
| ***EORTC QLQ-HCC18**** | | |
| Baseline | [redacted] | [redacted] |
| Cycle 3 Day 1 | [redacted] | [redacted] |
| Cycle 6 Day 1 | [redacted] | [redacted] |
| Cycle 9 Day 1 | [redacted] | [redacted] |
| Cycle 12 Day 1 | [redacted] | [redacted] |
| Cycle 15 Day 1 | [redacted] | [redacted] |
| Cycle 18 Day 1 | [redacted] | [redacted] |
| Off-Treatment† | [redacted] | [redacted] |
| ***EORTC QLQ-C30‡*** | | |
| Baseline | [redacted] | [redacted] |
| Cycle 3 Day 1 | [redacted] | [redacted] |
| Cycle 6 Day 1 | [redacted] | [redacted] |
| Cycle 9 Day 1 | [redacted] | [redacted] |
| Cycle 12 Day 1 | [redacted] | [redacted] |
| Cycle 15 Day 1 | [redacted] | [redacted] |
| Cycle 18 Day 1 | [redacted] | [redacted] |
| Off-Treatment† | [redacted] | [redacted] |
| ***EQ-5D§*** | | |
| Baseline | [redacted] | [redacted] |
| Cycle 3 Day 1 | [redacted] | [redacted] |
| Cycle 6 Day 1 | [redacted] | [redacted] |
| Cycle 9 Day 1 | [redacted] | [redacted] |
| Cycle 12 Day 1 | [redacted] | [redacted] |
| Cycle 15 Day 1 | [redacted] | [redacted] |
| Cycle 18 Day 1 | [redacted] | [redacted] |
| Off-Treatment† | [redacted] | [redacted] |

*Patients were administered the QLQ-HCC18 instrument at Baseline, Day 1 of each treatment cycle, and at the Off-treatment visit. The instrument includes 18 items which enhance the sensitivity and specificity of the QLQ-C30 in HCC-related issues. Items in the questionnaire are grouped into 8 scales and have response categories ranging from 1 to 4, where 1 = "not at all" and 4 = "very much"; †Visit occurred within 30 days following the final dose of study treatment; ‡Patients were administered the QLQ-C30 instrument at Baseline, Day 1 of each treatment cycle, and at the OTV. The instrument includes 30 items which are grouped into 5 functional domains and 9 symptom domains. There is also a single global Quality of Life (QoL)/global health status score. Items have response categories either ranging from 1 to 4, (where 1 = "not at all" and 4 = "very much") or 1 to 7 (where 1 = "very poor" and 7 = "excellent"). A summary score based on 27 of the 30 items can also be calculated. It is scaled from 0 to 100, where a higher score indicates more favourable patient outcomes; §Patients were administered the EQ-5D-3L instrument at Baseline, Day 1 of each treatment cycle and at the OTV. The instrument includes 5 items which address the domains of Mobility Self-Care Usual Activities Pain/Discomfort, and Anxiety/Depression. Items in questionnaire use the patient rating of 1 to 3, where 1 = "no problems" and 3 = "extreme problems." The scores of the 5 health state domains are scored into a single health utility index. The EQ-VAS (EuroQoL-Visual Analogue Scale) measures current health status from 0 to 100 where 0 is given the verbal label "worst imaginable health state" and 100 is given the label "best imaginable health state."

Abbreviations: EORTC, European Organisation for Research and Treatment of Cancer; HCC, hepatocellular carcinoma; HRQoL, health-related quality of life; EQ-5D, EuroQol-5 Dimensions; QLQ, Quality of Life Questionnaire.



Over the course of the study, HRQoL results became more variable and thus less interpretable due to attrition. Therefore, cross-sectional results were considered for Baseline and Cycles 3, 6, 9, 12, 15 and 18 only.

***EORTC QLQ-HCC18***

[redacted]

***EORTC QLQ-C30***

[redacted]

***EQ5D***

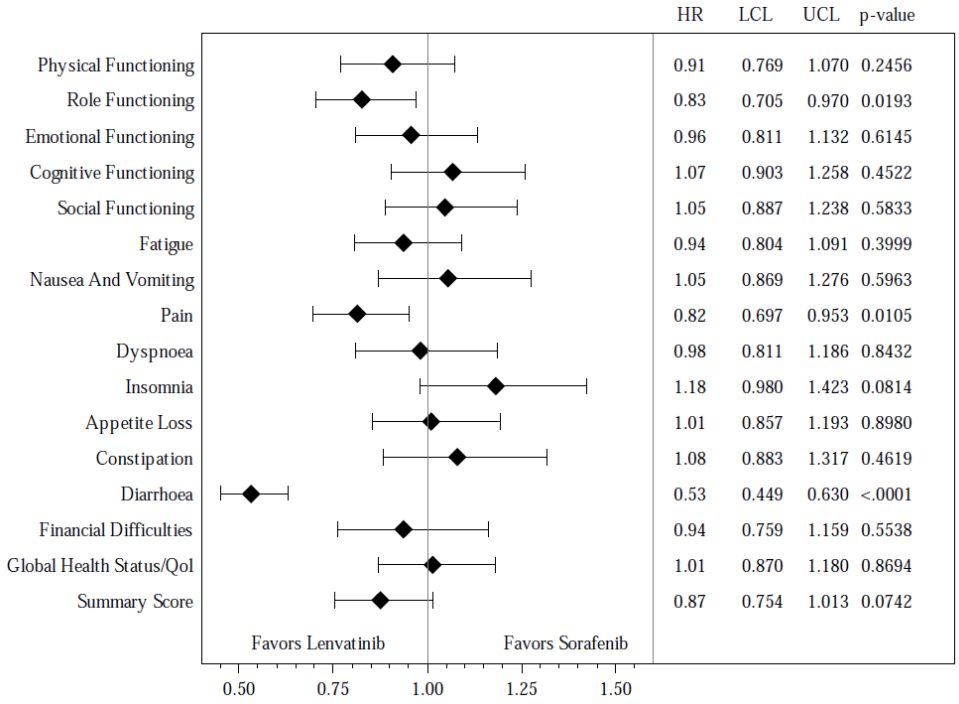
- Results of the EQ-5D showed that there were no statistically significant differences in HUI or Visual Analogue Scale (VAS) scores between treatment groups at Cycles 3, 6, 9, 12, 15, or 18 ($p>0.05$). The cross-sectional analysis of EQ-5D-HUI scores by timepoint is presented in Appendix L, Section L.4.

### *Clinically meaningful worsening in HRQoL*

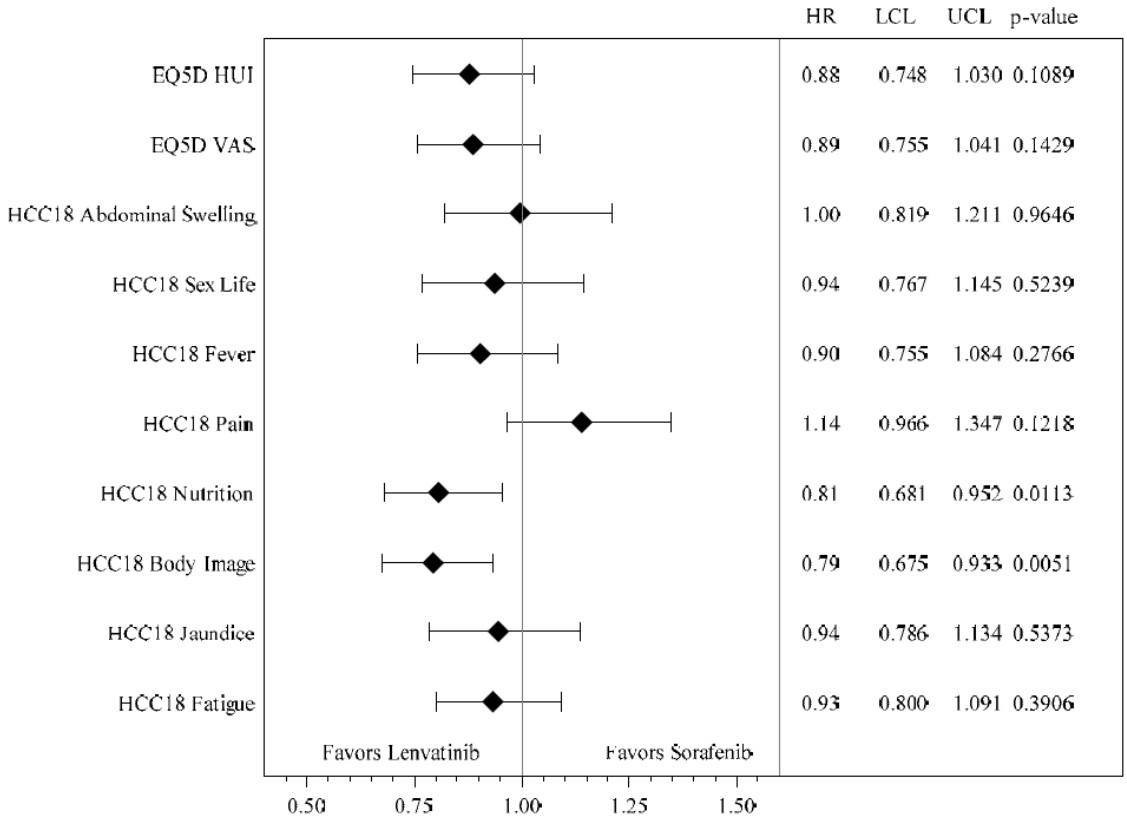
As highlighted in Figure 6 and Figure 7, patients treated with sorafenib experienced statistically greater levels of risk of a clinically meaningful worsening of Role Functioning (EORTC QLQ-C30), Pain (EORTC QLQ-C30), Diarrhoea (EORTC QLQ-C30), Body Image (EORTC QLQ-HCC18), and Nutrition (EORTC QLQ-HCC18) earlier in treatment over the course of the study compared with lenvatinib-treated patients.

**Figure 6: Hazard ratio of time to clinically meaningful worsening – EORTC QLQ-C30 domains**




Abbreviations: EORTC, European Organisation for Research and Treatment of Cancer; HR, hazard ratio; LCL, lower confidence level; QLQ, Quality of Life Questionnaire; QoL, quality of life; UCL, upper confidence level.

**Figure 7: Hazard ratio of time to clinically meaningful worsening – EQ-5D and EORTC QLQ-HCC18 domains**




Abbreviations: EORTC, European Organisation for Research and Treatment of Cancer; HCC, hepatocellular carcinoma; HR, hazard ratio; HUI, Health Utilities Index; LCL, lower confidence level; QoL, quality of life; QLQ, Quality of Life Questionnaire; UCL, upper confidence level; VAS, Visual Analogue Scale.

#### *B.2.6.1.6 Conclusion*

The REFLECT study met its primary endpoint, demonstrating non-inferiority of lenvatinib to sorafenib in terms of OS. Median OS was numerically longer but not statistically significant for lenvatinib (13.6 months) compared with sorafenib (12.3 months; HR=0.92 [95% CI 0.79, 1.06]) and superiority of lenvatinib over sorafenib could not be demonstrated. Lenvatinib has therefore demonstrated a proven overall survival benefit by statistical confirmation of non-inferiority when compared with sorafenib.

Baseline imbalances for the prognostic factor AFP ≥200 ng/mL and aetiology of HCV favoured sorafenib, and therefore do not affect the validity of the non-inferiority result for lenvatinib OS vs sorafenib. Following adjustment by baseline AFP category,

lenvatinib was nominally superior to sorafenib in prolonging OS (HR 0.856, p=0.0342).

Lenvatinib treatment resulted in statistically significant and clinically meaningful improvements compared with sorafenib for the secondary endpoints of PFS, TTP and ORR (based on investigator assessment by mRECIST; both p<0.00001). All results were consistent between the FAS and the PPS.

The QoL impact of lenvatinib and sorafenib was broadly equivalent across the majority of function and symptom areas, however there was a clinically meaningful delay in worsening for lenvatinib compared with sorafenib across several domains including diarrhoea, nutrition and pain.

### *B.2.7 Subgroup analysis*

Pre-planned subgroup analyses were performed in order to identify any differences in response to treatment as a result of baseline patient demographic or disease characteristics. Subgroups were based on randomisation stratification factors and other factors including age group, sex, aetiology, AFP at baseline and post-treatment anti-cancer therapy as described in Section B.2.3.1.11.

Demographics and disease characteristics by geographic region (Western and Asia-Pacific) are presented in Appendix E Section E.1. The demographics and disease characteristics for the remaining subgroups were not defined.

Results for the OS and PFS endpoints (reported as HR lenvatinib:sorafenib [95% CI]) are presented in Appendix E, Section E.2 and E.3, respectively.

The effect of lenvatinib and sorafenib on OS was generally consistent across subgroups (HR lenvatinib:sorafenib <1) and similar to the overall population. One exception was the Western region subgroup where the HR for lenvatinib:sorafenib (95% CI) was 1.08 (0.82, 1.42), numerically in favour of sorafenib. The effect of lenvatinib on OS was consistent across the Western and Asia-Pacific subgroups (13.6 months [95% CI: 11.5, 17.7] and 13.5 months [95% CI: 11.7,15.1], respectively), whereas median OS with sorafenib in the Western region was longer

than in the Asia-Pacific region (14.2 months [95% CI: 11.9, 18.0] and 11.0 months [95% CI: 9.6, 12.5], respectively). This result can be explained by a notable imbalance in post-treatment anti-cancer therapy during survival follow-up in the Western region (lenvatinib 43.1%, sorafenib 51.1%, respectively) which was an artefact of the clinical trial design (see Section B.2.6.1.1.3).

## *B.2.8 Meta-analysis*

REFLECT is the only RCT reporting on the efficacy and safety of lenvatinib in patients with unresectable HCC. Therefore, a meta-analysis was not required.

## *B.2.9 Indirect and mixed treatment comparisons*

As head-to-head clinical trial data were available for lenvatinib vs sorafenib (the only relevant comparator for the submission [see Table 1]), an indirect or mixed-treatment comparison was not necessary.

## *B.2.10 Adverse reactions*

### B.2.10.1 REFLECT

Adverse event data were recorded in the REFLECT trial. Data for the Safety Analysis Set, which included 476 lenvatinib-treated patients and 475 sorafenib-treated patients who received at least one dose of study drug is presented in this section.

At the time of data cut-off for REFLECT (13th November 2016), the median duration of treatment in the lenvatinib and sorafenib arms was 5.7 and 3.7 months, respectively (Table 19); median duration of lenvatinib therapy was approximately 1.5-times that of sorafenib. In both treatment arms, exposure (i.e. dose intensity) versus planned dose was high (Table 19).

**Table 19: Extent of exposure to study treatment – Safety Analysis Set**

| | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg* (N = 151) | 12 mg* (N = 325) | Total (N = 476) | 800 mg (N = 475) |
| Duration of Treatment† (months) | | | | |
| Mean (SD) | 7.6 (6.47) | 8.5 (7.27) | 8.2 (7.04) | 6.0 (6.47) |
| Median | 5.6 | 6.3 | 5.7 | 3.7 |
| Q1, Q3 | 2.4, 11.0 | 3.2, 12.0 | 2.9, 11.1 | 1.8, 7.4 |
| Min, max | 0.1, 33.7 | 0.0, 35.0 | 0.0, 35.0 | 0.1, 38.7 |
| Duration of Treatment (months), n (%) | | | | |
| 0 - <1 | 15 (9.9) | 24 (7.4) | 39 (8.2) | 31 (6.5) |
| 1 - <2 | 15 (9.9) | 35 (10.8) | 50 (10.5) | 124 (26.1) |
| 2 - <4 | 34 (22.5) | 60 (18.5) | 94 (19.7) | 114 (24.0) |
| 4 - <6 | 17 (11.3) | 42 (12.9) | 59 (12.4) | 53 (11.2) |
| 6 - <8 | 10 (6.6) | 36 (11.1) | 46 (9.7) | 47 (9.9) |
| 8 - <12 | 29 (19.2) | 50 (15.4) | 79 (16.6) | 45 (9.5) |
| 12 - <18 | 19 (12.6) | 39 (12.0) | 58 (12.2) | 28 (5.9) |
| ≥18 | 12 (7.9) | 39 (12.0) | 51 (10.7) | 33 (6.9) |
| No. of patient months‡ | 1141.3 | 2748.9 | 3890.2 | 2869.1 |
| Dose Intensity (mg/day/patient) | | | | |
| Mean (SD) | 7.0 (1.59) | 10.5 (6.54) | 9.4 (5.71) | 663.8 (173.15) |
| Median | 8.0 | 11.5 | 8.9 | 771.4 |
| Q1, Q3 | 6.3, 8.0 | 8.7, 12.0 | 7.9, 12.0 | 514.6, 800.0 |
| Min, Max | 2.1, 8.0 | 1.7, 120.0 | 1.7, 120.0 | 126.3, 800.0 |

* 8 mg and 12 mg were the starting doses of lenvatinib based on the patients' body weight (<60 kg, ≥60 kg) at Baseline; †Duration of treatment (months) = (Date of last dose of study drug - Date of first dose of study drug + 1) ÷30.4375; ‡Number of patient-months = Sum of all months that all patients received study drug based on treatment duration.

Abbreviations: SD, standard deviation.

Table 20 presents a summary of treatment-emergent adverse events (TEAEs) reported in the Safety Analysis Set of REFLECT. Because of the substantially longer exposure to lenvatinib than to sorafenib, TEAEs were also analysed adjusted by treatment exposure (episodes/patient year [PY]; Table 21).

**Table 20: Summary of TEAEs – Safety Analysis Set**

| Adverse events | Lenvatinib | Sorafenib |
|---|---|---|

| | 8 mg* N=151 n (%) | 12 mg* N=325 n (%) | Total N=476 n (%) | (N=475) n (%) |
|---|---|---|---|---|
| Patient with any TEAE | 151 (100.0) | 319 (98.2) | 470 (98.7) | 472 (99.4) |
| Patients with any related TEAE | 143 (94.7) | 304 (93.5) | 447 (93.9) | 452 (95.2) |
| Patients with any TEAE ≥Grade† 3 | 100 (66.2) | 257 (79.1) | 357 (75.0) | 316 (66.5) |
| Patient with any related TEAE ≥Grade† 3 | 70 (46.4) | 200 (61.5) | 270 (56.7) | 231 (48.6) |
| Number of patients with any serious AE‡ | 58 (38.4) | 147 (45.2) | 205 (43.1) | 144 (30.3) |
| Number of patients with any fatal SAE§ | 14 (9.3) | 47 (14.5) | 61 (12.8) | 36 (7.6) |
| Number of patients with non-fatal SAEs | 54 (35.8) | 135 (41.5) | 189 (39.7) | 128 (26.9) |
| Number of patients with:‡ | | | | |
| TEAEs leading to study drug withdrawal | 33 (21.9) | 61 (18.8) | 94 (19.7) | 69 (14.5) |
| TEAEs leading to study drug dose reduction | 43 (28.5) | 141 (43.4) | 184 (38.7) | 185 (38.9) |
| TEAEs leading to study drug interruption | 72 (47.7) | 176 (54.2) | 248 (52.1) | 193 (40.6) |
| TEAEs leading to study drug dose reduction or interruption | 81 (53.6) | 213 (65.5) | 294 (61.8) | 264 (55.6) |

*8 mg and 12 mg were the lenvatinib starting doses based on the subjects' body weight (<60 kg, ≥60 kg) at Baseline; †Adverse events were graded using CTCAE version 4.0; ‡Patients may be counted in more than 1 subcategory; §Category includes 70 subjects who had a TEAE ongoing at the time of death due to disease progression or whose cause of death was unknown.

Abbreviations: AE, adverse event; CI, confidence interval; CTCAE, Common Terminology Criteria for Adverse Events; HCC, hepatocellular carcinoma; SAE, serious adverse event; TEAE, treatment-emergent adverse event.

**Table 21: Summary of TEAEs by treatment exposure (episodes/PY) – Safety Analysis Set**

| Adverse events | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg* (N = 151) Total Duration= 95.1 years n (AE rate) | 12 mg* (N= 325) Total Duration= 229.1 years n (AE rate) | Total (N=496) Total Duration= 324.2 years n (AE rate) | 800 mg (N=475) Total Duration= 239.1 years n (AE rate) |
| Any TEAE episodes | 1,737 (18.26) | 4,387 (19.15) | 6,124 (18.89) | 4,718 (19.73 |
| Related TEAE episodes | 974 (10.24) | 2,572 (11.23) | 3,546 (10.94) | 2,865 (11.98) |
| Any TEAE ≥Grade 3 episodes | 278 (2.92) | 745 (3.25) | 1,023 (3.16) | 795 (3.33) |
| Related TEAE ≥Grade 3 | 126 (1.32) | 391 (1.71) | 517 (1.59) | 430 (1.80) |

| Adverse events | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg* (N = 151) Total Duration= 95.1 years n (AE rate) | 12 mg* (N= 325) Total Duration= 229.1 years n (AE rate) | Total (N=496) Total Duration= 324.2 years n (AE rate) | 800 mg (N=475) Total Duration= 239.1 years n (AE rate) |
| episodes | | | | |
| Serious AE episodes | 113 (1.19) | 296 (1.29) | 409 (1.26) | 232 (0.97) |
| Fatal serious TEAE episodes † | 14 (0.15) | 47 (0.21) | 61 (0.19) | 36 (0.15) |
| Non-fatal serious TEAE episodes | 108 (1.14) | 271 (1.18) | 379 (1.17) | 207 (0.87) |

*8 mg and 12 mg were the lenvatinib starting doses based on the subjects' body weight (<60 kg, ≥60 kg) at Baseline; †Fatal AE episodes were counted only once per patient, if more than 1 fatal AE was reported for the same patient.

Abbreviations: AE, adverse event; HCC, hepatocellular carcinoma; TEAE, treatment-emergent adverse event.

Table 22 presents a summary of TEAEs that occurred in ≥10% of patients in any group by patient incidence and by treatment exposure-adjusted rate. The relative risk, risk difference and associated 95% confidence intervals for each adverse reaction are presented in Appendix L, Section L.5.

**Table 22: TEAEs occurring in ≥10% of patients by preferred term – Safety Analysis Set**

| | Patient incidence | | | | Treatment exposure-adjusted rate (episodes/PY) | | | |
|---|---|---|---|---|---|---|---|---|
| | Lenvatinib | | | Sorafenib | Lenvatinib | | | Sorafenib |
| **Preferred term** | **8 mg* N=151 n (%)** | **12 mg* N=325 n (%)** | **Total N=476 n (%)** | **800 mg N=475 n (%)** | **8 mg* (N = 151) Total Duration= 95.1 years n (AE Rate)** | **12 mg* (N = 325) Total Duration= 229.1 years n (AE Rate)** | **Total (N = 476) Total Duration= 324.2 years n (AE Rate)** | **800 mg N = 475 Total Duration= 239.1 years n (AE Rate)** |
| Patient with any TEAE | 151 (100.0) | 319 (98.2) | 470 (98.7) | 472 (99.4) | 1,737 (18.26) | 4,387 (19.15) | 6,124 (18.89) | 4,718 (19.73) |
| Hypertension | 65 (43.0) | 136 (41.8) | 201 (42.2) | 144 (30.3) | 75 (0.79) | 178 (0.78) | 253 (0.78) | 166 (0.69) |
| Diarrhoea | 53 (35.1) | 131 (40.3) | 184 (38.7) | 220 (46.3) | 101 (1.06) | 226 (0.99) | 327 (1.01) | 351 (1.47) |
| Decreased appetite | 50 (33.1) | 112 (34.5) | 162 (34.0) | 127 (26.7) | 60 (0.63) | 136 (0.59) | 196 (0.60) | 139 (0.58) |
| Weight decreased | 43 (28.5) | 104 (32.0) | 147 (30.9) | 106 (22.3) | 48 (0.50) | 117 (0.51) | 165 (0.51) | 113 (0.47) |
| Fatigue | 42 (27.8) | 99 (30.5) | 141 (29.6) | 119 (25.1) | 49 (0.52) | 107 (0.47) | 156 (0.48) | 130 (0.54) |
| Palmar-plantar erythrodysaesthesia syndrome | 35 (23.2) | 93 (28.6) | 128 (26.9) | 249 (52.4) | 36 (0.38) | 106 (0.46) | 142 (0.44) | 289 (1.21) |
| Proteinuria | 37 (24.5) | 80 (24.6) | 117 (24.6) | 54 (11.4) | 53 (0.56) | 110 (0.48) | 163 (0.50) | 74 (0.31) |
| Dysphonia | 28 (18.5) | 85 (26.2) | 113 (23.7) | 57 (12.0) | 29 (0.30) | 102 (0.45) | 131 (0.40) | 66 (0.28) |
| Nausea | 24 (15.9) | 69 (21.2) | 93 (19.5) | 68 (14.3) | 30 (0.32) | 83 (0.36) | 113 (0.35) | 77 (0.32) |
| Platelet count decreased | 26 (17.2) | 61 (18.8) | 87 (18.3) | 58 (12.2) | 33 (0.35) | 84 (0.37) | 117 (0.36) | 67 (0.28) |
| Abdominal pain | 19 (12.6) | 62 (19.1) | 81 (17.0) | 87 (18.3) | 21 (0.22) | 87 (0.38) | 108 (0.33) | 106 (0.44) |
| Hypothyroidism | 25 (16.6) | 53 (16.3) | 78 (16.4) | 8 (1.7) | 25 (0.26) | 54 (0.24) | 79 (0.24) | 8 (0.03) |
| Vomiting | 22 (14.6) | 55 (16.9) | 77 (16.2) | 36 (7.6) | 30 (0.32) | 76 (0.33) | 106 (0.33) | 57 (0.24) |
| Constipation | 19 (12.6) | 57 (17.5) | 76 (16.0) | 52 (10.9) | 20 (0.21) | 69 (0.30) | 89 (0.27) | 60 (0.25) |

| | Patient incidence | | | | Treatment exposure-adjusted rate (episodes/PY) | | | |
|---|---|---|---|---|---|---|---|---|
| | Lenvatinib | | | Sorafenib | Lenvatinib | | | Sorafenib |
| **Preferred term** | **8 mg* N=151 n (%)** | **12 mg* N=325 n (%)** | **Total N=476 n (%)** | **800 mg N=475 n (%)** | **8 mg* (N = 151) Total Duration= 95.1 years n (AE Rate)** | **12 mg* (N = 325) Total Duration= 229.1 years n (AE Rate)** | **Total (N = 476) Total Duration= 324.2 years n (AE Rate)** | **800 mg N = 475 Total Duration= 239.1 years n (AE Rate)** |
| Blood bilirubin increased | 23 (15.2) | 48 (14.8) | 71 (14.9) | 63 (13.3) | 31 (0.33) | 66 (0.29) | 97 (0.30) | 75 (0.31) |
| Pyrexia | 24 (15.9) | 45 (13.8) | 69 (14.5) | 63 (13.3) | 28 (0.29) | 52 (0.23) | 80 (0.25) | 72 (0.30) |
| Ascites | 21 (13.9) | 47 (14.5) | 68 (14.3) | 44 (9.3) | 22 (0.23) | 55 (0.24) | 77 (0.24) | 51 (0.21) |
| Oedema peripheral | 23 (15.2) | 43 (13.2) | 66 (13.9) | 33 (6.9) | 23 (0.24) | 62 (0.27) | 85 (0.26) | 40 (0.17) |
| Aspartate aminotransferase increased | 21 (13.9) | 44 (13.5) | 65 (13.7) | 80 (16.8) | 28 (0.29) | 54 (0.24) | 82 (0.25) | 100 (0.42) |
| Abdominal pain upper | 21 (13.9) | 37 (11.4) | 58 (12.2) | 40 (8.4) | 30 (0.32) | 43 (0.19) | 73 (0.23) | 48 (0.20) |
| Asthenia | 14 (9.3) | 40 (12.3) | 54 (11.3) | 48 (10.1) | 16 (0.17) | 60 (0.26) | 76 (0.23) | 54 (0.23) |
| Alanine aminotransferase increased | 17 (11.3) | 36 (11.1) | 53 (11.1) | 52 (10.9) | 21 (0.22) | 46 (0.20) | 67 (0.21) | 66 (0.28) |
| Back pain | 11 (7.3) | 39 (12.0) | 50 (10.5) | 31 (6.5) | 12 (0.13) | 43 (0.19) | 55 (0.17) | 32 (0.13) |
| Rash | 18 (11.9) | 28 (8.6) | 46 (9.7) | 76 (16.0) | 20 (0.21) | 31 (0.14) | 51 (0.16) | 87 (0.36) |
| Stomatitis | 11 (7.3) | 34 (10.5) | 45 (9.5) | 56 (11.8) | 12 (0.13) | 43 (0.19) | 55 (0.17) | 67 (0.28) |
| Alopecia | 5 (3.3) | 9 (2.8) | 14 (2.9) | 119 (25.1) | 6 (0.06) | 9 (0.04) | 15 (0.05) | 123 (0.51) |

*8 mg and 12 mg were the lenvatinib starting doses based on the subjects' body weight (<60 kg, ≥60 kg) at Baseline.

Abbreviations: AE, adverse event; HCC, hepatocellular carcinoma; TEAE, treatment-emergent adverse event.

Table 23 presents details of AEs that are incorporated into the economic model (see Section B.3.3.2). These include Grade 3 or 4 TEAEs that occurred in ≥5% or more of patients in either treatment arm, and additional Grade 3 or 4 TEAEs occurring in <5% of either treatment arm that were considered to be clinically or economically significant by UK clinical experts (Section B.3.3.3). The relative risk, risk difference and associated 95% confidence intervals for each of the AEs considered in the model is presented in Appendix L, Section L.5.

**Table 23: Grade 3 or 4 TEAEs occurring in ≥5% of patients in any treatment group or considered clinically relevant by preferred term – Safety Analysis Set**

| | Lenvatinib 8 or 12 mg (N=476) n (%) | | Sorafenib 800 mg (N=475) n (%) | |
|---|---|---|---|---|
| | Grade 3 | Grade 4 | Grade 3 | Grade 4 |
| Patients with any Grade 3 or 4 TEAEs | 260 (54.6) | 36 (7.6) | 248 (52.2) | 32 (6.7) |
| Hypertension | 111 (23.3) | 0 (0.0) | 68 (14.3) | 0 (0.0) |
| Weight decreased | 36 (7.6) | 0 (0.0) | 14 (2.9) | 0 (0.0) |
| Proteinurea | 27 (5.7) | 0 (0.0) | 8 (1.7) | 0 (0.0) |
| Platelet count decreased | 26 (5.5) | 0 (0.0) | 14 (2.9) | 2 (0.4) |
| Blood bilirubin increased | 24 (5.0) | 7 (1.5) | 21 (4.4) | 2 (0.4) |
| Aspartate aminotransferase increased | 21 (4.4) | 3 (0.6) | 32 (6.7) | 6 (1.3) |
| Gamma-glutamyl transferase increased | 20 (4.2) | 6 (1.3) | 16 (3.4) | 3 (0.6) |
| Palmar-plantar erythrodysaesthesia syndrome | 14 (2.9) | 0 (0.0) | 54 (11.4) | 0 (0.0) |
| Diarrhoea | 20 (4.2) | 0 (0.0) | 20 (4.2) | 0 (0.0) |
| Fatigue | 18 (3.8) | 0 (0.0) | 17 (3.6) | 0 (0.0) |
| Asthenia | 14 (2.9) | 0 (0.0) | 11 (2.3) | 0 (0.0) |

Off-Abbreviations: HCC, hepatocellular carcinoma; TEAE, treatment-emergent adverse event.

### B.2.10.2 Additional studies

There are no additional studies reporting adverse reactions to lenvatinib other than those reported in Section B.2.10.1.

### B.2.10.3 Safety overview

- The overall lenvatinib AE profile from the REFLECT study was consistent with that observed in previous studies of its other indications (radioiodine-refractory differentiated thyroid cancer and renal cell carcinoma).

- Lenvatinib and sorafenib have safety profiles that are consistent with other VEGF/VEGFR-targeted therapies; however, the nature and extent of AEs differed between the two agents based on their different mechanisms of action.
- Nearly all patients treated with either lenvatinib or sorafenib experienced a TEAE and most patients experienced an AE that was considered related to study treatment (93.9% of lenvatinib-treated patients and 95.2% of sorafenib-treated patients).
- The most frequently reported TEAEs (>30% of patients) with lenvatinib were hypertension, diarrhoea, decreased appetite, and weight decreased. The most frequently reported TEAEs (>30% of patients) with sorafenib were palmar-plantar erythrodysaesthesia syndrome, diarrhoea, and hypertension.
- The median duration of exposure to therapy in the REFLECT study was approximately 1.5x higher with lenvatinib than sorafenib; because of the substantially longer exposure to lenvatinib, TEAEs were also analysed adjusted by treatment exposure.
- Although grade ≥3 TEAEs occurred at a higher frequency in the lenvatinib arm (75%) than the sorafenib arm (66.5%), when adjusted for treatment duration, the event rate was similar (3.16 and 3.33 episodes/PY, respectively.
- Grade 3 or 4 TEAEs that occurred in ≥5% of the lenvatinib arm were hypertension, weight decreased, proteinurea, platelet count decreased, blood bilirubin increased, aspartate aminotransferase increased and gamma-glutamyl transferase increased. Grade 3 or 4 TEAEs that occurred in ≥5% of the sorafenib arm were hypertension, aspartate aminotransferase increased, blood bilirubin increased and palmar-plantar erythrodysaesthesia syndrome.
- Fatal AEs were more frequent in the lenvatinib arm (12.8% of patients vs 7.6% of patients in the sorafenib arm), however when adjusted for treatment exposure, the rate was similar (0.19 episodes/PY in the lenvatinib arm vs 0.15 episodes/PY in the sorafenib arm).

### *B.2.11 Ongoing studies*

There are no additional ongoing Phase II or Phase III studies investigating the efficacy and/or safety of lenvatinib in HCC.

### *B.2.12 Innovation*

Lenvatinib is considered innovative as it is a multiple receptor tyrosine kinase (RTK) inhibitor with a novel binding mode that inhibits the kinase activities of VEGF receptors (VEGFR1, VEGFR2 and VEGFR3) and FGF receptors (FGFR1, FGFR2, FGFR3 and FGFR4) in addition to other proangiogenic and oncogenic pathway-related RTKs (including platelet derived growth factor receptor PDGFRα; KIT; and RET) involved in tumour proliferation.

Sorafenib is currently the only available systemic treatment option for patients with advanced HCC in England and Wales and there is a clear unmet need for new treatments which delay progression and improve survival without negatively impacting patients' quality of life.

Lenvatinib has demonstrated a proven overall survival benefit by statistical confirmation of non-inferiority when compared with sorafenib in the REFLECT trial. Furthermore, lenvatinib demonstrated a significant and clinically meaningful benefit in terms of PFS, ORR and TTP compared with sorafenib, with a manageable AE profile which may be favourable compared with sorafenib from a patient perspective due to lower rates of palmar plantar erythrodysesthesia syndrome and diarrhoea, which can have a substantial impact on QoL. Importantly, results from the EORTC-QLQ (C30 and HCC18) suggest that lenvatinib treatment leads to a clinically meaningful delay in several HRQoL domains, including Role Functioning, Pain, Diarrhoea, Body Image, and Nutrition.

Lenvatinib is the only therapy to demonstrate comparable survival rates to sorafenib; over the past decade, several other investigational therapies have failed to meet the endpoints of non-inferiority or superiority for OS compared with sorafenib (sunitinib (27), brivanib (28), linifanib (29), sorafenib + erlotinib (30), sorafenib + doxorubicin (31)).

## *B.2.13 Interpretation of clinical effectiveness and safety evidence*

### B.2.13.1 Principal (interim) findings from the clinical evidence highlighting the clinical benefits and harms of the technology

The clinical benefits of lenvatinib were demonstrated in a pivotal, Phase III active controlled RCT, REFLECT. Lenvatinib has demonstrated a proven overall survival benefit by statistical confirmation of non-inferiority when compared with sorafenib. Whilst OS was numerically higher in the lenvatinib group (13.6 months) than the sorafenib group (12.3 months [HR 0.92, 95% CI: 0.79-1.06]), the difference was not large enough to achieve statistical superiority. Lenvatinib demonstrated statistically significant and clinically meaningful improvements versus sorafenib for the secondary endpoints of PFS, TTP and ORR. Although HRQoL scores were generally similar between the lenvatinib and sorafenib, there were several domains in which lenvatinib patients were observed to have nominally significantly better and clinically meaningful results compared to patients treated with sorafenib at various cycles of therapy in the cross-sectional and longitudinal analyses.

The effect of lenvatinib and sorafenib on OS was generally consistent (HR lenvatinib:sorafenib <1) across a range of subgroups (including age, sex and body weight) and similar to the overall population. One exception was the Western region subgroup, where the HR (95% CI) for lenvatinib:sorafenib was 1.08 (0.82, 1.42). While the effect of lenvatinib on OS was consistent across the Western and Asia-Pacific subgroups (13.6 months [95% CI: 11.5, 17.7] and 13.5 months [95% CI: 11.7,15.1], respectively), median OS with sorafenib was longer in the Western region than in the Asia-Pacific region (14.2 months [95% CI: 11.9, 18.0] and 11.0 months [95% CI: 9.6, 12.5], respectively). However, there was a notable imbalance between the lenvatinib and sorafenib arms of the Western subgroup regarding the proportion of patients who received post-treatment anti-cancer therapy during survival follow-up (43.1% and 51.1%, respectively) that was not observed in the Asia-Pacific region. When OS was adjusted for post-treatment anti-cancer therapy, the HR (95% CI) for lenvatinib:sorafenib [redacted]. The impact of post-treatment anti-cancer therapy on the OS results is discussed further in Section B.2.13.2.

The proportion of patients with baseline AFP levels ≥200 ng/mL, which is a marker for poor HCC prognosis (16, 17), was greater in the lenvatinib arm (46.4%) than in the sorafenib arm (39.3%). After adjustment for baseline AFP, lenvatinib was found to be nominally superior to sorafenib in terms of OS (median OS 13.6 months vs 12.3 months, respectively; HR 0.856 [95% CI:0.736, 0.995], p=0.0342).

Lenvatinib demonstrated statistically significant and clinically meaningful improvements versus sorafenib for the secondary endpoints of PFS, TTP and ORR. Median PFS was 7.4 months in the lenvatinib arm compared with 3.7 months in the sorafenib arm (HR: 0.66; 95% CI: 0.57, 0.77; p<0.00001). PFS results from a post-hoc, blinded, retrospective IIR using both mRECIST and RECIST 1.1 criteria were consistent with the investigator assessments of PFS. The effect of lenvatinib on median PFS across subgroups was consistent with the results for the overall population.

Meaningful tumour shrinkage is an important management tool for HCC. The promising ORR results observed with lenvatinib in the REFLECT trial may potentially enable tumour downstaging, and thereby, allow patients with unresectable tumours, or those that cannot receive other local interventions (such as transplantation, radiofrequency ablation, percutaneous ethanol injection, TACE), to become eligible for these types of treatments. While this was not an option in REFLECT, based on the design of the trial, it is conceivable and would be appealing in clinical practice.

The analysis of the EORTC QLQ-C30, QLQ-HCC18 and EQ-5D responses generally indicated equivalent scores between the lenvatinib and sorafenib treatment arms. Although QoL scores worsened with both treatments over the course of the study, a clinically meaningful delay in deterioration for multiple domains, including diarrhoea, nutrition, and pain (all of which specifically affect the patients' daily lives), was observed with lenvatinib compared with sorafenib.

The overall lenvatinib AE profile from the REFLECT study was consistent with that observed in previous studies of its other indications (radioiodine-refractory differentiated thyroid cancer and renal cell carcinoma). While lenvatinib and sorafenib have safety profiles that are consistent with other VEGF/VEGFR-targeted

therapies, the nature and extent of AEs differed between the two agents based on their different mechanisms of action. The most frequently reported TEAEs (>30% of patients) with lenvatinib were hypertension, diarrhoea, decreased appetite, and weight decreased. The most frequently reported TEAEs (>30% of patients) with sorafenib were palmar-plantar erythrodysaesthesia syndrome, diarrhoea, and hypertension. Although grade ≥3 TEAEs and fatal AEs occurred at a higher frequency in the lenvatinib arm than the sorafenib arm, the median duration of exposure to therapy was 1.5x longer with lenvatinib than sorafenib. When AEs were adjusted for treatment duration, the rate of grade ≥3 TEAEs and fatal AEs was similar between the treatments.

The REFLECT trial is the first study in 10 years to demonstrate a proven overall survival benefit by statistical confirmation of non-inferiority when compared with sorafenib, and as such represents an advancement in HCC therapy. Several other therapies, including sunitinib, brivanib, linifanib, sorafenib + erlotinib and sorafenib + doxorubicin have failed to demonstrate non-inferiority or superiority for OS compared with sorafenib in RCTs (27-31). In addition, lenvatinib also demonstrated statistically significant and clinically meaningful benefits compared with sorafenib regarding disease progression and overall response and is therefore a highly relevant additional treatment for the management of HCC.

### B.2.13.2 Strengths and limitations of the clinical evidence base for the technology

#### *Internal validity*

REFLECT was a large, multinational, well-controlled and well-conducted study. An open-label design was chosen in the interests of patient safety because dose modification guidelines due to toxicity were different for lenvatinib and sorafenib for the same toxicities. In addition, the differences in formulation between lenvatinib (capsule) and sorafenib (tablet) would have required preparation of multiple matching placebo capsules or tablets to permit dose reductions. This would have been confusing for very ill patients and would have resulted in a high risk for dosing errors.

The sponsor created a Data Integrity Protection Plan, which ensured that relevant data fields were masked so that the clinical and statistical team members were blinded to treatment for each patient in order to minimise bias.

The entry criteria for REFLECT were carefully defined based on those used in the historical sorafenib trials that evaluated the effectiveness of first-line treatment of HCC and agreed with the regulatory authorities. Concomitant medications were allowed during the study only to relieve symptoms or to treat complications; no anti-VEGF and no other systemic or localised anti-cancer therapy for HCC was permitted, thus reducing any possibility of distorting the perceived effects of lenvatinib and sorafenib.

Sorafenib was selected as the comparator because it was the only drug approved for the systemic treatment of unresectable HCC in over 100 countries at the time of study design. Sorafenib remains the only approved systemic therapy for advanced HCC in the first-line setting and is currently the only NICE-recommended systemic therapy for HCC (24).

The dropout rate for reasons other than radiologic or clinical progression or toxicity was low and balanced between the 2 treatment arms; 2.5% of lenvatinib and 1.3% of sorafenib patients were lost to follow-up or withdrew consent.

The primary efficacy endpoint was OS (the gold standard), which is the most robust and reliable efficacy endpoint and is not subject to response assessment bias. PFS, TTP, and ORR were assessed as secondary endpoints. These endpoints are assessed prior to survival benefit and therefore are not subject to the potential confounding effect of subsequent therapy, making them useful endpoints in addition to OS (46). All efficacy endpoints selected are standard outcomes for oncology trials. Tumour assessments for secondary endpoints were performed using mRECIST criteria (as opposed to RECIST 1.1) as it most appropriately reflects changes in intrahepatic lesions (and the effects of RTK inhibitors/anti-VEGF treatment) by measuring only the viable portions of the lesions. PFS results from a post-hoc, blinded, retrospective IIR using RECIST 1.1 criteria were consistent with the investigator and post-hoc IIR assessments of PFS using mRECIST.

There were some imbalances at baseline in factors which were shown to impact treatment efficacy. Baseline AFP levels and aetiology were imbalanced in such a way that would likely have favoured the sorafenib treatment arm and generated more conservative estimates of efficacy for lenvatinib. After adjustment for baseline aetiology, the HR for OS remained numerically in favour of lenvatinib. After adjustment for baseline AFP levels, the HR for OS was nominally superior in favour of lenvatinib. These results demonstrate that despite some baseline imbalances in REFLECT which may have improved the relative performance of sorafenib, the validity of the treatment effect of lenvatinib on OS, which was non-inferior to sorafenib, was not affected.

***External validity***

The evidence base for lenvatinib reflects the proposed licensed indication and anticipated use in clinical practice in the UK.

The overall patient populations of the REFLECT study and the sorafenib SHARP study (used to support the current NICE recommendation) were similar with regard to demographic and key disease characteristics and considered to reflect UK clinical practice. Both studies enrolled patients with unresectable HCC who had no previous systemic treatment, were Child Pugh Class A (≥95% of patients) and had BCLC stage B or C disease. This is also consistent with the patient population from a published UK audit of sorafenib-treated patients (10). The proportion of sorafenib-treated patients with BCLC Stage C (advanced disease) and macroscopic portal vein invasion/extrahepatic spread was similar between the REFLECT and the SHARP trial.

The REFLECT study enrolled more male than female patients, reflective of the gender distribution observed in clinical practice in England (5) and as seen in the UK audit data (10). The proportion of patients with cirrhosis at baseline in REFLECT (50.8% in the lenvatinib arm and 48.5% in the sorafenib arm was lower than would be expected in clinical practice (approximately 70-90% (7, 8) . However, during a blinded, post-hoc independent imaging review of tumour response, the presence of cirrhosis at Baseline (based on Screening images) was assessed by the reviewers in

which approximately 75% of patients in both treatment arms had radiographic evidence of cirrhosis at Baseline, which is closer to the expected incidence in UK clinical practice.

In the REFLECT trial, 70% of patients in the lenvatinib arm and 69% of patients in the sorafenib arm were of Asian race. In total, 28% of the lenvatinib arm and 29% of patients in the sorafenib arm were White. While the proportion of patients of Asian race in REFLECT was higher than that of the UK population, and some disease characteristics, such as disease aetiology are expected to differ between western and Asian populations, feedback from a practising UK oncologist is that the aetiology of HCC does not influence clinical practice. The REFLECT study population is generally comparable to the typical population of patients with advanced HCC for most aspects and therefore the results for key outcomes (OS and PFS) in the overall population are considered to be applicable to the UK population.

Firstly, although the HR for OS (lenvatinib:sorafenib) in the Western population subgroup was higher than for the overall trial population, exceeding 1 (HR 1.08 [95% CI 0.82, 1.42]), this result can be explained by a notable imbalance in post-treatment anti-cancer therapy during survival follow-up in the Western region (lenvatinib 43.1%, sorafenib 51.1%, respectively). Such an imbalance was not observed in the Asia-Pacific region. After discontinuing therapy in REFLECT, patients in the sorafenib arm were eligible for second-line trials that enrolled patients who did not tolerate or had failed treatment with sorafenib, whereas lenvatinib-treated patients were ineligible for these trials as prior investigational agents were not permitted. Thus, this imbalance in post-treatment therapy was an artefact of the clinical trial design rather than reflecting clinical practice. When OS was adjusted for post-treatment anti-cancer therapy, the HR (95% CI) for lenvatinib:sorafenib [redacted]. In contrast, PFS was not influenced by post-progression therapies and the treatment effect was consistent between the Western region, the Asia-Pacific region and the overall population.

Secondly, the median OS of 12.3 months in the sorafenib arm of the REFLECT trial was longer than that reported for sorafenib in any of the randomised Phase III sorafenib studies to date (including SHARP), which reported OS ranging from 6.5–10.7 months (27-30, 43, 44). The median OS in the sorafenib arm of REFLECT was also longer than that reported in a UK audit of patients who received sorafenib through the Cancer Drugs Fund (median OS 9.5 months for patients with Child-Pugh Class A liver function) (10). However, in REFLECT, the TTP in the sorafenib arm was 3.7 months, which is consistent with that reported in most of these historical studies (2.8–5.5 months). These findings also suggest that the greater use of post-treatment anti-cancer therapies, especially investigational drugs that were not available to lenvatinib-treated patients, may have contributed to the longer median OS observed with sorafenib in the Western region compared with the Asia-Pacific region in REFLECT, and compared with the historical sorafenib studies. Other factors may have also contributed to the longer OS observed with sorafenib in REFLECT compared with the historic SHARP and Asia-Pacific trials. These include an improvement in supportive care for patients with advanced HCC in the 10 years since the two historic sorafenib trials were conducted, and greater familiarity among clinicians with sorafenib, leading to better outcomes.

As such, taking into account the potential impact of post-progression therapies on the OS results, particularly in the Western region subgroup, the overall trial population is considered to be most reflective of the HCC population in UK clinical practice.

### B.2.13.3 End of life

Lenvatinib meets the full criteria for an end-of-life treatment (Table 24). The current life expectancy for patients with advanced HCC is less than 1 year (21). The results of the cost-effectiveness analysis show an incremental mean overall survival benefit for lenvatinib of 3.1 months compared with sorafenib.

**Table 24: 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 | Yes | Section B.1.3 page 12 |
| 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 | Yes | Section B.3.11 page 116 |

# B.3 Cost effectiveness

## *B.3.1 Published cost-effectiveness studies*

An SLR was conducted to identify the economic implications, including the cost-effectiveness/utility and resource use, associated with treatment-naive advanced unresectable, and/or metastatic HCC. Full details of the SLR methodology are presented in Appendix G. In total, 8 records were identified which reported cost-effectiveness analyses conducted from a UK perspective and are therefore considered to be relevant to clinical practice in England. These studies are presented in Table 25.

**Table 25: Summary list of published cost-effectiveness studies relevant to clinical practice in England**

| Study | Year | Summary of model | Intervention and comparator | Patient population (average age in years) | QALYs (intervention, comparator) | Costs (currency) (intervention, comparator) | ICER (per QALY gained) |
|---|---|---|---|---|---|---|---|
| Chaplin et al, 2015 (47) | 2015 | Markov model with three health states (stable disease, progression and death), horizon 10-year (lifetime). | Y-90 vs SOR | First-line HCC (average age NR) | Y-90 1.12<br>SOR 0.85 | Y-90 £21,441<br>SOR £34,050 | Y-90 had a QALY gain of 0.27 at a lower cost, therefore dominating SOR |
| Connock et al, 2010 (48) | 2010 | Markov model: All patients start in the 1st line – no progression health state. Patients receive first-line treatment (with sorafenib or BSC) until documentation of further disease progression or until a treatment limiting AE occurs. At the point of progression, patients may either continue on first-line treatment (with sorafenib) or switch to BSC (palliative care). At any point in the model, patients may die due to all cause (general) mortality. | SOR plus BSC vs BSC alone | Advanced HCC with Child-Pugh A (average age NR) | NR | NR | Manufacturer's ICER: £64,754/QALY; ERG ICERs ranged from £76,000-£85,805/QALY.<br><br>With PAS scheme, manufacturer's ICER was £51,899/QALY and ERG ICERs ranged from £52,641-£58,147/QALY |
| Palmer et al, 2017 (49) | 2017 | A cost-minimisation model, with equal efficacy assumed between Y-90 resin microspheres and sorafenib. | Y-90 vs SOR | HCC BCLC stage C (average age NR) | Y-90 provided 0.0079 (95% CI 0.0046-0.0111) more QALYs vs SOR | Y-90 saved £8,909 (95% CI £3,257-£14,570) vs SOR | Y-90 dominated SOR |
| NICE 2017 (24) | 2017 | Model: Markov with half-cycle correction<br>Economic analysis: Cost-effectiveness analysis and cost-utility analysis | SOR vs BSC | Advanced HCC (average age NR) | Incremental QALYs:0.3588 | SOR: £30,602<br>BSC: £20,863 | £20,863/0.3588 = £58,147/QALY |

| Study | Year | Summary of model | Intervention and comparator | Patient population (average age in years) | QALYs (intervention, comparator) | Costs (currency) (intervention, comparator) | ICER (per QALY gained) |
|---|---|---|---|---|---|---|---|
| | | Health States:<br>(A) non-progressive advanced disease<br>(B) progressive disease<br>(C) BSC<br>(D) death<br>Cost perspective: UK NHS<br>Model duration: 14 years<br>Time horizon: Lifetime<br>Discount (utilities and costs): 3.5% | | | | | |
| NICE 2010 (24) | 2010 | Model: Markov with half-cycle correction<br>Economic analysis: Cost-effectiveness analysis and cost-utility analysis<br>Health States<br>(A) non-progressive advanced disease<br>(B) progressive disease<br>(C) BSC<br>(D) death<br>Cost perspective: UK NHS Model duration: 14 years<br>Time horizon: Lifetime<br>Discount (utilities and costs): 3.5% | SOR vs BSC | Advanced HCC (average age 64.9 for sorafenib patients, 66.3 for BSC patients – assumed from SHARP trial) | Base case analysis:<br>SOR 1.08<br>BSC 0.72<br><br>ERG estimates:<br>SOR 1.06-1.08<br>BSC 0.72-0.73 | Base case analysis:<br>SOR £32,971<br>BSC £9,739<br><br>ERG estimates:<br>SOR £32,971-39,910<br>BSC £9,480-9,739 | Base case ICER = £64,754/QALY<br><br>ERG estimates:<br>£64,754-85,804/QALY |
| SMC 2008 | 2008 | Model type: Markov with half-cycle correction | SOR vs BSC | Advanced HCC (average | QALYs NR | Costs NR | Cost/QALY NR<br>Cost per LYG |

| Study | Year | Summary of model | Intervention and comparator | Patient population (average age in years) | QALYs (intervention, comparator) | Costs (currency) (intervention, comparator) | ICER (per QALY gained) |
|---|---|---|---|---|---|---|---|
| (50) | | Economic analysis: Cost-effectiveness/cost-utility<br>Health states:<br>(A) non-progressive advanced disease<br>(B) progressive disease<br>(C) BSC<br>(D) death<br>Cost perspective: UK NHS<br>Model duration: 14 years<br>Time horizon: Lifetime<br>Discount (utilities and costs): 3.5% | | age NR) | | | £45,596 |
| SMC 2011 (51) | 2010 | Model type: Markov model<br>Economic analysis type: Cost-utility<br>Health states:<br>(A) PFS<br>(B) Progressed disease<br>(C) Death<br>Cost perspective: SMC NHS<br>Model duration: NR<br>Time horizon: 15 years<br>Discount (utilities and costs): NR | SOR vs BSC | Advanced HCC (average age NR) | Incremental QALY gain 0.36 | Costs NR | £67,012/QALY |

| Study | Year | Summary of model | Intervention and comparator | Patient population (average age in years) | QALYs (intervention, comparator) | Costs (currency) (intervention, comparator) | ICER (per QALY gained) |
|---|---|---|---|---|---|---|---|
| SMC 2016 (52) | 2015 | Model type: Markov model<br>Economic analysis type: cost-utility<br>Health states:<br>(A) PFS<br>(B) Progressed disease<br>(C) Death<br>Cost perspective: SMC NHS<br>Model duration: NR<br>Time horizon: 15 years<br>Discount (utilities and costs): NR | SOR vs BSC | Advanced HCC (average age NR) | Incremental QALYs: 0.367 | Incremental cost with PAS: £13,809 | ICER with PAS £37,670/QALY |

Abbreviations: AE, adverse event; BCLC, Barcelona Clinic Liver Cancer; BSC, best supportive care; CI, confidence interval; ERG, Evidence Review Group; HCC, hepatocellular carcinoma; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NHS, National Health Service; NR, not reported; PAS, Patient Access Scheme; PD, progressive disease; PFS, progression-free survival; QALYs, quality-adjusted life years; SIRT, selective internal radiation therapy; SOR, sorafenib.

## *B.3.2 Economic analysis*

None of the cost-effectiveness analyses (CEAs) identified in the economic SLR (Appendix G) included lenvatinib as a comparator. Therefore, it was necessary to include a de novo economic model in this submission. Previous published economic analyses have been used to inform the model's structure, assumptions and data sources (24).

### B.3.2.1 Patient population

The economic evaluation includes adults with untreated advanced or unresectable HCC and Child-Pugh Class A liver function as described in Table 1. It utilises data from the FAS of the pivotal REFLECT Phase III clinical trial population which included some patients with BCLC stage B (those who were considered ineligible for TACE), with most patients having BCLC stage C disease. In total, 99% of patients in REFLECT had Child-Pugh class A liver function (the remainder were Child-Pugh class B). As described in Table 1, this population is consistent with the NICE recommended population for the use of sorafenib in HCC, the SHARP study which was the evidence base for this recommendation, and with UK clinical practice as specified in the sorafenib NICE recommendation (i.e. predominantly BCLC stage C (advanced) disease, predominantly good liver function (Child-Pugh Class A) and good ECOG performance status (0–2)).

### B.3.2.2 Model structure

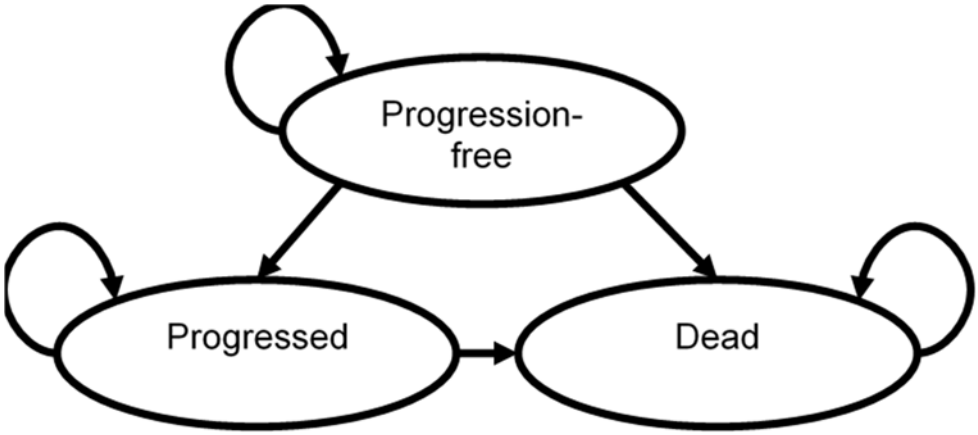
The analysis uses a partitioned survival model (PSM) developed in Microsoft® Excel to determine the cost-effectiveness of lenvatinib versus sorafenib in the treatment of HCC, over a lifetime time horizon. Partitioned survival models are commonly used in late-stage oncology modelling and have been utilised in numerous NICE single technology appraisals (STAs) and multiple technology appraisals (MTAs), including those for HCC (24). Patients must be in one of three distinct and mutually exclusive health states (Figure 8) at the end of each 28-day model cycle:

- Progression-free
- Progressed
- Dead

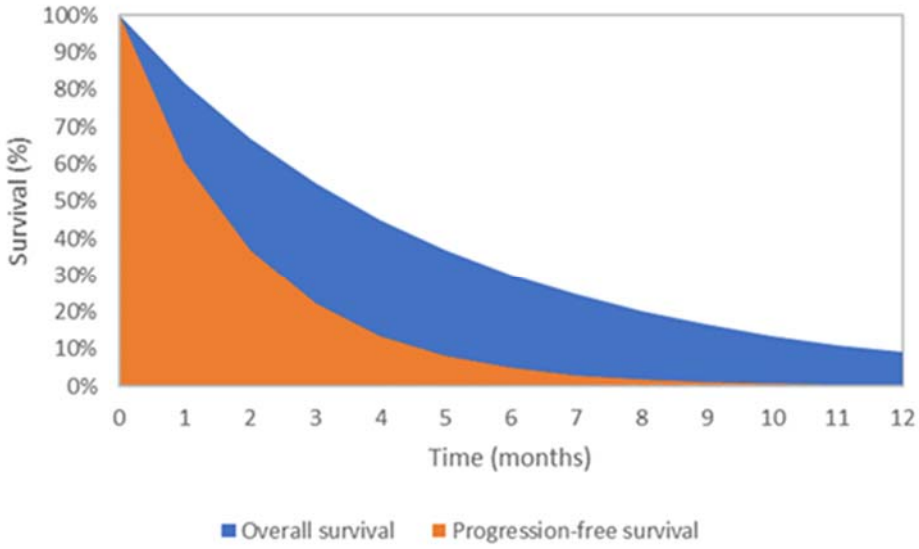
All patients enter the model in the progression-free health state and remain in this state until disease progression or death. Once in the progressed state, patients remain there until death. The model's outputs include costs incurred and quality-adjusted life-years (QALYs) gained, enabling calculation of the incremental cost-effectiveness ratio (ICER), expressed as the cost per QALY gained.

**Figure 8: Model structure**




Ovals (Figure 8) represent health states and arrows represent hypothetical transitions between the health states. However, with a PSM approach, explicit movement between health states is not modelled. Rather, the distribution of patients across all health states at each cycle is modelled (Figure 9), defined by OS and PFS curves.

**Figure 9: Example calculation of distribution of patients**



By incorporating health states for pre-progression (Progression-free), post-progression (Progressed) and death (Dead), the model captures the chronic and progressive nature of HCC and reflects the key outcomes of the pivotal Phase III lenvatinib trial (REFLECT). These are the three most relevant health states from a patient, clinician and NHS perspective, as two of the primary objectives of treating advanced HCC are prolonging life and avoiding disease progression (53).

The PSM approach is associated with advantages and disadvantages compared with multi-state transition modelling. The PSM is intuitive, easy to implement, generally predicts trial outcomes well for the within-trial period, and is well aligned with the outcomes of REFLECT (54). The disadvantages of the PSM approach are consequences of the assumption that survival outcomes are independent of each other; extrapolations for a given outcome reflect within-trial trends in that outcome alone, and the characterisation of uncertainty is more difficult (because, for example, varying parameters which define PFS does not affect survival) (54).

However, in this context, REFLECT provides relatively complete observed PFS and OS data; 64.4% of patients in the lenvatinib arm had experienced disease progression and 73.4% had died at the end of REFLECT (data cut-off 13th November 2016). The requirement for extrapolation in this instance is therefore modest, and the advantages of the PSM approach were judged to outweigh the disadvantages.

Costs incurred in the progression-free health state are those associated with active disease management. These include drug costs, costs associated with medical management of the condition, and the cost of treating grade 3–4 TEAEs. In the progressed health state, patients continue to incur costs associated with medical management in addition to palliative care costs.

Utilities are applied to each health state; patients experience a higher utility score in the progression-free state than in the progressed state.

The submission's inputs and assumptions were validated by 6 practising NHS UK clinicians, including consultant medical oncologists from the Clatterbridge Cancer

Centre and the Christie NHS Foundation Trust. These were selected based on their expertise in HCC and knowledge of UK clinical practice (see Section B.3.3.3).

Lenvatinib is associated with improved OS and PFS versus sorafenib over the lifetime of individuals (see Section B.2.6.1.1 for OS results and Section B.2.6.1.2 for PFS results from REFLECT), so these differences are modelled over a lifetime time horizon†, in line with current NICE guidance (55). A half-cycle correction has been applied. Costs and outcomes are discounted at 3.5%, in line with the NICE reference case (55).

**Table 26: Features of the economic analysis**

| Factor | Previous relevant appraisals | Current appraisal | |
|---|---|---|---|
| | TA474 (Sorafenib for advanced HCC) (24) | Chosen values | Justification |
| Time horizon | Lifetime | Lifetime | Lenvatinib is associated with improved OS and PFS versus sorafenib over the lifetime of individuals. |
| Treatment waning effect? | None | None | PFS and OS data from REFLECT are almost complete; as such, the impact of modelling a waning treatment effect would be minimal. |
| Source of clinical outcomes data | SHARP | REFLECT | REFLECT contains outcomes data for lenvatinib and sorafenib and is the most relevant source of data to inform the appraisal. |
| Source of utilities | Mapping from FACT-G to a set of time trade-off utility values using an algorithm developed by Dobrez et al (56) | REFLECT | Utility values and efficacy data are taken from the same source for consistency. Utility values in the sorafenib submission were not considered to be plausible by the appraisal committee, given that the utility value for the |

† Fewer than 1% of modelled patients remain alive at 13 years and 10 years in the lenvatinib and sorafenib arms, respectively. A time horizon of 20 years is therefore considered sufficient to capture lifetime differences between lenvatinib and sorafenib.

| | | | |
|---|---|---|---|
| | | | progressed state was higher than that for the progression-free state. |
| Source of costs | NHS reference costs (57); Personal Social Services Research Unit (58); NHS Health and Social Care Information Centre (59); Newcastle Upon Tyne 2006/07 tariffs*; Plymouth Hospital NHS Trust 2008*; UCL lab tariff 2007*; Mullhaven Medical Laboratory 2008*; British National Formulary (60) | NHS reference costs (57); Personal Social Services Research Unit (58); British National Formulary (60); Sorafenib submission to NICE (2016) (24); Nuffield Trust (61); Marie Curie Cancer Care (62) | An SLR was conducted to identify studies reporting cost and resource use data (see Appendix I). None of the identified studies were conducted in the UK and therefore they were not considered to be relevant to clinical practice in England. |
| Source of medical resource use data | Sorafenib resource use survey (63) | Sorafenib resource use survey (63) | No further sources of medical resource use data were identified |

*These references provided in the TA474 submission (24) are no longer available.

Abbreviations: FACT-G, Functional Assessment of Cancer Therapy - General; NHS, National Health Service; OS, overall survival; PFS, progression-free survival; SLR, systematic literature review; UCL, University College London.

### B.3.2.3 Intervention technology and comparators

Marketing Authorisation for lenvatinib in HCC is expected to be granted by the European Commission by June 2018. The current draft indication for lenvatinib is for the treatment of adult patients who have received no prior systemic therapy for HCC (see Table 1 and Appendix C). The comparator technology is sorafenib, which is licensed for treatment of HCC.

Lenvatinib and sorafenib are included in the model for the treatment of adult patients with untreated advanced or unresectable HCC and Child-Pugh Class A liver function, in line with the patient populations from the REFLECT and SHARP Phase III trials and with current NICE guidance for sorafenib (Table 1).

The model assumes that lenvatinib and sorafenib will be administered until progression, death, or other withdrawal in line with the protocol of the REFLECT Phase III clinical trial, the draft SmPC (Appendix C) and expected UK clinical practice.

## *B.3.3 Clinical parameters and variables*

The economic model relies on patient-level data from the REFLECT Phase III trial (see Section B.2.6.1):

- OS
- PFS
- Time to discontinuation (TTD)
- AEs

The methodology used to estimate survival curves for OS and PFS is described in Section B.3.3.1. AE data are presented in Section B.3.3.2.

### B.3.3.1 Survival analyses

#### *B.3.3.1.1 General approach*

At the data cut-off of 13th November 2016, 73.4% of patients in the lenvatinib arm and 73.5% of patients in the sorafenib arm had died. The Kaplan-Meier estimator for sorafenib PFS was 6% at the last observed data point. Therefore, there was a requirement to extrapolate beyond the end of REFLECT. This extrapolation was achieved using parametric survival analysis, performed in accordance with the relevant NICE Decision Support Unit (DSU) guidance (64). The Kaplan-Meier estimators for TTD were almost complete (0% and 4% for lenvatinib and sorafenib, respectively). The Kaplan-Meier curves were therefore used directly to inform TTD and extrapolation was not performed for this outcome‡.

As described in Section B.2.3.2, there were baseline imbalances between the lenvatinib and sorafenib treatment arms regarding the proportion of patients with AFP levels ≥200 ng/mL, and in the aetiology of HCC (HBV, HCV, alcohol). Of principal concern was the imbalance in baseline AFP levels between treatment arms; AFP has been demonstrated to be a strong independent predictor of outcomes regardless of treatment type (16), and the proportion of patients with AFP levels

‡ For simplicity, it was conservatively assumed that the remaining 4% of patients receiving sorafenib treatment discontinued treatment at the end of trial follow-up.

≥200 ng/mL was higher in the lenvatinib arm (46.4%) than the sorafenib arm (39.3%). Covariate analyses were performed to evaluate baseline factors that may have impacted OS in the overall study population, including AFP and HCC aetiology; when adjusting for baseline AFP, lenvatinib was nominally superior to sorafenib (median OS 13.6 months vs 12.3 months, respectively; HR 0.856 with upper limit of the 95% CI <1 [95% CI:0.736, 0.995], p=0.0342).

- **Base-case and scenario analyses adjust for these imbalances in baseline characteristics**. Current EMA guidance on adjustment for baseline characteristics in clinical trials suggests that in the presence of imbalances for strong predictors of outcomes, adjustment for such covariates generally improves the precision and efficiency of the analysis and avoids conditional bias from chance covariate imbalance (65).
- The base-case analysis was based on multivariable adjustments to the PFS and OS curves. In order to explore the impact of these multivariable adjustments on the results of the economic evaluation, **a scenario analysis is presented based on unadjusted parametric models** (based on an ITT approach).
- **A further scenario analysis is presented which adjusts for AFP only (in addition to stratification factors)**. This is intended to provide an intermediary scenario to the multivariable (base-case) and unadjusted (scenario) analyses.

In order to perform extrapolation, a three-stage process was followed:

1. Assessment of the proportional hazards (PH) assumption (Section B.3.3.1.1.1)
2. Identification of prognostic factors upon which to base adjustment (Section B.3.3.1.2)
3. Estimation of parametric survival models to allow prediction of event rates (Section B.3.3.1.3)

#### B.3.3.1.1.1 Proportional hazards assessment

The validity of the PH assumption between treatments was assessed. This was tested using:

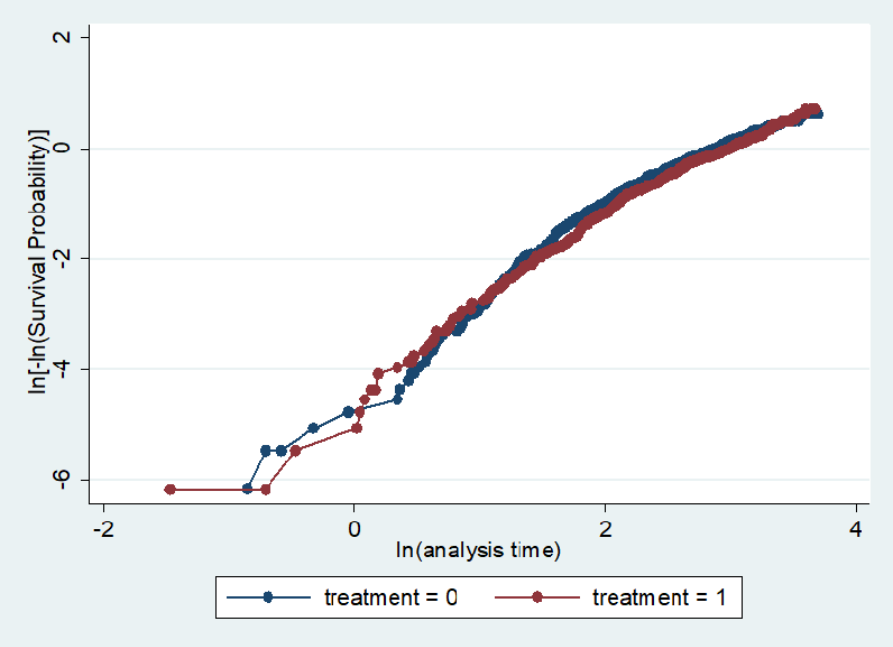

 

- Visual inspection of the log-cumulative hazard plots, and
- PH global test (Schoenfeld residual test)

#### B.3.3.1.1.2 Overall survival

The log-cumulative hazard plot for OS is presented in Figure 10.

**Figure 10: Log-cumulative hazard plot – OS data**




Treatment = 0 (sorafenib); treatment = 1 (lenvatinib).
Abbreviations: OS, overall survival.

The PH global test yielded a p-value of 0.2902, which indicates no statistical deviation to the PH assumption. The Kaplan-Meier curve stays generally parallel (see Section B.2.6.1.1). The visual inspection of the hazard plot shows that the lines converge and cross at the beginning and end but remain relatively straight (with a slight curve) and parallel.

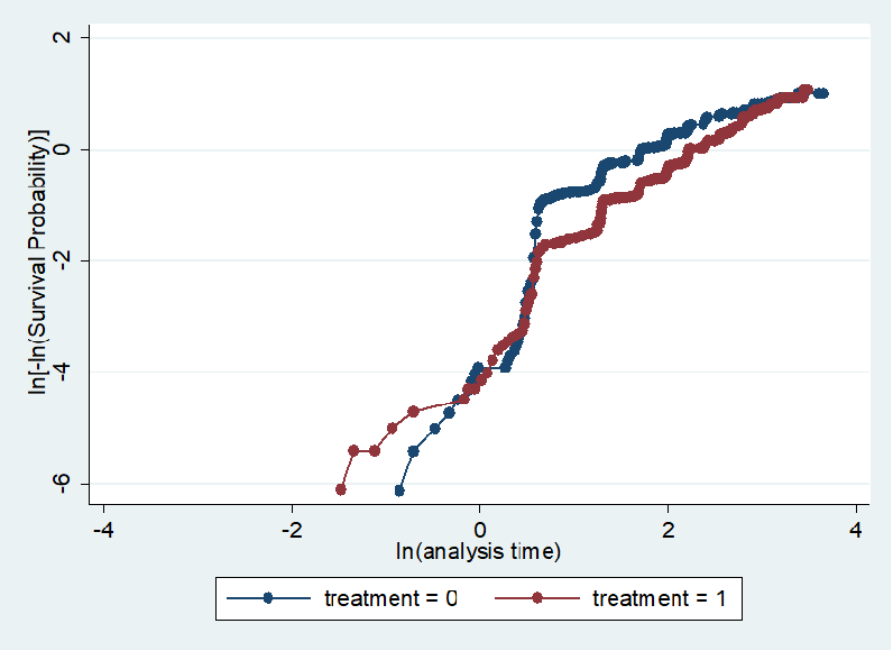
According to NICE DSU TSD 14 (64), pooled PH models with a treatment covariate should be used when the cumulative hazard plots of two (or more) arms are parallel and straight. In the context of this trial, cumulative hazard plots are not completely parallel, nor are they completely straight. Because of this, and the observation that the curves cross at the beginning of the study, further analyses are based on

independent statistical models for each arm. This also provides consistency with the modelling of PFS (see Section B.3.3.1.1.3).

#### B.3.3.1.1.3 Progression-free survival

The log-cumulative hazard plot for PFS is presented in Figure 11.

**Figure 11: Log-cumulative hazard plot – PFS data**




Treatment = 0 (sorafenib); treatment = 1 (lenvatinib).
Abbreviations: PFS, progression-free survival.

The PH global test yielded a p-value <0.0001, which indicates a statistical deviation to the PH assumption. The visual inspection of the hazard plot shows that the lines cross at the start and merge towards the end of the study. This suggests a change in the trend of the hazard. In the absence of support for the PH assumption further analyses were based on independent statistical models for each arm.

#### *B.3.3.1.2 Identification of prognostic factors*

[redacted]

[redacted]

Post-treatment anti-cancer therapy, while a sub-group, was not included as a candidate variable due to the inherent lack of randomisation. However, as stated in Section B.2.6.1.1.3, adjusting for post-treatment anti-cancer therapy does have a downward effect on the OS HR.

[redacted]

[redacted]). This finding is consistent with the results of the pre-planned supportive analyses adjusting for imbalanced baseline characteristics (see Section B.2.6.1.1.2), which demonstrate that lenvatinib was nominally superior to sorafenib when adjusted for AFP (HR 0.856 [95% CI: 0.736, 0.995], p=0.0342). Please note however that this statistical model, and hazard ratio, is not used directly in the economic evaluation, which is based on multivariable parametric models estimated separately for each arm (Sections B.3.3.1.1.1 and B.3.3.1.3).

Multivariable parametric models were used to generate predictions for outcomes based on the REFLECT study population. Specifically, the analysis uses the mean of covariates method to make predictions.

Under the mean of covariates method, the mean value of each covariate used in the prediction equation, obtained from a standard multivariable regression, is used to predict outcomes for an average patient.

For example, consider the log-logistic distribution, with survivor function given by:

$$S(t) = \left\{1 + \left(\exp(-\boldsymbol{x_j}\boldsymbol{\beta}\right)t^{1/\gamma}\right\}^{-1}$$

Where $t$ is time, $\gamma$ is the scale parameter, $\boldsymbol{x_j}$ is a vector of covariates and $\boldsymbol{\beta}$ is a vector of model coefficients. In order to implement the mean of covariates method, each covariate in the vector of covariates is set equal to their respective mean value (i.e. $\bar{\boldsymbol{x}}_{\boldsymbol{j}}$), such that predicted survival becomes:

$$S_{MoC}(t) = \left\{1 + \left(\exp(-\bar{\boldsymbol{x}}_{\boldsymbol{j}}\boldsymbol{\beta}\right)t^{1/\gamma}\right\}^{-1}$$

The resulting predicted survival is representative of an average patient. In non-linear models, the limitation of this approach is that the predicted survival for the average patient is not necessarily the same as the average predicted survival across all patients. Nevertheless, this method is transparent to implement and allows for straightforward characterisation of uncertainty in probabilistic sensitivity analysis. Assessment of uncertainty is more difficult in other approaches such as the corrected group prognosis (in which a survival curve is calculated for each level of covariates, after which an average survival curve is calculated as a weighted average of the survival curves for each level of covariates).

#### *B.3.3.1.3 Choice of distribution*

Statistical analyses of survival data were performed in accordance with NICE DSU TSD 14 (64) where possible. All analyses described below were based on patient-level data from REFLECT.

Six distributions were investigated for all outcomes: Weibull, exponential, log-logistic, log-normal, gamma and Gompertz.

The most appropriate distribution was selected based on (a) assessment of the statistical goodness of fit (measured using the Akaike Information Criteria [AIC] and Bayesian Information Criteria [BIC]) and (b) consistency with previous findings of extrapolation methods in advanced HCC. Statistical goodness of fit was considered more relevant to distribution choice for REFLECT due to the maturity of data than in

cases where follow-up is less complete. Visual inspection was not used, as after adjustment for multiple covariates, comparison to the unadjusted Kaplan-Meier curve was not considered an appropriate method for assessing goodness-of-fit.

During the 2009 appraisal of sorafenib in advanced HCC, the log-normal curve was accepted by the Appraisal Committee and was used in the manufacturer's base case for OS. During the appraisal, it was commented that the Weibull curve might also provide an acceptable fit to the data. During the 2016 reconsideration, two pieces of real-world data were presented by the manufacturer. Both data sources show that a small proportion of patients survive for an extended period of time, indicating that the Weibull curve does not fit with the survival observed in clinical practice and that the log-normal curve is indeed a better fit (24).

For the base case analysis (i.e. adjusting for baseline imbalances), model fit statistics (AIC/BIC) for the lenvatinib and sorafenib data suggested that the log-logistic distribution was preferred for OS, with log-normal and gamma distributions associated with similarly good fits to the observed data (

Table 27 and

Table 28). The use of the log-logistic distribution is considered to be consistent with the findings of the 2016 sorafenib reconsideration, as for certain ranges of the parameters, the shape of the log-normal and log-logistic hazard functions can be very similar in nature (66). Alternative distributions are presented in scenario analyses.

**Table 27: AIC and BIC: Lenvatinib, OS (multivariable analyses)**

| Model | Obs | df | AIC | BIC | Mean OS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; ll, log likelihood; Obs, observations; OS, overall survival.

**Table 28: AIC and BIC: Sorafenib, OS (multivariable analyses)**

| Model | Obs | df | AIC | BIC | Mean OS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; ll, log likelihood; Obs, observations; OS, overall survival.

For PFS, the statistical measures indicated that the log-normal was preferred for lenvatinib, with gamma and log-logistic similarly well performing (Table 29). For the sorafenib arm, the gamma distribution was preferred (Table 30). However, use of the gamma distribution led to implausible extrapolations in which PFS for sorafenib exceeded that of lenvatinib. This was not considered a clinically plausible scenario, and therefore the gamma distribution was not considered further. The log-normal distribution was the next best fitting distribution and provided clinically plausible results and was therefore used in the base-case.

**Table 29: AIC and BIC: Lenvatinib, PFS (multivariable analyses)**

| Model | Obs | df | AIC | BIC | Mean PFS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; ll, log likelihood; Obs, observations; PFS, progression-free survival.

**Table 30: AIC and BIC: Sorafenib, PFS (multivariable analyses)**

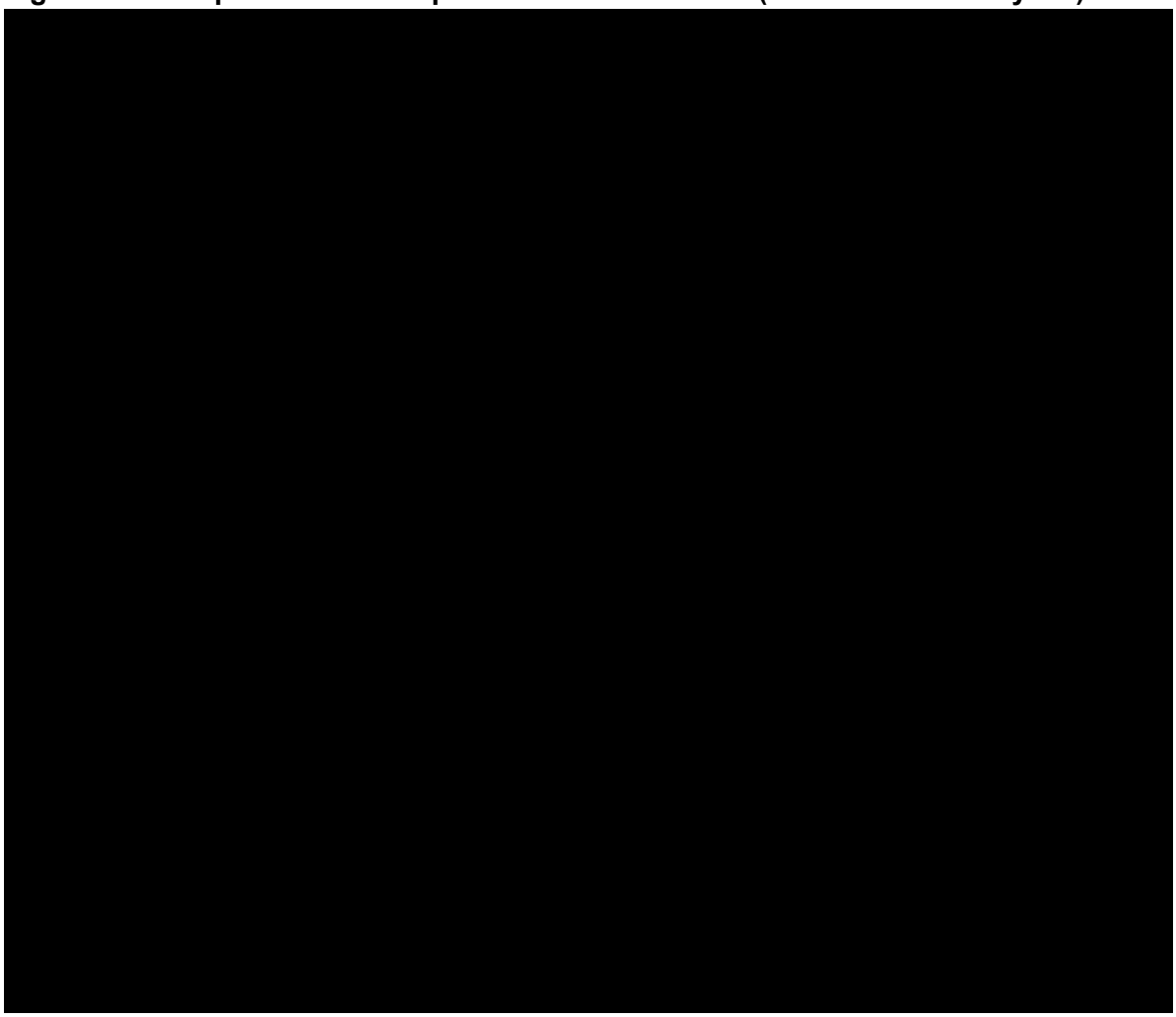
| Model | Obs | df | AIC | BIC | Mean PFS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; ll, log likelihood; Obs, observations; PFS, progression-free survival.

Alternative extrapolations for each model outcome are presented in Figure 12 and Figure 13. In general, the differences in extrapolated outcomes between distributions were modest, a consequence of the relatively mature data.

**Figure 12: Comparison of extrapolations for lenvatinib (multivariable analyses)**

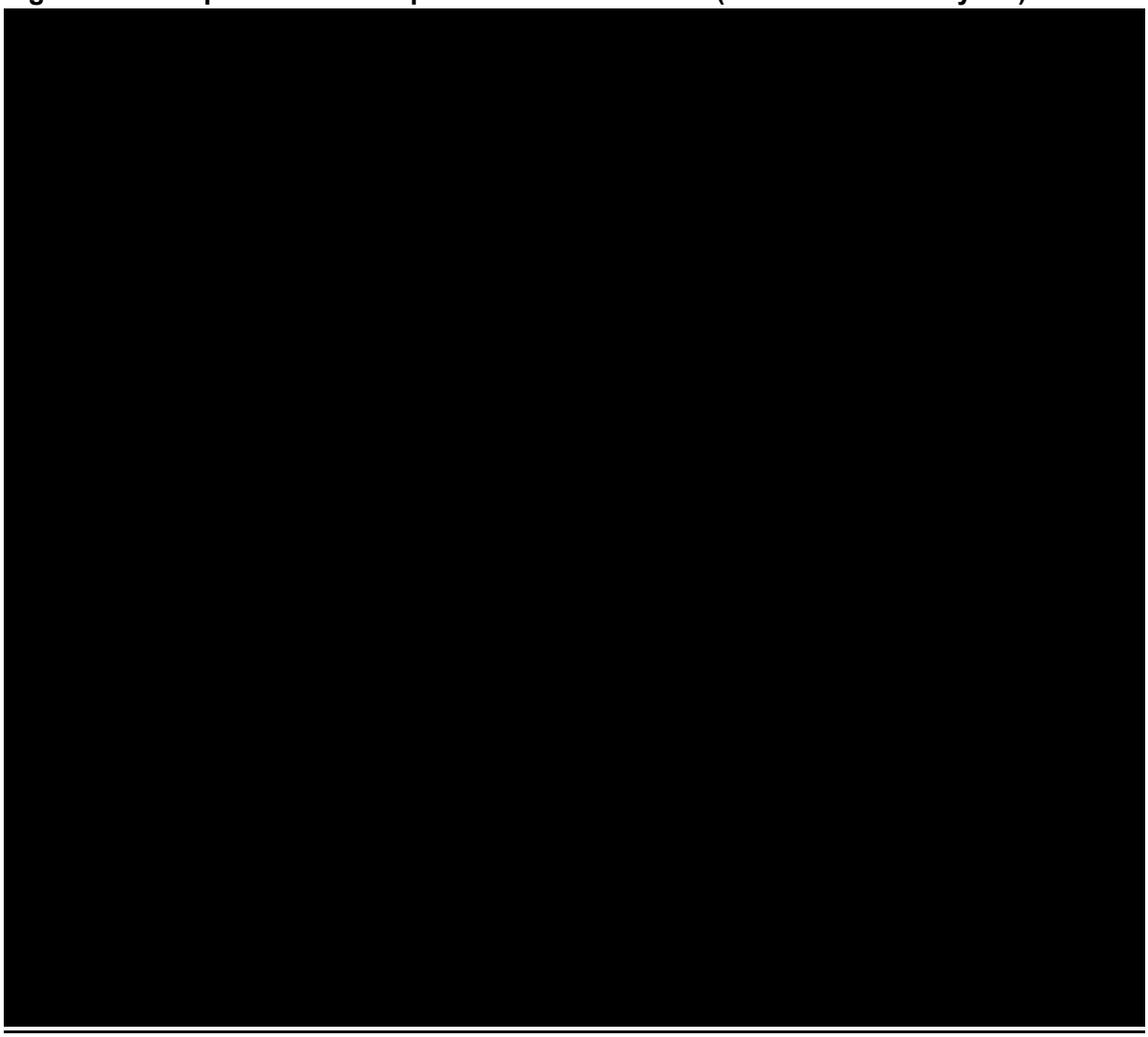


Abbreviations: OS, overall survival; PFS, progression-free survival.



**Figure 13: Comparison of extrapolations for sorafenib (multivariable analyses)**



Abbreviations: OS, overall survival; PFS, progression-free survival.

Final statistical models used in the base-case are presented in Appendix O. As described in Section B.3.3.1.1, scenario analysis considers use of the unadjusted analysis (in addition to analysis adjusting for AFP and stratification factors only). For completeness, the predicted vs observed curves and model fit statistics based on an unadjusted analysis are presented in Appendix P. The model fit statistics from the unadjusted analysis were consistent with the preferred models based on the multivariable analysis.

### B.3.3.2 Adverse events

AEs considered in the economic model include Grade 3 or 4 TEAEs occurring in ≥5% of patients in either treatment arm of REFLECT (see Table 31). Additional Grade 3 or 4 TEAEs that occurred in <5% of patients in either treatment arm were included if identified as being clinically or economically significant by UK clinical experts (see Section B.3.3.3). The additional TEAEs identified by the experts were diarrhoea, asthenia and fatigue.

**Table 31: List of AEs**

| AEs | Number of patients, n (%) | | Average number of episodes per patient | |
|---|---|---|---|---|
| | Lenvatinib | Sorafenib | Lenvatinib | Sorafenib |
| Hypertension | 111 (23.3%) | 68 (14.3%) | 1.12 | 1.09 |
| Weight decreased | 36 (7.6%) | 14 (2.9%) | 1.03 | 1.00 |
| Blood bilirubin increased | 31 (6.5%) | 23 (4.8%) | 1.10 | 1.04 |
| Proteinuria | 27 (5.7%) | 8 (1.7%) | 1.04 | 1.00 |
| Gamma-glutamyltransferase increased | 26 (5.5%) | 19 (4.0%) | 1.04 | 1.05 |
| Platelet count decreased | 26 (5.5%) | 16 (3.4%) | 1.12 | 1.00 |
| Aspartate aminotransferase increased | 24 (5.0%) | 38 (8.0%) | 1.04 | 1.11 |
| Diarrhoea | 20 (4.2%) | 20 (4.2%) | 1.20 | 1.20 |
| Fatigue† | 18 (3.8%) | 17 (3.6%) | 1.07 | 1.00 |
| Palmar-plantar erthrodysaesthesia syndrome | 14 (2.9%) | 54 (11.4%) | 1.07 | 1.17 |
| Asthenia | 14 (2.9%) | 11 (2.3%) | 1.07 | 1.00 |

†Average number of episodes per patient for fatigue is based on asthenia.
Abbreviations: AE, adverse event; CI, confidence interval.

### B.3.3.3 Clinical expert opinion

The submission inputs and assumptions were validated by 6 practising NHS UK clinicians, including consultant medical oncologists from the Clatterbridge Cancer Centre and the Christie NHS Foundation Trust. These were selected based on their expertise in HCC and knowledge of UK clinical practice. Three of the experts were investigators in the REFLECT trial and had at least one patient in the lenvatinib arm.

For selection of clinically significant covariates (see Section B.3.3.1.2), 5 clinical experts participated (two of these were UK clinicians involved in the validation described above). The background information provided, and interview guide used, is provided in Appendix M.

## *B.3.4 Measurement and valuation of health effects*

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

EQ-5D-3L data were collected in the REFLECT clinical trial. Patients completed the questionnaire at the Baseline visit, on Day 1 of each subsequent treatment cycle, and at the Off-Treatment visit (see Section B.2.3.1.9). For the economic analysis, the UK EQ-5D-3L value set was applied to responses at each time point (67).

The UK EQ-5D-3L value set was developed using data collected from 2,997 general population respondents using a time-trade-off methodology. The EQ-5D is NICE's preferred measure of HRQoL, and therefore use of EQ-5D data from REFLECT to inform the economic analysis is consistent with the NICE reference case (55).

EQ-5D-3L data from REFLECT were analysed to generate mean utility values at baseline, in the progression-free health state and in the progressed health state (Table 32). Additional analyses were conducted based on the lenvatinib and sorafenib arms of REFLECT separately to determine:

- Mean utility value at baseline and in the progression-free and progressed health states
- Adjusted mean utility value at baseline and in the progression-free and progressed health states, controlling for prior treatment, age, sex, geographical region, baseline EQ-5D and baseline ECOG-PS; adjustment was performed using a linear mixed model

Adjusted mean utility values for the progression-free and progressed health states were similar between the lenvatinib and sorafenib arms, with a small numerical difference in favour of lenvatinib. It is therefore conservatively assumed that utility

values in the lenvatinib and sorafenib arms are equal to the mean values in the full REFLECT population.

**Table 32: Summary of EQ-5D**

| | | Full REFLECT population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| Baseline | N | 921 | 463 | 458 |
| | Mean (SE) | 0.829 (0.0067) | 0.823 (0.0101) | 0.836 (0.0088) |
| | Adjusted mean (SE) | N/A | 0.784 (0.0113) | 0.798 (0.0113) |
| Progression-free | N | 852 | 421 | 431 |
| | Mean (SE) | 0.745 (0.0079) | 0.750 (0.0105) | 0.740 (0.0118) |
| | Adjusted mean (SE) | N/A | 0.745 (0.0116) | 0.737 (0.0115) |
| Progressed | N | 755 | 373 | 382 |
| | Mean (SE) | 0.678 (0.0118) | 0.678 (0.0163) | 0.679 (0.0170) |
| | Adjusted mean (SE) | N/A | 0.665 (0.0188) | 0.656 (0.0185) |

Abbreviations: N/A, not applicable; SE, standard error.

### B.3.4.2 Mapping

EQ-5D-3L data were collected in REFLECT, therefore no mapping was required.

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

A SLR was conducted to identify studies reporting on the HRQoL of patients with advanced/unresectable/metastatic HCC. Full details of the methodology and results of included studies are presented in Appendix H.

### B.3.4.4 Adverse reactions

Disutilities associated with AEs were not explicitly modelled. Given that adjusted mean utility values for lenvatinib and sorafenib are similar (with lenvatinib associated with a small numerical increase), it is not considered that the AE profiles for lenvatinib and sorafenib result in differing utility values, despite the expected differences in impact of the AE profiles on quality of life.

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

The utility value for the progression-free health state was estimated to be 0.745 (95% CI: 0.730, 0.760). Utility in the progressed health state was 0.678 (95% CI: 0.701, 0.655) (Table 33).

**Table 33: Summary of utility values for cost-effectiveness analysis**

| Health state | Utility value: mean (standard error) | Reference in submission (section and page number) | Justification |
|---|---|---|---|
| Progression-free | 0.745 (0.0079) | Section B.3.4.1, page 91 | EQ-5D-3L from REFLECT was considered to be the most appropriate source of HRQoL data given the consistency with the reference case and applicability of the patient population. |
| Progressed | 0.678 (0.0118) | Section B.3.4.1, page 91 | |

Abbreviations: EQ-5D-3L, EuroQol-5 Dimensions (3 level version).

The relatively small difference between progression-free and progressed utility values is believed to reflect the data collection schedule of REFLECT and the proximity of the post-progression measurement to progression itself. Clinical experts believed this estimate of post-progression HRQoL to be higher than might be expected, as HCC has a significant impact on the patient's functioning and well-being (68). Nevertheless, the EQ-5D-3L from REFLECT was considered to be the most appropriate source of HRQoL data given the consistency with the reference case and applicability of the patient population.

HRQoL was assumed to be constant within each health state, but as more patients moved into the progressed health state, the average HRQoL experienced by patients decreased over time. Given the challenges associated with post-progression utility values outlined above, the analysis may fail to adequately capture the reduced HRQoL experienced post-progression, particularly towards the end of life. Given that lenvatinib is associated with delayed progression, it may be expected that the model is not able to capture the full HRQoL benefit associated with lenvatinib. Alternative

sources of HRQoL (Table 34) were used in scenario analysis (scenario 22 in Table 44).

**Table 34: Alternative sources of utility values**

| State | Committee preference in NICE appraisal of sorafenib (24) | NICE submission for regorafenib (69) | Hypothetical scenario with lower post-progression utility |
|---|---|---|---|
| Progression-free | 0.760 | 0.797 | 0.745 |
| Progressed | 0.680 | 0.749 | 0.500 |

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

Details of the systematic identification of studies reporting cost and resource use data are presented in Appendix I. In total, 19 studies were identified which reported cost or resource use data relating to the management of advanced HCC. None of these studies were conducted in the UK and were therefore not considered to be relevant to clinical practice in England.

Parameters used in the economic evaluation are presented in Appendix N.

### B.3.5.1 Intervention and comparators' costs and resource use

#### *B.3.5.1.1 Technology costs*

Technology costs were taken from the British National Formulary (60) for sorafenib. The cost used for lenvatinib is the approved patient access scheme (PAS) price. Table 35 presents the costs of each product based on these prices and mean dose observed in REFLECT. This approach was consistent with the methods requested by the BMJ Evidence Review Group during the previous appraisal of lenvatinib in renal cell carcinoma (RCC) (70).

Total drug cost is calculated based on the mean dose observed in REFLECT. In the base-case, the costs of drug wastage were not included. A scenario is presented in which discontinuation is associated with wastage of 7 days' worth of drug costs.

The final cost per cycle of lenvatinib was [redacted] with the application of the agreed PAS. The corresponding cost per cycle for sorafenib was £2,968.

Drug costs were applied to the proportion of patients remaining on treatment in each model cycle, based on Kaplan-Meier curves generated using patient-level data from REFLECT. At the end of trial follow-up, all non-censored patients in the lenvatinib arm had discontinued, and 4% of non-censored patients in the sorafenib arm remained on treatment. It was conservatively assumed that the proportion of patients receiving sorafenib dropped to zero after the end of trial follow-up.

**Table 35: Drug dosing (per day)**

| Drug | Dose per pill (mg) | Price per pill | Mean dose (mg)* | Cost per day |
|---|---|---|---|---|
| LEN (PAS) | 4 | [redacted] | 9.4 | [redacted] |
| SOR | 200 | £31.93 | 663.8 | £105.99 |

*Source: mean measure LEN and SOR - e7080-g000-304 Clinical Study Report Table 31, Eisai, 2017.
Abbreviations: LEN, lenvatinib; PAS, patient access scheme; SOR, sorafenib.

#### *B.3.5.1.2 Administration*

Lenvatinib treatment should be initiated and supervised by a health care professional experienced in the use of anti-cancer therapies (Appendix C). In the model this was assumed to occur in an outpatient setting for both products, with the cost captured as part of background medical management (see Section B.3.5.2). After initiation it was assumed that both lenvatinib and sorafenib would not be associated with administration costs and that patients would orally self-administer both products.

#### *B.3.5.1.3 Monitoring*

Both sorafenib and lenvatinib are associated with additional monitoring requirements. For lenvatinib (Appendix C), specified monitoring activities include the following:

- Blood pressure should be monitored after 1 week of treatment with lenvatinib, then every 2 weeks for the first 2 months, and monthly thereafter.
- Urine protein should be monitored regularly.
- Periodic monitoring of electrocardiogram and electrolytes (magnesium, potassium and calcium) should be considered during treatment.

- Patients should be monitored for clinical symptoms or signs of cardiac decompensation, as dose interruptions, adjustments, or discontinuation may be necessary.
- Liver function tests should be monitored before initiation of treatment, then every 2 weeks for the first 2 months, and monthly thereafter during treatment.
- Thyroid function should be monitored before initiation of, and periodically throughout, treatment with lenvatinib.

For sorafenib (71), these monitoring activities include:

- Regular blood pressure monitoring
- Further monitoring is advised for special populations:
    - The use of caution is advised when using sorafenib in patients who have, or may develop, prolongation of QTc. When using sorafenib in these patients periodic monitoring with on-treatment electrocardiograms and electrolytes (magnesium, potassium, calcium) should be considered.
    - Patients taking concomitant warfarin or phenprocoumon should be monitored regularly.
    - Monitoring of renal function should be considered in elderly people.

A clinical expert suggested that in practice, monitoring requirements for both products would be the same. The Assessment Group in the recent MTA of lenvatinib and sorafenib in the treatment of differentiated thyroid cancer assumed that both products would be associated with the same monitoring and testing costs whilst on treatment (72). Given this evidence, and in the absence of any product-specific data, the same monitoring costs were assumed for both sorafenib and lenvatinib. These monitoring and testing activities are combined with other resource use and are detailed in Section B.3.5.1.4.

#### *B.3.5.1.4 Summary of intervention and comparator costs*

**Table 36: Unit costs associated with the technology in the economic model**

| Items | Lenvatinib (with PAS) | Sorafenib | Reference in submission |
|---|---|---|---|
| Technology cost | [redacted] per pack of 30 x 4 mg capsules | £3576.56 per pack of 112 x 200mg tablets (73) | Section B.3.5.1.1 |
| Mean cost of technology treatment/28 days | [redacted] | £2,968 | Section B.3.5.1.1 |
| Administration cost | £0 | £0 | Section B.3.5.1.2 |
| Monitoring cost* | £0 | £0 | Sections B.3.5.1.3 and B.3.5.2 |
| Tests* | £0 | £0 | Section B.3.5.2 |
| **Total** | [redacted] | **£2,968** | - |

*Costs for monitoring and testing are incorporated in health state costs (see Section B.3.5.2).
Abbreviations: PAS, patient access scheme.

### B.3.5.2 Health-state unit costs and resource use

In the absence of evidence from either REFLECT or the published literature, estimates of resource use in the progression-free and progressed health states were based on a resource use survey commissioned by the manufacturer of sorafenib, presented in TA189 and updated in TA474 (24). During the 2016 reconsideration of sorafenib, the Appraisal Committee expressed a preference for the pooling of the original and updated results of the resource use survey. Due to differences in data reported in the two publicly available manufacturer's submissions, a two-stage approach was adopted in order to achieve this:

- For the original survey results: new unit costs were applied to the presented resource use estimates.
- For the updated survey results: as resource use was not provided directly, the 2015/16 costs presented were uplifted to 2016/17 using the Hospital and Community Health Services (HCHS) Pay and Prices Index (58).

A weighted average cost based on the number of clinicians responding to each survey was then calculated. Unit costs are presented in Table 37 and weighted average costs are presented in Table 38.

Unit costs were taken from published national sources including NHS reference costs 2016 to 2017 (57) and Personal Social Services Research Unit (PSSRU) Unit Costs of Health and Social Care 2017 (58). The cost of hospice care was taken from a report by Marie Curie Cancer Care (62) and inflated from 2003/04 to 2016/17 prices using the HCHS Pay and Prices Index (58).

**Table 37: Health states and associated unit costs in the economic model**

| | Unit cost | Reference |
|---|---|---|
| **Physician visits** | | |
| Appointment with oncologist | £172.67 | NHS reference costs 2016/17. WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology) (57) |
| Appointment with hepatologist | £216.66 | NHS reference costs 2016/17. WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (hepatology) (57) |
| Appointment with Macmillan nurse | £42.00 | PSSRU, Unit Costs of Health and Social Care 2017. Nurse (GP practice). Cost per hour, including qualifications (58) |
| Appointment with gastroenterologist | £140.80 | NHS reference costs 2016/17. WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (gastroenterology) (57) |
| Appointment with radiologist | £73.88 | NHS reference costs 2016/17. WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (interventional radiology) (57) |
| Appointment with clinical nurse specialist | £42.00 | PSSRU, Unit Costs of Health and Social Care 2017. Nurse (GP practice). Cost per hour, including qualifications (58) |
| Appointment with palliative care physician/nurse | £42.00 | PSSRU, Unit Costs of Health and Social Care 2017. Nurse (GP practice). Cost per hour, including qualifications (58) |
| **Radiological tests** | | |
| Abdominal CT | £101.57 | NHS reference costs 2016/17. Average of all computerised tomography currency codes (adult only), weighted by activity (RD20A, RD21A, RD22Z, RD23Z, RD24Z, RD25Z, RD26Z, RD27Z, RD28Z) (57) |
| Abdominal MRI | £144.88 | NHS reference costs 2016/17. Average of all magnetic resonance imaging currency codes (adult only, excluding cardiac magnetic resonance imaging), weighted by activity (RD01A, RD02A, RD03Z, RD04Z, RD05Z, |

| | Unit cost | Reference |
|---|---|---|
| | | RD06Z, RD07Z) (57) |
| **Laboratory tests** | | |
| AFP test | £25.02 | NHS reference costs 2016/17. Average of DAPS01 and DAPS02 (cytology, and histopathology and histology) (57) |
| Liver function test | £25.02 | NHS reference costs 2016/17. Average of DAPS01 and DAPS02 (cytology, and histopathology and histology) (57) |
| INR | £2.60 | NHS reference costs 2016/17. Average of DAPS03, DAPS05 and DAPS08 (integrated blood services, haematology and phlebotomy) (57) |
| Complete blood count | £2.60 | NHS reference costs 2016/17. Average of DAPS03, DAPS05 and DAPS08 (integrated blood services, haematology and phlebotomy) (57) |
| Biochemistry | £1.13 | NHS reference costs 2016/17. DAPS04 (clinical biochemistry) (57) |
| Endoscopy | £499.80 | NHS reference costs 2016/17. FE50A (Wireless Capsule Endoscopy, 19 years and over). Outpatient procedures (57) |
| **Hospitalisation** | | |
| Hospitalisation | £1,924.44 | NHS reference costs 2016/17. Average of all hospitalisations for malignant, hepatobiliary or pancreatic disorders, weighted by activity (GC12C, GC12D, GC12E, GC12F, GC12G, GC12H, GC12J, GC12K) (57) |
| **Hospital follow-up** | | |
| Specialist | £216.66 | NHS reference costs 2016/17. WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (hepatology) (57) |
| GP | £37.00 | PSSRU, Unit Costs of Health and Social Care 2017. General practitioner, cost per surgery consultation lasting 9.22 minutes (including direct care staff costs, with qualification costs) (58) |
| Nurse | £42.00 | PSSRU, Unit Costs of Health and Social Care 2017. Nurse (GP practice). Cost per hour, including qualifications (58) |
| **Social care** | | |
| Residential care (daily cost) | £110.65 | PSSRU, Unit Costs of Health and Social Care 2017. Average of two sources: private sector residential care for older people (£632 per week = £90.29 per day) and private sector residential care homes for adults requiring physical support (£131 per day) (58) |
| Day care (daily cost) | £91.00 | PSSRU, Unit Costs of Health and Social Care 2017. Day care for adults requiring physical support (58) |
| Home care (daily cost) | £164.56 | PSSRU, Unit Costs of Health and Social Care 2017. Home care for adults requiring physical support. Average of two hourly rates: £25.62 (services provided in-house) |

| | Unit cost | Reference |
|---|---|---|
| | | and £15.52 (provision by external providers). Assumes 8 hours per day (58) |
| Hospice (daily cost) | £502.94 | Marie Curie Cancer Care – Understanding the cost of end of life care in different settings (62) |

Abbreviations: CT, computerised tomography; GP, general practitioner; INR, international normalised ratio; MRI, magnetic resonance imaging; NHS, National Health Service; PSSRU, Personal Social Services Research Unit.

**Table 38: Per-cycle health state costs (weighted averages)**

| | Progression-free | Progressed |
|---|---|---|
| Physician visits | £159.63 | £384.40 |
| Laboratory tests | £161.78 | £135.56 |
| Radiological tests | £30.04 | £27.25 |
| Hospitalisation | £91.52 | £196.78 |
| Hospital follow-up† | £168.50 | £726.26 |
| Social care† | £21.19 | £1,066.07 |
| **Total** | **£632.67** | **£2,536.32** |

†Not reported in the 2016 reconsideration of sorafenib by NICE; these values are therefore based only on the survey results presented in the original sorafenib submission to NICE.

### B.3.5.3 Adverse reaction unit costs and resource use

AEs considered in the economic model include grade 3 or 4 TEAEs occurring in ≥5% of patients in either treatment arm of REFLECT. Additional Grade 3 or 4 TEAEs that occurred in <5% of patients in either treatment arm were included if identified as being clinically or economically significant by UK clinical experts (see Section B.3.3.3).

Resource use for asthenia, diarrhoea, fatigue, hypertension, palmar-plantar erythrodysaesthesia syndrome, proteinuria and weight decreased are as used in the manufacturer's submission to NICE for lenvatinib in the treatment of differentiated thyroid cancer (ID1059) (72). In that submission adverse event costs were informed by four practising NHS clinicians who provided input at an advisory board. Resource use associated with the other four adverse reactions (aspartate aminotransferase increased; blood bilirubin increased; gamma-glutamyl transferase increased; and platelet count decreased) was estimated by a UK clinician.

Given the lenvatinib and sorafenib AE profiles (see Section B.2.10), drug costs of treating AEs were considered negligible and therefore not included in the analysis.

**Table 39: List of adverse reactions and summary of costs in the economic model**

| Adverse reactions | Items | Unit cost | Reference |
|---|---|---|---|
| Aspartate aminotransferase increased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| Asthenia | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| | Nurse visit | £42.00 | PSSRU 2017 – Nurse (GP practice) – cost per hour, including qualifications (58) |
| | **Total** | £659.11 | - |
| Blood bilirubin increased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| | Outpatient contact | £172.67 | NHS reference costs 2016/17 – WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology) (57) |
| | CT scan | £101.57 | NHS reference costs 2016/17. Average of all computerised tomography currency codes (adult only), weighted by activity (RD20A, RD21A, RD22Z, RD23Z, RD24Z, RD25Z, RD26Z, RD27Z, RD28Z) (57) |
| | **Total** | £891.35 | - |
| Diarrhoea | Hospitalisation | £588.54 | NHS reference costs 2016/17 – FZ91K Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 6-10 – non-elective short-stay (57) |
| Fatigue | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| | Nurse visit | £42.00 | PSSRU 2017 – Nurse (GP practice) – cost per hour, including qualifications (58) |
| | **Total** | £659.11 | - |
| Gamma-glutamyl transferase increased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |

| Adverse reactions | Items | Unit cost | Reference |
|---|---|---|---|
| Hypertension | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| | Outpatient contact | £172.67 | NHS reference costs 2016/17 – WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology) (57) |
| | GP contacts (x2) | £37.00 (x2) | PSSRU 2017 – General practitioner – cost per surgery consultation lasting 9.22 minutes – including direct care staff costs, with qualification costs (58) |
| | **Total** | £863.78 | - |
| Palmar-plantar erthrodysaesthesia syndrome | Hospitalisation | £431.64 | NHS reference costs 2016/17 – JD07J Skin Disorders without Interventions, with CC score 2-5 – non-elective short stay (57) |
| Platelet count decreased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| Proteinuria | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| | Outpatient contact | £172.67 | NHS reference costs 2016/17 – WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology) (57) |
| | **Total** | £789.78 | - |
| Weight decreased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay (57) |
| | Dietician | £30.00 | PSSRU 2017 – dietitians/speech and language therapists - cost per working hour, Band 4 (58) |
| | **Total** | £647.11 | - |

Abbreviations: GP, general practitioner; NHS, National Health Service; PSSRU, Personal Social Services Research Unit.

### B.3.5.4 Miscellaneous unit costs and resource use

Additional costs considered in the analysis included those for end of life care (Table 40) and post-progression therapies (Table 41).

Costs for end of life care were taken from a 2014 report from the Nuffield Trust (61). Where resource use estimates were provided directly in this source, total costs have

been generated using unit costs from published national sources. Where only total costs were provided, modelled costs have been generated using inflation factors from the HCHS Pay and Prices Index (58).

**Table 40: Estimated costs in the last three months of life (cancer diagnosis) (61)**

| Resource use | Estimated costs/activity |
|---|---|
| All hospital contacts | £6,129.25 |
| Local authority-funded social care | £462.04 |
| Nurse visits | £317.10 |
| GP visits | £421.80 |
| **Total** | £7,330.19 |

Abbreviations: GP, general practitioner.

During REFLECT, use of post-progression therapies differed by treatment arm – there was an imbalance in post-treatment anti-cancer therapy during survival follow-up in the Western region. After discontinuing therapy in REFLECT, patients in the sorafenib arm were eligible for second-line trials (including for regorafenib) that enrolled patients who did not tolerate or had failed treatment with sorafenib, whereas lenvatinib-treated patients were ineligible for these trials as prior investigational agents were not permitted (see Section B.2.13.2). Patients in the sorafenib arm were permitted to continue receiving sorafenib as a post-progression therapy, and lenvatinib patients were permitted to switch to sorafenib following discontinuation. Some patients received other post-progression therapies in REFLECT, however only sorafenib and regorafenib are considered in the model as they are the only therapies licensed for this indication. No costs are applied for post-progression therapies other than sorafenib and regorafenib.

Multiplying the cost per day of sorafenib and regorafenib (Table 41) with the weighted duration of post-progression treatment (Table 42) provides the total cost of post-progression therapy (applied to each patient leaving the progression-free state, in the cycle in which this occurs) for each treatment arm:

- Lenvatinib - [redacted]
- Sorafenib - £1,900

**Table 41: Costs of post-progression therapies**

| | Units/pack* | mg/unit* | Cost/pack* | Cost/mg | Mean dose (mg) | Cost/day |
|---|---|---|---|---|---|---|
| Sorafenib | 112 | 200 | £3,576.56 | £0.16 | 663.8† | £105.99 |
| Regorafenib | 84 | 40 | £3,744.00 | £1.11 | 144‡ | £160.46 |

*Drug costs, pack and dosing information obtained from BNF (60); †mean dose assumed to be the same as first line; ‡mean dose taken from the RESORCE trial (74).

**Table 42: Use of post-progression therapies**

| | % of patients leaving progression-free state* | | Mean duration of post-progression therapy (days) | |
|---|---|---|---|---|
| | Lenvatinib arm | Sorafenib arm | Lenvatinib arm | Sorafenib arm |
| Sorafenib | [redacted] | [redacted] | [redacted] | [redacted] |
| Regorafenib | [redacted] | [redacted] | [redacted] | [redacted] |

*Calculated as the number of patients using each therapy divided by the number of patients experiencing either progression or death from the progression-free state. A weighted average cost is applied to all individuals leaving the PFS state in the cycle in which this occurs.

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

### B.3.6.1 Summary of base-case analysis inputs

A table of inputs and variables used in the base-case analysis is provided in Appendix N.

### B.3.6.2 Assumptions

| Component | Assumption | Justification |
|---|---|---|
| Clinical efficacy | Imbalances at baseline are adjusted using multivariable regression | In the presence of imbalances for strong predictors of outcomes, adjustment for such covariates generally improves the precision and efficiency of the analysis and avoids conditional bias from chance covariate imbalance. |
| | Independent statistical models are estimated for each arm | The proportional hazards assumption was not supported for PFS and was ambiguous for OS. |
| | Base-case distribution – OS | The log-logistic distribution was the preferred distribution based on standard statistical tests and is considered consistent with recent appraisals. |

| | | |
|---|---|---|
| | Base-case distribution – PFS | The log-normal distribution was the preferred distribution based on standard statistical tests for lenvatinib; although the gamma distribution was preferred for sorafenib, the resulting extrapolations were considered clinically implausible. |
| HRQoL | HRQoL depends only on health state and does not differ by treatment arm | Following adjustment for baseline characteristics, utility values for lenvatinib were similar to but numerically higher than those for sorafenib. It was therefore conservatively assumed that utility values do not differ by treatment arm. |
| Resource use and costs | All AE costs are applied in the first model cycle | This approach was adopted for simplicity. As observed outcomes are relatively complete, it is not expected that many further TEAEs would be observed. |
| | Excluded AEs are not associated with substantial costs | AEs considered include Grade 3 or 4 TEAEs occurring in ≥5% of patients in either treatment arm of REFLECT. Additional Grade 3 or 4 TEAEs that occurred in <5% of patients in either treatment arm were included if identified as being clinically or economically significant by UK clinical experts and are therefore expected to be the main determinants of cost. |
| | Medical resource use in the progression-free and progressed states is the same for lenvatinib and sorafenib | This assumption was made in the absence of other data, and was confirmed by a UK clinician. |
| | A one-off cost of post-progression therapies is applied in the cycle in which a patient leaves the progression-free state | This simplifying assumption avoids the use of multiple tunnel states to track use of post-progression therapies. |
| | The only post-progression therapies that are costed are sorafenib and regorafenib | These are the only licenced products in HCC. |

Abbreviations: AE, adverse event; HCC, hepatocellular carcinoma; HRQoL, health-related quality of life; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year.

## *B.3.7 Base-case results*

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

Base-case incremental cost-effectiveness analysis results are presented in Table 43. Lenvatinib is associated with lower total costs and higher total QALYs compared with sorafenib; lenvatinib is therefore considered to be dominant versus sorafenib.

**Table 43: Base-case results**

| Technologies | Total costs (£) | Total LYG | Total QALYs | Incremental costs (£) | Incremental LYG | Incremental QALYs | ICER incremental (£/QALY) |
|---|---|---|---|---|---|---|---|
| Lenvatinib | [redacted] | 1.66 | 1.18 | - | - | - | - |
| Sorafenib | £64,617 | 1.44 | 1.01 | [redacted] | -0.22 | -0.17 | Dominated |

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

Clinical outcomes from the model and disaggregated results of the base-case cost-effectiveness are provided in Appendix J.

## *B.3.8 Sensitivity analyses*

### B.3.8.1 Probabilistic sensitivity analysis

Joint parameter uncertainty was explored through probabilistic sensitivity analysis (PSA), in which all parameters are assigned distributions and varied jointly.[§] 10,000 Monte Carlo simulations were recorded. Where the covariance structure between parameters was known, correlated random draws were sampled from a multivariate normal distribution[**]. Results were plotted on the cost-effectiveness plane (CEP) and a cost-effectiveness acceptability curve (CEAC) was generated.

Parameters, estimates of uncertainty, and distributional assumptions used in PSA are presented in Appendix N.

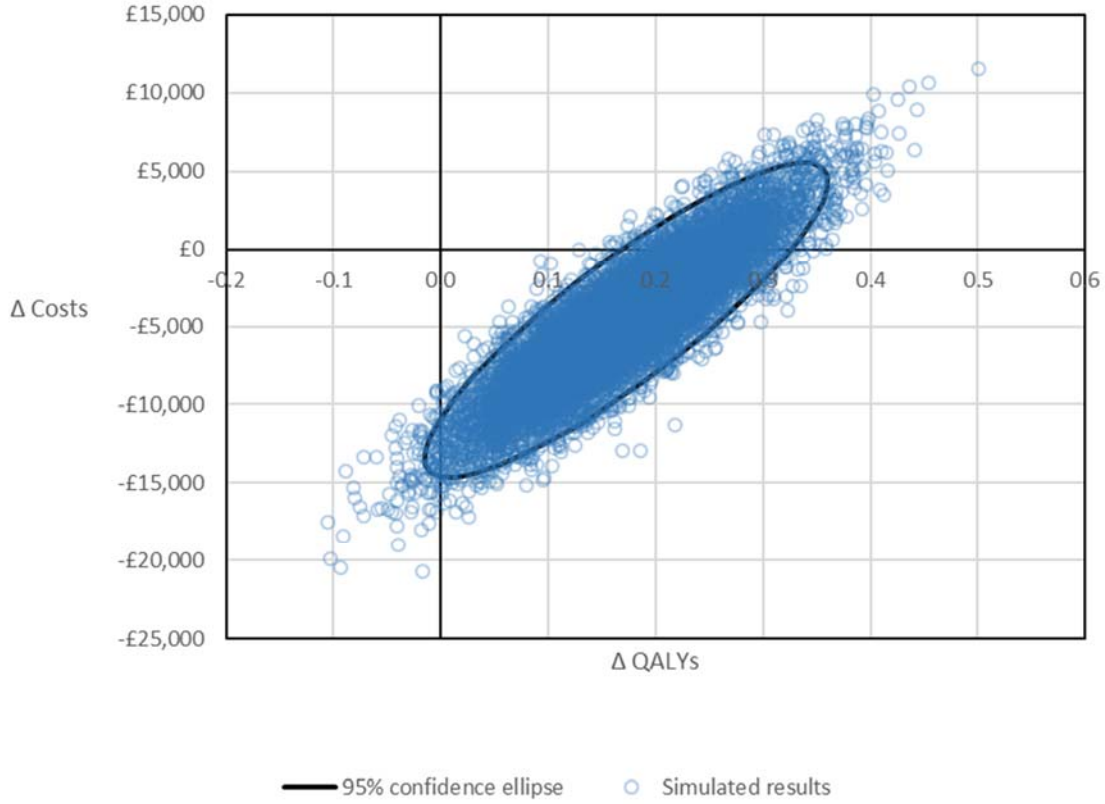
Figure 14 and Figure 15 present the CEP and CEAC, respectively. The probability that lenvatinib was cost-effective at a threshold of £50,000 per QALY was 100%. Across 10,000 PSA simulations, lenvatinib was associated with mean cost-savings of [redacted] (95% CI: [redacted], [redacted]) and mean incremental QALYs of 0.17 (95% CI: 0.02, 0.32). These results are considered to be congruent with deterministic cost-savings of [redacted] and a deterministic increase in QALYs of 0.17.

[§] Given that the utility values for progression-free and progressed disease are ordered (i.e. utility value for progression-free disease > utility value for progressed disease), the method described by Ren et al (20) has been used to ensure that this order is maintained for all simulations.

[**] Sampling from multivariate distributions is performed using code developed by the Centre for Bayesian Statistics in Health Economics (CHEBS); taken from http://www.shef.ac.uk/chebs/software.

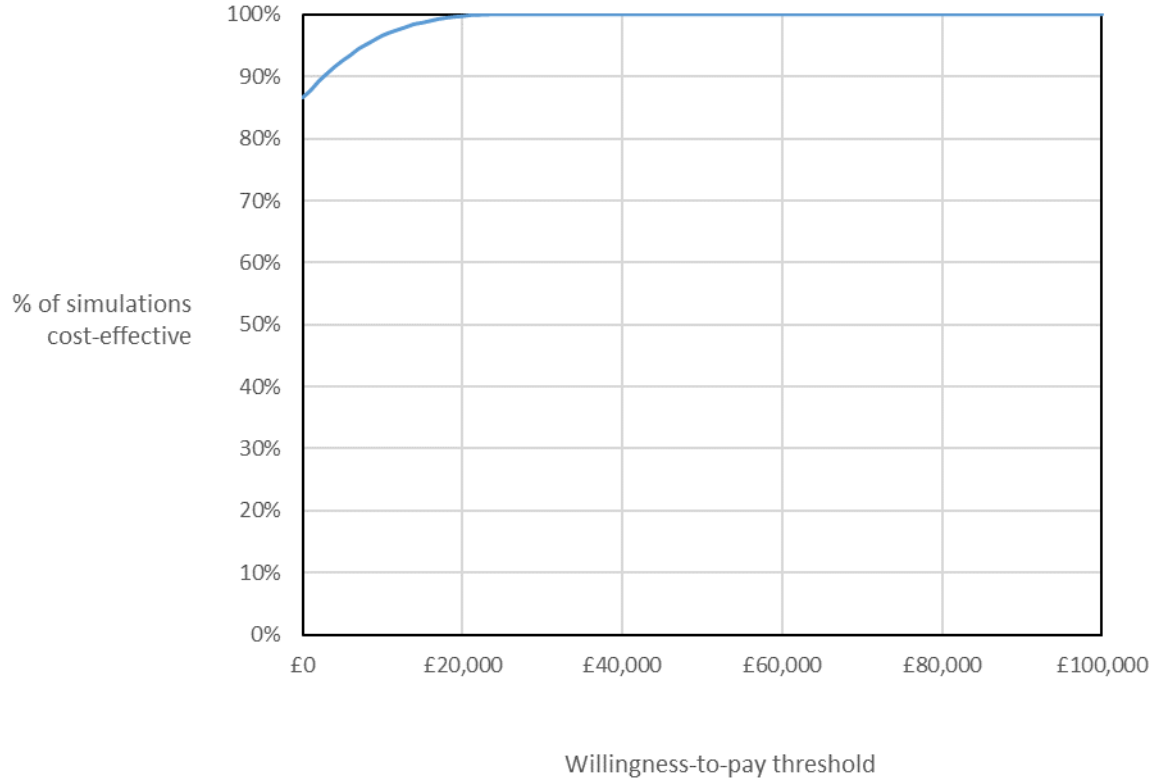

 

**Figure 14: Cost-effectiveness plane**




Abbreviations: QALYs, quality-adjusted life years.

**Figure 15: Cost-effectiveness acceptability curve**




### B.3.8.2 Deterministic sensitivity analysis

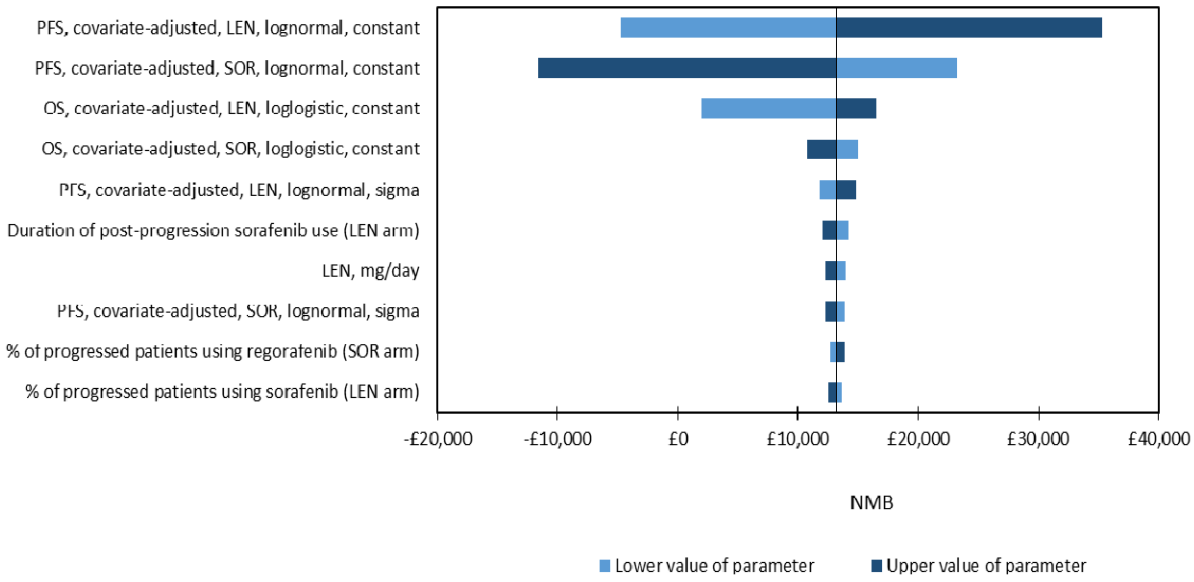
Parameter uncertainty was tested using univariate sensitivity analysis, in which all model parameters were systematically and independently varied over a plausible range determined by either the 95% confidence interval, or +/- 15% where no estimates of precision were available. Net monetary benefit was recorded at the upper and lower values to produce a tornado diagram, assuming a cost-effectiveness threshold of £50,000 per QALY (see Section B.2.13.3).

Upper and lower ranges of included parameters are presented in Appendix N.

Figure 16 presents the results of the univariate sensitivity analysis in the form of a tornado diagram. The most influential parameters were found to be the constant terms for the base-case PFS and OS models for each of the lenvatinib and sorafenib arms, which represent the baseline hazards of events. Other influential parameters included the ancillary parameters for the base-case PFS models in each of the lenvatinib and sorafenib arms; the proportion of lenvatinib patients using sorafenib as

a post-progression therapy, and the duration of this use; the proportion of sorafenib patients using regorafenib as a post-progression therapy; and the mean daily dose of lenvatinib.

**Figure 16: Results of univariate sensitivity analysis (tornado diagram)**




Abbreviations: LEN, lenvatinib; NMB, net monetary benefit; OS, overall survival; PFS, progression-free survival; SOR, sorafenib.

### B.3.8.3 Scenario analysis

The results of scenario analysis are presented in Table 44. All scenarios showed lenvatinib to be cost-effective versus sorafenib, assuming a cost-effectiveness threshold of £50,000 per QALY. All scenarios, with the exception of applying a 60% discount to the sorafenib list price, showed lenvatinib to be dominant versus sorafenib.

**Table 44: Scenario analysis**

| # | Scenario | Incremental Costs | Incremental QALYs | NMB | % change from base-case NMB | ICER |
|---|---|---|---|---|---|---|
| - | Base-case | [redacted] | 0.17 | £13,125 | 0% | Dominant |
| 1 | Drug wastage included | -£4,793 | 0.17 | £13,430 | 2% | Dominant |
| 2 | Mortality costs excluded | -£4,425 | 0.17 | £13,063 | 0% | Dominant |
| 3 | No covariate adjustment | -£6,814 | 0.11 | £12,552 | -4% | Dominant |
| 4 | Adjustment for AFP and stratification factors only | -£3,779 | 0.18 | £12,926 | -2% | Dominant |
| 5 | OS distribution: log-normal | -£4,287 | 0.18 | £13,128 | 0% | Dominant |
| 6 | OS distribution: gamma | -£6,681 | 0.13 | £13,043 | -1% | Dominant |
| 7 | OS distribution: Weibull | -£5,650 | 0.14 | £12,727 | -3% | Dominant |
| 8 | OS distribution: Gompertz | -£6,220 | 0.12 | £12,406 | -5% | Dominant |
| 9 | OS distribution: exponential | -£5,487 | 0.15 | £13,030 | -1% | Dominant |
| 10 | PFS distribution: log-logistic | -£5,669 | 0.18 | £14,465 | 10% | Dominant |
| 11 | PFS distribution: gamma† | -£1,647 | 0.17 | £9,902 | -25% | Dominant |
| 12 | PFS distribution: Weibull | -£2,746 | 0.17 | £11,150 | -15% | Dominant |
| 13 | PFS distribution: Gompertz† | -£2,133 | 0.17 | £10,454 | -20% | Dominant |
| 14 | PFS distribution: exponential | -£3,354 | 0.17 | £11,839 | -10% | Dominant |
| 15 | Resource use costs halved (all states) | -£3,966 | 0.17 | £12,604 | -4% | Dominant |
| 16 | Resource use costs doubled (all states) | -£5,530 | 0.17 | £14,167 | 8% | Dominant |
| 17 | Target dose assumed | -£6,225 | 0.17 | £14,863 | 13% | Dominant |
| 18 | Discount rates of 1.5% | -£4,180 | 0.19 | £13,489 | 3% | Dominant |
| 19 | Time horizon: 1 year | -£5,900 | 0.03 | £7,242 | -45% | Dominant |

| # | Scenario | Incremental Costs | Incremental QALYs | NMB | % change from base-case NMB | ICER |
|---|---|---|---|---|---|---|
| 20 | Time horizon: 2 years | -£6,237 | 0.07 | £9,823 | -25% | Dominant |
| 21 | Time horizon: 5 years | -£6,032 | 0.13 | £12,683 | -3% | Dominant |
| 22 | Sorafenib utility data used (committee preference) | -£4,487 | 0.18 | £13,366 | 2% | Dominant |
| 23 | Regorafenib utility data used | -£4,487 | 0.18 | £13,591 | 4% | Dominant |
| 24 | Post-progression utility of 0.5 | -£4,487 | 0.19 | £14,154 | 8% | Dominant |
| 25 | 60% discount applied to sorafenib | £6,872 | 0.17 | £1,765 | -87% | [redacted] |

Abbreviations: AFP, alpha fetoprotein; ICER, incremental cost-effectiveness ratio; NMB, net monetary benefit; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year.

† In these scenarios the PFS curves for lenvatinib and sorafenib are assumed equivalent at the point they would otherwise cross each other.

#### B.3.8.4 Summary of sensitivity analyses results

The results of PSA were found to be highly congruent with the deterministic base-case results and showed lenvatinib to be cost-effective versus sorafenib in 100% of simulations, assuming a cost-effectiveness threshold of £50,000 per QALY.

The most influential parameters in deterministic sensitivity analysis were the constant terms (representative of the baseline hazard of events) for the base-case PFS and OS models for each of the lenvatinib and sorafenib arms. The effects of other model parameters on the base-case NMB were found to be modest.

All scenarios, with the exception of the application of a 60% discount for sorafenib, showed lenvatinib to be dominant versus sorafenib; lenvatinib remained cost-effective versus sorafenib in this scenario.

### *B.3.9 Subgroup analysis*

No subgroup analyses are presented, as described in Section B.3.2.1.

### *B.3.10 Validation*

#### B.3.10.1 Validation of cost-effectiveness analysis

##### *B.3.10.1.1 Quality assurance*

Two economists external to the model development process performed quality assurance. Quality assurance was performed using cell-by-cell checks and logical tests; key parts of the model were also rebuilt to ensure consistency.

##### *B.3.10.1.2 Internal validity*

The internal validity of the electronic model was assessed by comparing model outcomes to those observed in REFLECT. Table 45 presents the results of this comparison. The economic model was considered consistent with REFLECT.

**Table 45: Comparison of model outcomes and REFLECT**

| Outcome | Economic model base-case | | Economic model scenario, using unadjusted ITT population | | Unadjusted REFLECT | |
|---|---|---|---|---|---|---|
| | Lenvatinib | Sorafenib | Lenvatinib | Sorafenib | Lenvatinib | Sorafenib |
| Median OS (months) | 13.8 | 12.0 | 13.8 | 12.9 | 13.6 | 12.3 |
| Median PFS (months) | 7.4 | 4.6 | 7.4 | 4.6 | 7.4 | 3.7 |
| Proportion surviving at month 12 | 56.6% | 50.8% | 55.8% | 51.7% | 55.0% | 50.0% |
| Proportion surviving at month 24 | 25.7% | 20.4% | 28.2% | 25.1% | 29.9% | 26.2% |

Abbreviations: ITT, intent-to-treat; OS, overall survival; PFS, progression-free survival.

#### *B.3.10.1.3 External validity*

In order to assess the external validity of the economic evaluation to UK clinical practice, model outputs were compared to observational data sources. The GIDEON (75) and King et al (10) studies provide observational data for the use of sorafenib as first line systemic treatment for the treatment of advanced HCC.

King et al, 2017 (10) performed an audit of NHS patients who had received sorafenib as first-line systemic therapy, identified using local Cancer Drugs Fund records or locally held databases. The median OS for evaluable sorafenib-treated with Child-Pugh A was 9.5 months. This compares to an unadjusted median OS for sorafenib of 12.3 months in REFLECT (see Section B.2.6.1.1), and a model base-case estimate of 12 months.

GIDEON is a global (39 countries), prospective, observational registry study evaluating the safety of sorafenib and treatment practices in HCC (75). Participants were candidates for systemic therapy in whom a decision to treat with sorafenib had been made. Patients were followed from the start of therapy with sorafenib to withdrawal of consent, death, or end of study. In the intent-to-treat population, median OS in patients with patients with Child-Pugh A was 13.6 months.

Given the survival estimates of sorafenib patients in REFLECT and the economic model were intermediate between both identified real-world data sources, the results of the economic evaluation were considered consistent with those expected in clinical practice.

### *B.3.11 Interpretation and conclusions of economic evidence*

A systematic review of the economic literature did not identify any published economic evaluations for lenvatinib in HCC relevant to the UK (see Section B.3.1), and so it was necessary to develop a de novo economic model. The model structure adopted is consistent with clinical practice and previous modelling approaches in HCC, and oncology more broadly.

The core assumptions of the economic evaluation, including the relevance of comparators and the identification of prognostic covariates, were informed and validated by UK-based clinical experts (see Section B.3.3.3), and unit costs for resource use were taken from UK sources. The overall trial population of REFLECT is considered to be reflective of the HCC population in UK clinical practice (see Section B.2.13.2). The economic evaluation was therefore considered highly relevant to the HCC population in England and Wales.

The use of multivariable parametric survival models allowed us to address imbalances in important prognostic variables identified at baseline (see Section B.3.3.1.2) and to improve precision, avoiding conditional bias from covariate imbalance. This analysis adjusted for baseline characteristics only; the imbalance in post-treatment therapy observed during REFLECT was not adjusted for. Estimates of relative efficacy may therefore be considered conservative (see Section B.2.13.2). The relatively complete data observed in REFLECT meant that reliance on extrapolation was limited, and the collection of EQ-5D permitted reliable estimates of HRQoL which were consistent with the clinical data and definitions used throughout the economic model.

The main external data sources used within the economic evaluation were the original, and updated, resource use surveys presented in previous Technology

Appraisals for sorafenib (24). The appraisal committees during those appraisals have previously highlighted this to be a source of uncertainty, and this remains the case within our economic evaluation.

Mean OS for individuals receiving sorafenib was predicted to be 18 months in the base-case, with a mean predicted survival gain of 3.1 months associated with lenvatinib. These estimates were considered to be robust and based on plausible predictions of OS. We believe lenvatinib therefore represents a life-extending treatment at the end of life (55).

The base-case analysis, including the lenvatinib PAS, suggests that lenvatinib would lead to cost-savings for the NHS compared to sorafenib. After inclusion of an assumed 60% discount for sorafenib, lenvatinib was associated with an ICER of [redacted]. We therefore believe that lenvatinib represents a cost-effective use of NHS resources.

## B.4 References


1. Office for National Statistics. Cancer Registration Statistics, England Statistical bulletins. 2016. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/previousReleases. Accessed on: February 2018.
2. Finn RS, Zhu AX. Targeting angiogenesis in hepatocellular carcinoma: focus on VEGF and bevacizumab. Expert Rev Anticancer Ther. 2009;9(4):503-9.
3. Yamaguchi R, Yano H, Iemura A, Ogasawara S, Haramaki M, Kojiro M. Expression of vascular endothelial growth factor in human hepatocellular carcinoma. Hepatology. 1998;28(1):68-77.
4. Lee HJ, Kang HJ, Kim KM, Yu ES, Kim KH, Kim SM, et al. Fibroblast growth factor receptor isotype expression and its association with overall survival in patients with hepatocellular carcinoma. Clin Mol Hepatol. 2015;21(1):60-70.
5. National Cancer Registration and Analysis Service. Trends in incidence of primary liver cancer subtypes. Available from: http://www.ncin.org.uk/publications/data_briefings/trends_in_incidence_of_primary_liver_cancer_subtypes. Accessed on: October 2017.
6. Cancer Research UK. Liver cancer statistics. 2017. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer#heading-One. Accessed on: October 2017.
7. Pons F, Varela M, Llovet JM. Staging systems in hepatocellular carcinoma. HPB (Oxford). 2005;7(1):35-41.
8. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology. 2007;132(7):2557-76.
9. Parkin DM. 11. Cancers attributable to infection in the UK in 2010. Br J Cancer. 2011;105 Suppl 2:S49-56.
10. King J, Palmer DH, Johnson P, Ross P, Hubner RA, Sumpter K, et al. Sorafenib for the Treatment of Advanced Hepatocellular Cancer - a UK Audit. Clin Oncol (R Coll Radiol). 2017;29(4):256-62.
11. Das D, Chattopadhyay D, Aslam T, Patanwala I, Walia D, Rose J, et al. NAFLD and the changing face of hepatocellular carcinoma (HCC). Gut. 2011;60:A243.
12. Williams R, Alexander G, Armstrong I, Baker A, Bhala N, Camps-Walsh G, et al. Disease burden and costs from excess alcohol consumption, obesity, and viral hepatitis: fourth report of the Lancet Standing Commission on Liver Disease in the UK. Lancet. 2017.
13. Younossi ZM, Blissett D, Blissett R, Henry L, Stepanova M, Younossi Y, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. Hepatology. 2016;64(5):1577-86.
14. Office for National Statistics. Cancer survival in England: Patients diagnosed between 2011 and 2015 and followed up to 2016. 2017. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed. Accessed on: January 2018.



15. Pascual S, Herrera I, Irurzun J. New advances in hepatocellular carcinoma. World J Hepatol. 2016;8(9):421-38.
16. Silva JP, Gorman RA, Berger NG, Tsai S, Christians KK, Clarke CN, et al. The prognostic utility of baseline alpha-fetoprotein for hepatocellular carcinoma patients. J Surg Oncol. 2017;116(7):831-40.
17. Llovet JM, Pena CE, Lathia CD, Shan M, Meinhardt G, Bruix J, et al. Plasma biomarkers as predictors of outcome in patients with advanced hepatocellular carcinoma. Clin Cancer Res. 2012;18(8):2290-300.
18. National Institute for Health and Care Excellence. Sorafenib for treating advanced hepatocellular carcinoma resource impact template. 2017. Available from: Available at https://www.nice.org.uk/guidance/ta474/resources. Accessed
19. Llovet JM, Bruix J. Novel advancements in the management of hepatocellular carcinoma in 2008. J Hepatol. 2008;48 Suppl 1:S20-37.
20. European Association For The Study Of The Liver, European Organisation For Research And Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. Eur J Cancer. 2012;48(5):599-641.
21. Verslype C, Rosmorduc O, Rougier P, Group EGW. Hepatocellular carcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012;23 Suppl 7:vii41-8.
22. National Cancer Registration and Analysis Service. Cancer Outcomes and Services Dataset Version 8.0. 2017. Available from: http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd_downloads_v8. Accessed on: October 2017.
23. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. 2018.
24. National Institute for Health and Care Excellence. Sorafenib for treating advanced hepatocellular carcinoma. 2017. Available from: https://www.nice.org.uk/guidance/ta474. Accessed on: October 2017.
25. Colagrande S, Inghilesi AL, Aburas S, Taliani GG, Nardi C, Marra F. Challenges of advanced hepatocellular carcinoma. World J Gastroenterol. 2016;22(34):7645-59.
26. Barone C, Koeberle D, Metselaar H, Parisi G, Sansonno D, Spinzi G. Multidisciplinary approach for HCC patients: hepatology for the oncologists. Ann Oncol. 2013;24 Suppl 2:ii15-23.
27. Cheng AL, Kang YK, Lin DY, Park JW, Kudo M, Qin S, et al. Sunitinib versus sorafenib in advanced hepatocellular cancer: results of a randomized phase III trial. J Clin Oncol. 2013;31(32):4067-75.
28. Johnson PJ, Qin S, Park JW, Poon RT, Raoul JL, Philip PA, et al. Brivanib versus sorafenib as first-line therapy in patients with unresectable, advanced hepatocellular carcinoma: results from the randomized phase III BRISK-FL study. J Clin Oncol. 2013;31(28):3517-24.
29. Cainap C, Qin S, Huang WT, Chung IJ, Pan H, Cheng Y, et al. Linifanib versus Sorafenib in patients with advanced hepatocellular carcinoma: results of a randomized phase III trial. J Clin Oncol. 2015;33(2):172-9.
30. Zhu AX, Rosmorduc O, Evans TR, Ross PJ, Santoro A, Carrilho FJ, et al. SEARCH: a phase III, randomized, double-blind, placebo-controlled trial of

sorafenib plus erlotinib in patients with advanced hepatocellular carcinoma. J Clin Oncol. 2015;33(6):559-66.
31. Abou-Alfa GK, Niedzwieski D, Knox JJ, Kaubisch A, Posey J, Tan BR, et al. Phase III randomized study of sorafenib plus doxorubicin versus sorafenib in patients with advanced hepatocellular carcinoma (HCC): CALGB 80802 (Alliance). Journal of Clinical Oncology. 2016;34(15_suppl):4003-.
32. Eisai Inc. A Multicenter, Randomized, Open-Label, Phase 3 Trial to Compare the Efficacy and Safety of Lenvatinib (E7080) Versus Sorafenib in First-Line Treatment of Subjects With Unresectable Hepatocellular Carcinoma (E7080-G000-304). Clinical Study Report. 2017;Data on file.
33. Eisai Inc. Statistical analysis plan: E7080-G000-304. Data on file. 2016.
34. Eisai Inc. Data on file: OS adjusted by post-treatment. 2017.
35. Eisai Inc. Data on file: Cross sectional analysis of EQ-5D. 2017.
36. Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet. 2018;Published Online February 9, 2018.
37. Eisai Inc. Phase 1/2 Study of E7080 in Patients With Advanced Hepatocellular Carcinoma (HCC) (E7080-J081-202). Clinical Study Report. 2015;Data on file.
38. Tamai T, Hayato S, Hojo S, Suzuki T, Okusaka T, Ikeda K, et al. Dose Finding of Lenvatinib in Subjects With Advanced Hepatocellular Carcinoma Based on Population Pharmacokinetic and Exposure-Response Analyses. J Clin Pharmacol. 2017;57(9):1138-47.
39. Ikeda K, Kudo M, Kawazoe S, Osaki Y, Ikeda M, Okusaka T, et al. Phase 2 study of lenvatinib in patients with advanced hepatocellular carcinoma. J Gastroenterol. 2017;52(4):512-9.
40. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30(1):52-60.
41. Lencioni R, Kudo M, Finn RS, Qin S, Han KH, Ikeda K, et al. Independent imaging review (IIR) results in a phase 3 trial of lenvatinib (LEN) vs sorafenib (SOR) in first-line treatment of patients (pts) with unresectable hepatocellular carcinoma (uHCC). Poster presented at the 2018 Gastrointestinal Cancers Symposium, January 18-20, San Francisco, California, USA. 2018.
42. Jackson R, Psarelli EE, Berhane S, Khan H, Johnson P. Impact of Viral Status on Survival in Patients Receiving Sorafenib for Advanced Hepatocellular Cancer: A Meta-Analysis of Randomized Phase III Trials. J Clin Oncol. 2017;35(6):622-8.
43. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008;359(4):378-90.
44. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol. 2009;10(1):25-34.
45. Food and Drug Administration. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics. 2007. Available from:



https://www.fda.gov/downloads/Drugs/Guidances/ucm071590.pdf. Accessed on: October 2017.
46. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008;100(10):698-711.
47. Chaplin S, Taylor M, Lapon J, White J. Economic evaluation of glass yttrium-90 microspheres versus sorafenib for the treatment of advanced hepatocellular carcinoma: cost effectiveness analysis in the United Kingdom. Cardiovascular and Interventional Radiology. 2015;38(3 Supplement 1):S279-80.
48. Connock M, Round J, Bayliss S, Tubeuf S, Greenheld W, Moore D. Sorafenib for the treatment of advanced hepatocellular carcinoma. Health Technol Assess. 2010;14 Suppl 1:17-21.
49. Palmer D, Ross P, Shah T, Yu D, Shergill S, Patterson K, et al. Cost effectiveness of selective internal radiation therapy (SIRT) with Y-90 resin microspheres versus sorafenib in Barcelona Clinic Liver Cancer (BCLC) stage C hepatocellular carcinoma patients in the UK. Annals of Oncology. 2017;28(Suppl 5):703P.
50. Scottish Medicines Consortium. Sorafenib, 200mg tablets (Nexavar®) No.482/08. 2008. Available from: https://www.scottishmedicines.org.uk/files/sorafenib__Nexavar___FINAL_JUNE_2008_for_website.pdf. Accessed
51. Scottish Medicines Consortium. Sorafenib, 200mg tablets (Nexavar®) No.482/08 Resubmission. 2011. Available from: https://www.scottishmedicines.org.uk/files/advice/sorafenib_Nexavar_RESUBMISSION_FINAL_DECEMBER_2010.doc_for_website.pdf. Accessed
52. Scottish Medicines Consortium. Sorafenib, 200mg tablets (Nexavar®) No.482/08 Second Resubmission. 2016. Available from: https://www.scottishmedicines.org.uk/SMC_Advice/Advice/482_08_sorafenib__Nexavar_/sorafenib_Nexavar_2nd_Resubmission. Accessed
53. Molla N, AlMenieir N, Simoneau E, Aljiffry M, Valenti D, Metrakos P, et al. The role of interventional radiology in the management of hepatocellular carcinoma. Curr Oncol. 2014;21(3):e480-e92.
54. Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU Technical Support Document 19. Partitioned Survival Analysis for Decision Modelling in Health Care: A Critical Review. 2017. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf. Accessed on: January 2018.
55. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword. Accessed on: November 2017.
56. Dobrez D, Cella D, Pickard AS, Lai JS, Nickolov A. Estimation of patient preference-based utility weights from the functional assessment of cancer therapy - general. Value Health. 2007;10(4):266-72.



57. National Health Service. Reference costs 2016/17. 2017. Available from: https://improvement.nhs.uk/resources/reference-costs/. Accessed on: December 2017.
58. Personal Social Services Research Unit. Unit Costs of Health and Social Care: Personal Social Services Research Unit; 2017.
59. National Health Service. Personal Social Services: Expenditure and Unit Costs, England - 2015-16. 2016. Available from: https://digital.nhs.uk/catalogue/PUB22240. Accessed on: January 2018.
60. British National Formulary. BNF. 2017. Available from: https://bnf.nice.org.uk/. Accessed on: December 2017.
61. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. Nuffield Trust, 2014.
62. Marie Curie Cancer Care. Understanding the cost of end of life care in different settings. 2012.
63. National Institute for Health and Care Excellence. Sorafenib for the treatment of advanced hepatocellular carcinoma. 2010. Available from: https://www.nice.org.uk/guidance/ta189. Accessed
64. Latimer N. NICE DSU Technical Support Document 14. Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2013. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf. Accessed on: December 2017.
65. European Medicines Agency. Guideline on adjustment for baseline covariates in clinical trials. 2015. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500184923.pdf. Accessed on: November 2017.
66. Dey AK, Kundu D. Discriminating Between the Log-Normal and Log-Logistic Distributions. Communications in Statistics - Theory and Methods. 2009;39(2):280-92.
67. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
68. Gandhi S, Khubchandani S, Iyer R. Quality of life and hepatocellular carcinoma. J Gastrointest Oncol. 2014;5(4):296-317.
69. National Institute for Health and Care Excellence. Regorafenib for previously treated unresectable hepatocellular carcinoma. NICE ID991. 2018.
70. National Institute for Health and Care Excellence. Lenvatinib with everolimus for previously treated advanced renal cell carcinoma. 2018. Available from: https://www.nice.org.uk/guidance/ta498. Accessed on: January 2018.
71. Bayer plc. Summary of product characteristics: Nexavar 200mg film-coated tablets. 2017. Available from: https://www.medicines.org.uk/emc/medicine/18520. Accessed on: December 2017.
72. National Institute for Health and Care Excellence. Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [ID1059]. 2018. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10101. Accessed

73. British National Formulary. Sorafenib. 2018. Available from: https://bnf.nice.org.uk/medicinal-forms/sorafenib.html. Accessed on: January 2018.
74. Bruix J., Qin S., Merle P., Granito A., Huang Y.H., Bodoky G., et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2017;389(10064):56-66.
75. Marrero JA, Kudo M, Venook AP, Ye SL, Bronowicki JP, Chen XP, et al. Observational registry of sorafenib use in clinical practice across Child-Pugh subgroups: The GIDEON study. J Hepatol. 2016;65(6):1140-7.

## B.5 Appendices

**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 analyses**

**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: Additional supporting data from the REFLECT pivotal trial**

**Appendix M: KOL input on covariate adjustments for the economic model**

**Appendix N: Variables applied in the base-case of the economic model**

**Appendix O: Statistical models**

**Appendix P: Unadjusted parametric models**

**Appendix Q: Appendix references**

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

**Single technology appraisal**

**Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma ID1089**

Dear Eisai

The Evidence Review Group, BMJ, and the technical team at NICE have looked at the submission received on 15 February 2018 from Eisai. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **Monday 26 March 2018**.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Abi Senthinathan, Technical Lead (Abitha.senthinathan@nice.org.uk). Any procedural questions should be addressed to Stephanie Callaghan, Project Manager (Stephanie.Callaghan@nice.org.uk).

Yours sincerely

Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

## Section A: Clarification on effectiveness data

### Clinical effectiveness evidence

A1. **Priority question**: Please provide a rationale for patients being censored for PFS for, "No progression at the time of treatment discontinuation" (Table 14 of the company submission), when page 20 states that, "patients who discontinued study treatment for any reason other than disease progression were followed in the Randomisation Phase until disease progression or start of another anti-cancer therapy".

A2. **Priority question**: Please clarify whether 'objectively documented disease progression' as stated on page 20 of the company submission regarding discontinuation of study treatment was according to investigator-assessed modified RECIST, or another set of criteria?

A3. **Priority question**: Please provide a full breakdown for each group of the type of anti-cancer medications and procedures received after discontinuation of the study medication for:

a. The full population;

b. The Western subgroup

A4. **Priority question**: For the Western subgroup, please provide mean (SD) and median (Q1, Q3) number of cycles of treatment for each group.

A5. **Priority question**: Kudo 2018 states that, "Modifications to sorafenib doses were implemented according to prescribing information in each region"; for each treatment group by geographic region, please provide mean (SD) and median (Q1, Q3) dose intensity, and rates of dose interruption and reduction.

A6. **Priority question**: Please provide equivalent adverse event data used in the model for the Western subgroup, unadjusted for treatment exposure (i.e. Grade 3 or above occurring in at least 5% of either group, plus diarrhoea, asthenia and fatigue).

A7. Please provide details of the number of patients in each group whose survival follow-up was discontinued by the sponsor (company submission, page 20), and reasons why.

A8. Please explain the clinical reasoning behind AFP levels being dichotomised at 200 ng/mL, and whether differences between mean and median baseline levels are clinically meaningful.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A9. Please clarify whether there was an investigator assessment of PFS using RECIST v1.1 in REFLECT and, if there was, please provide the results.

**Literature searching**

A10. Please provide details of any clinical effectiveness studies that were excluded from the systematic literature review for the reason, "Studies where the primary endpoint was not met", as listed in Table 1 of Appendix D.1.2.

**Section B: Clarification on cost-effectiveness data**

**Survival analysis**

B1. **Priority question**: For the covariate adjusted survival analyses for independent imaging review (IIR) PFS and OS using the full population, please provide an additional set of analyses using the corrected group prognosis (CGP) method and include an option in the economic model to alternate between that and the mean of covariates (MoC) method.

B2. **Priority question**: Please perform adjusted and unadjusted survival analyses for IIR PFS and OS using the Western population subgroup of the REFLECT trial and update the economic model to include this subgroup analysis as an option. For the adjusted analyses, please use both the CGP and MoC methods and include both as options in the economic model.

B3. **Priority question**: Please provide an additional set of OS parametric models adjusted for post-progression anti-cancer therapies, for the full population and the Western subgroup, and include these as options in the economic model.

**Data**

B4. **Priority question:** For IIR PFS, OS, and time to treatment discontinuation, please provide the following for both the full population and Western subgroup of the REFLECT trial:

a. individual times for each event/censor;

b. reasons for censoring (depending on the response to A1, please consider providing an alternative IIR PFS analysis that does not censor for "no progression at the time of treatment discontinuation").

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

**Adverse events**

B5. Please clarify why grade 5 adverse events are not included in the economic model. Also, please provide a scenario analysis that includes grade 5 treatment-emergent adverse events (regardless of incidence level) using appropriate costs.

**Health-related quality of life**

B6. Given that disutilities associated with adverse events were not explicitly modelled, please consider providing adjusted and unadjusted treatment-specific utilities for the on- and off-treatment periods for those who are progression-free, for both the full population and the Western subgroup in the REFLECT trial. Please also add these as options in the economic model.

B7. Please provide Table 32 of the company submission (Summary of EQ-5D data) for the Western subgroup.

B8. Please clarify how covariates for the linear mixed model were chosen. If AFP levels and HCC aetiology were not included as covariates, please justify their exclusion.

B9. Please provide further details on the steps taken to adjust mean utility values, including the coefficients and p-values for the covariates included in the linear mixed model.

**Resource use and costs**

B10. Please provide a scenario, applying appropriate resources and costs, which includes the full list of post-progression therapies and procedures that patients in REFLECT received, including treatments without marketing authorisation in the UK. Please provide results for the full population and Western subgroup and include this as an option in the economic model.

B11. Please provide a scenario which excludes post-progression therapy costs (sorafenib and regorafenib). Please provide results for the full population and Western subgroup and include this as an option in the economic model.

B12. Please clarify why no monitoring or testing costs are applied to patients who receive post-progression therapies (sorafenib or regorafenib).

B13. Please clarify why mortality costs and post-progression drug costs are applied using the OS curves (column G in Appendix Transition (LEN) and (SOR)) rather than the health state values (column L in Appendix Transition (LEN) and (SOR)) in which a half cycle correction has been applied. The ERG considers there to be an inconsistent application of the half-cycle correction for each cost category.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B14. The ERG notes that Georghiou and Bardsley (2014) report end-of-life costs using a 2010/11 cost year rather than a 2013/14 cost year assumed by the company. Please consider inflating this cost appropriately.

B15. Please amend the drug wastage scenario to account for the wastage associated with post-progression drug costs, as the current approach is inconsistent.

B16. In the drug wastage scenario, please clarify why primary drug wastage costs are informed by the TTD curves rather than the on-treatment health states in which a half-cycle correction has been applied.

B17. Please explain the difference between the social care costs applied to patients with progressed disease and end-of-life care costs. The ERG is concerned end-of-life care costs have been double counted.

**Systematic literature review**

B18. Please clarify why TA189 and TA474 were not included as studies reporting relevant resource and cost use data in the literature review.

B19. Please clarify why the NICE submission for regorafenib (reference 69 in company submission) in Table 34 of the company submission was not included in the literature review for studies reporting HRQOL data.

B20. Please clarify why the Cochrane library was restricted by intervention (Lenvatinib or Sorafenib) for the HRQOL search.

B21. Figure 5 in Appendix H does not include the number of records identified via grey literature searches. Please add this information to figure 5.

B22. Please provide excluded studies lists with reasons for exclusion for the cost-effectiveness, resource use and HRQOL searches.

B23. The ERG has identified discrepancies between Table 18 of Appendix J and the economic model, specifically:

- post-progression drug costs for lenvatinib (£[redacted] vs £[redacted]);
- total costs for lenvatinib (£[redacted] vs [redacted];
- primary drug costs for sorafenib (£10,582 vs £21,163);
- post-progression drug costs for sorafenib (£1,269 vs 1,864);
- total costs for sorafenib (£53,440 vs £64,617).

Please clarify if the values in the economic model are correct.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

**Section C: Textual clarifications and additional points**

C1. Please provide the listings and appendices for the CSR.

**Single technology appraisal**

**Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma ID1089**

Dear Eisai

The Evidence Review Group, BMJ, and the technical team at NICE have looked at the submission received on 15 February 2018 from Eisai. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **Monday 26 March 2018**.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Abi Senthinathan, Technical Lead (Abitha.senthinathan@nice.org.uk). Any procedural questions should be addressed to Stephanie Callaghan, Project Manager (Stephanie.Callaghan@nice.org.uk).

Yours sincerely

Frances Sutcliffe
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

## Section A: Clarification on effectiveness data

### Clinical effectiveness evidence

A1. **Priority question**: Please provide a rationale for patients being censored for PFS for, "No progression at the time of treatment discontinuation" (Table 14 of the company submission), when page 20 states that, "patients who discontinued study treatment for any reason other than disease progression were followed in the Randomisation Phase until disease progression or start of another anti-cancer therapy".

The primary PFS analysis was conducted according to FDA guidance (1), where patients were censored when they discontinued treatment for any reason other than disease progression. A sensitivity analysis was conducted based on EMA guidance (2), where patients were not censored at discontinuation if they did not have disease progression. The results of this sensitivity analysis (presented in Table 1 below and Table 14.2.2.2 of the CSR) were consistent with the primary PFS analysis based on FDA censoring guidance:

- Primary analysis: Median PFS 7.4 months in the lenvatinib arm and 3.7 months in the sorafenib arm (HR: 0.66; 95% CI: 0.57, 0.77; p<0.00001)
- Sensitivity analysis: Median PFS [redacted] months in the lenvatinib arm and [redacted] months in the sorafenib arm (HR: [redacted]; 95% CI: [redacted]; [redacted]).

The primary analysis conducted according to FDA guidance on censoring was accepted by the EMA and will be included in the lenvatinib summary of product characteristics (Company Submission Document B, Appendix C).

**Table 1: Progression-free survival sensitivity analysis based on randomisation stratification factors and treating all PDs and deaths as events – Full Analysis Set**

| | Lenvatinib N=478 | Sorafenib N=476 |
|---|---|---|
| Patients with events, n (%) | [redacted] | [redacted] |
| Progressive disease | [redacted] | [redacted] |
| Death | [redacted] | [redacted] |
| Censored patients, n (%) | [redacted] | [redacted] |
| No baseline tumour assessment | [redacted] | [redacted] |
| No post-baseline tumour assessment | [redacted] | [redacted] |
| No progression at the time of data cut-off | [redacted] | [redacted] |
| No progression at the time of consent withdrawal | [redacted] | [redacted] |
| Progression-free survival (months)* | | |
| Median (95% CI) | [redacted] | [redacted] |
| Q1 (95% CI) | [redacted] | [redacted] |
| Q3 (95% CI) | [redacted] | [redacted] |
| Progression-free survival rate (%) (95% CI)† at | | |
| 6 Months | [redacted] | [redacted] |
| 12 Months | [redacted] | [redacted] |

| | Lenvatinib N=478 | Sorafenib N=476 |
|---|---|---|
| 18 Months | [redacted] | [redacted] |
| 24 Months | [redacted] | [redacted] |
| Stratified cox model hazard ratio (95% CI)‡,§ | [redacted] | |
| Stratified log-rank test p-value§ | [redacted] | |

*Quartiles are estimated by Kaplan-Meier method, and the 95% confidence intervals are estimated with a generalised Brookmeyer and Crowley method; †PFS rate and 95% CI were calculated using the Kaplan-Meier product-limit method and the Greenwood Formula; ‡Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor; §Stratified by region (Region 1: Asia-Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg);

Abbreviations: CI, confidence interval; PD, progressive disease; PFS, progression-free survival; Q, quartile.

A2. **Priority question**: Please clarify whether 'objectively documented disease progression' as stated on page 20 of the company submission regarding discontinuation of study treatment was according to investigator-assessed modified RECIST, or another set of criteria?

Objectively documented disease progression refers to radiological progression based on investigator-assessed mRECIST.

A3. **Priority question**: Please provide a full breakdown for each group of the type of anti-cancer medications and procedures received after discontinuation of the study medication for:

a. The full population;

b. The Western subgroup

A full breakdown of anti-cancer procedures received after discontinuation of study medication during survival follow-up is provided for the full population in Table 2 (Table 14.1.6.9 in the CSR) and for the Western population in Table 3 (Table 14.1.6.9.1 of the CSR). A full breakdown of anti-cancer medications received after discontinuation of study medication during survival follow-up is provided for the full population in Table 4 (Table 14.1.6.11 of the CSR) and the Western population in Table 5 (Table 14.1.6.11.1 of the CSR).

**Table 2: Anti-cancer procedures by System Organ Class and Preferred Term during survival follow-up – Full Analysis Set**

| System Organ Class<br>Preferred Term | Lenvatinib (N=478)<br>n (%) | Sorafenib (N=476)<br>n (%) |
|---|---|---|
| Patients with any anti-cancer procedure during survival follow-up | 122 (25.5) | 130 (27.3) |
| Surgical and medical procedures | [redacted] | [redacted] |

| System Organ Class<br>Preferred Term | Lenvatinib<br>(N=478)<br>n (%) | Sorafenib<br>(N=476)<br>n (%) |
|---|---|---|
| Abdominal cavity drainage | [redacted] | [redacted] |
| Brachytherapy | [redacted] | [redacted] |
| Cancer surgery | [redacted] | [redacted] |
| Central venous catheterisation | [redacted] | [redacted] |
| Chemotherapy | [redacted] | [redacted] |
| Gamma radiation therapy | [redacted] | [redacted] |
| Gamma radiation therapy to brain | [redacted] | [redacted] |
| Gamma radiation therapy to lung | [redacted] | [redacted] |
| Hepatectomy | [redacted] | [redacted] |
| Hepatic embolisation | [redacted] | [redacted] |
| High frequency ablation | [redacted] | [redacted] |
| High intensity focused ultrasound | [redacted] | [redacted] |
| Lung lobectomy | [redacted] | [redacted] |
| Magnetic therapy | [redacted] | [redacted] |
| Mass excision | [redacted] | [redacted] |
| Microwave therapy | [redacted] | [redacted] |
| Percutaneous ethanol injection therapy | [redacted] | [redacted] |
| Pulmonary resection | [redacted] | [redacted] |
| Radioembolisation | [redacted] | [redacted] |
| Radiotherapy | [redacted] | [redacted] |
| Radiotherapy to abdomen | [redacted] | [redacted] |
| Radiotherapy to adrenal gland | [redacted] | [redacted] |
| Radiotherapy to bone | [redacted] | [redacted] |
| Radiotherapy to brain | [redacted] | [redacted] |
| Radiotherapy to joint | [redacted] | [redacted] |
| Radiotherapy to liver | [redacted] | [redacted] |
| Radiotherapy to lung | [redacted] | [redacted] |
| Radiotherapy to lymph nodes | [redacted] | [redacted] |
| Regional chemotherapy | [redacted] | [redacted] |
| Rib excision | [redacted] | [redacted] |
| Spinal operation | [redacted] | [redacted] |
| Supportive care | [redacted] | [redacted] |
| Therapeutic embolisation | [redacted] | [redacted] |

| System Organ Class<br>Preferred Term | Lenvatinib<br>(N=478)<br>n (%) | Sorafenib<br>(N=476)<br>n (%) |
|---|---|---|
| Transcatheter arterial chemoembolisation | [redacted] | [redacted] |
| Tumour excision | [redacted] | [redacted] |

**Table 3: Anti-cancer procedures by System Organ Class and Preferred Term during survival follow-up for the Western region – Full Analysis Set**

| System Organ Class<br>Preferred Term | Lenvatinib<br>(N=157)<br>n (%) | Sorafenib<br>(N=157)<br>n (%) |
|---|---|---|
| Patients with any anti-cancer procedure during survival follow-up | 11 (7.0) | 18 (11.5) |
| Surgical and medical procedures | [redacted] | [redacted] |
| Abdominal cavity drainage | [redacted] | [redacted] |
| Brachytherapy | [redacted] | [redacted] |
| Cancer surgery | [redacted] | [redacted] |
| Central venous catheterisation | [redacted] | [redacted] |
| Chemotherapy | [redacted] | [redacted] |
| Gamma radiation therapy | [redacted] | [redacted] |
| Gamma radiation therapy to brain | [redacted] | [redacted] |
| Gamma radiation therapy to lung | [redacted] | [redacted] |
| Hepatectomy | [redacted] | [redacted] |
| Hepatic embolisation | [redacted] | [redacted] |
| High frequency ablation | [redacted] | [redacted] |
| High intensity focused ultrasound | [redacted] | [redacted] |
| Lung lobectomy | [redacted] | [redacted] |
| Magnetic therapy | [redacted] | [redacted] |
| Mass excision | [redacted] | [redacted] |
| Microwave therapy | [redacted] | [redacted] |
| Percutaneous ethanol injection therapy | [redacted] | [redacted] |
| Pulmonary resection | [redacted] | [redacted] |
| Radioembolisation | [redacted] | [redacted] |
| Radiotherapy | [redacted] | [redacted] |
| Radiotherapy to abdomen | [redacted] | [redacted] |
| Radiotherapy to adrenal gland | [redacted] | [redacted] |
| Radiotherapy to bone | [redacted] | [redacted] |
| Radiotherapy to brain | [redacted] | [redacted] |

| System Organ Class<br>Preferred Term | Lenvatinib<br>(N=157)<br>n (%) | Sorafenib<br>(N=157)<br>n (%) |
|---|---|---|
| Radiotherapy to joint | [redacted] | [redacted] |
| Radiotherapy to liver | [redacted] | [redacted] |
| Radiotherapy to lung | [redacted] | [redacted] |
| Radiotherapy to lymph nodes | [redacted] | [redacted] |
| Regional chemotherapy | [redacted] | [redacted] |
| Rib excision | [redacted] | [redacted] |
| Spinal operation | [redacted] | [redacted] |
| Supportive care | [redacted] | [redacted] |
| Therapeutic embolisation | [redacted] | [redacted] |
| Transcatheter arterial chemoembolisation | [redacted] | [redacted] |
| Tumour excision | [redacted] | [redacted] |

**Table 4: Anti-cancer medications by ATC level and Preferred Term during survival follow-up – Full Analysis Set**

| Anatomical class (ATC level 1)<br>Pharmacological class (ATC level 3)<br>WHO Drug Name (Preferred Term)[†] | Lenvatinib (N=478)<br>n (%) | Sorafenib<br>(N=476)<br>n (%) |
|---|---|---|
| Patients with any anti-cancer medication (not given for any procedure) during survival follow-up | 156 (32.6) | 184 (38.7) |
| Antineoplastic and immunomodulating agents | [redacted] | [redacted] |
| Antimetabolites | [redacted] | [redacted] |
| Capecitabine | [redacted] | [redacted] |
| Floxuridine | [redacted] | [redacted] |
| Fluorouracil | [redacted] | [redacted] |
| Gemcitabine | [redacted] | [redacted] |
| Gimeracil | [redacted] | [redacted] |
| Gimeracil w/oteracil potassium/tegafur | [redacted] | [redacted] |
| Raltitrexed | [redacted] | [redacted] |
| Tegafur | [redacted] | [redacted] |
| Uftoral | [redacted] | [redacted] |
| Uracil | [redacted] | [redacted] |
| Cytotoxic antibiotics and related substances | [redacted] | [redacted] |
| Doxorubicin | [redacted] | [redacted] |
| Epirubicin | [redacted] | [redacted] |
| Mitomycin | [redacted] | [redacted] |
| Mitoxantrone | [redacted] | [redacted] |

| Anatomical class (ATC level 1)<br>Pharmacological class (ATC level 3)<br>WHO Drug Name (Preferred Term)[†] | Lenvatinib (N=478)<br>n (%) | Sorafenib (N=476)<br>n (%) |
|---|---|---|
| Hormone antagonists and related agents | [redacted] | [redacted] |
| Enzalutamide | [redacted] | [redacted] |
| Immunostimulants | [redacted] | [redacted] |
| Interferon | [redacted] | [redacted] |
| Immunosuppressants | [redacted] | [redacted] |
| Lenalidomide | [redacted] | [redacted] |
| Thalidomide | [redacted] | [redacted] |
| Other antineoplastic agents | [redacted] | [redacted] |
| Axitinib | [redacted] | [redacted] |
| Bevacizumab | [redacted] | [redacted] |
| Cabozantinib | [redacted] | [redacted] |
| Cisplatin | [redacted] | [redacted] |
| Erlotinib | [redacted] | [redacted] |
| Everolimus | [redacted] | [redacted] |
| Ipilimumab | [redacted] | [redacted] |
| Lambrolizumab | [redacted] | [redacted] |
| Lenvatinib | [redacted] | [redacted] |
| Monoclonal antibodies | [redacted] | [redacted] |
| Nintedanib | [redacted] | [redacted] |
| Nivolumab | [redacted] | [redacted] |
| Oteracil | [redacted] | [redacted] |
| Other antineoplastic agents | [redacted] | [redacted] |
| Oxaliplatin | [redacted] | [redacted] |
| Pazopanib | [redacted] | [redacted] |
| Protein kinase inhibitors | [redacted] | [redacted] |
| Ramucirumab | [redacted] | [redacted] |
| Regorafenib | [redacted] | [redacted] |
| Sorafenib | [redacted] | [redacted] |
| Sunitinib | [redacted] | [redacted] |
| Tivantinib | [redacted] | [redacted] |
| Vorinostat | [redacted] | [redacted] |
| Plant alkaloids and other natural products | [redacted] | [redacted] |
| Brucea javanica | [redacted] | [redacted] |

| Anatomical class (ATC level 1)<br>Pharmacological class (ATC level 3)<br>WHO Drug Name (Preferred Term)† | Lenvatinib (N=478)<br>n (%) | Sorafenib (N=476)<br>n (%) |
|---|---|---|
| Etoposide | [redacted] | [redacted] |
| Paclitaxel | [redacted] | [redacted] |
| Musculoskeletal system | [redacted] | [redacted] |
| Drugs affecting bone structure and mineralisation | [redacted] | [redacted] |
| Zoledronic acid | [redacted] | [redacted] |
| Various | [redacted] | [redacted] |
| All other therapeutic products | [redacted] | [redacted] |
| Folinic acid | [redacted] | [redacted] |
| Investigational drug | [redacted] | [redacted] |
| Investigational drug | [redacted] | [redacted] |
| Protein supplements | [redacted] | [redacted] |
| Protein supplements | [redacted] | [redacted] |
| Unspecified herbal and traditional medicine | [redacted] | [redacted] |
| Coix lacryma-jobi | [redacted] | [redacted] |
| Coix spp. | [redacted] | [redacted] |
| Panax ginseng | [redacted] | [redacted] |
| Unspecified herbal and traditional medicine | [redacted] | [redacted] |

†WHODD MAR2016 HD B2 was used to code post-treatment anticancer medications not given for any procedure.
Abbreviations: ATC, anatomical class; WHO, World Health Organization.

**Table 5: Anti-cancer medications by ATC level and Preferred Term during survival follow-up in the Western region – Full Analysis Set**

| Anatomical class (ATC Level 1)<br>Pharmacological class (ATC Level 3)<br>WHO Drug Name (Preferred Term)† | Lenvatinib (N=157)<br>n (%) | Sorafenib (N=157)<br>n (%) |
|---|---|---|
| Patients with any anti-cancer medication (not given for any procedure) during survival follow-up | 41 (26.1) | 61 (38.9) |
| Antineoplastic and immunomodulating agents | [redacted] | [redacted] |
| Antimetabolites | [redacted] | [redacted] |
| Capecitabine | [redacted] | [redacted] |
| Floxuridine | [redacted] | [redacted] |
| Fluorouracil | [redacted] | [redacted] |
| Gemcitabine | [redacted] | [redacted] |
| Gimeracil | [redacted] | [redacted] |
| Gimeracil w/oteracil Potassium/tegafur | [redacted] | [redacted] |
| Raltitrexed | [redacted] | [redacted] |

| Anatomical class (ATC Level 1)<br>Pharmacological class (ATC Level 3)<br>WHO Drug Name (Preferred Term)[†] | Lenvatinib (N=157) n (%) | Sorafenib (N=157) n (%) |
|---|---|---|
| Tegafur | [redacted] | [redacted] |
| Uftoral | [redacted] | [redacted] |
| Uracil | [redacted] | [redacted] |
| Cytotoxic antibiotics and related substances | [redacted] | [redacted] |
| Doxorubicin | [redacted] | [redacted] |
| Epirubicin | [redacted] | [redacted] |
| Mitomycin | [redacted] | [redacted] |
| Mitoxantrone | [redacted] | [redacted] |
| Hormone antagonists and related agents | [redacted] | [redacted] |
| Enzalutamide | [redacted] | [redacted] |
| Immunostimulants | [redacted] | [redacted] |
| Interferon | [redacted] | [redacted] |
| Immunosuppressants | [redacted] | [redacted] |
| Lenalidomide | [redacted] | [redacted] |
| Thalidomide | [redacted] | [redacted] |
| Other antineoplastic agents | [redacted] | [redacted] |
| Axitinib | [redacted] | [redacted] |
| Bevacizumab | [redacted] | [redacted] |
| Cabozantinib | [redacted] | [redacted] |
| Cisplatin | [redacted] | [redacted] |
| Erlotinib | [redacted] | [redacted] |
| Everolimus | [redacted] | [redacted] |
| Ipilimumab | [redacted] | [redacted] |
| Lambrolizumab | [redacted] | [redacted] |
| Lenvatinib | [redacted] | [redacted] |
| Monoclonal antibodies | [redacted] | [redacted] |
| Nintedanib | [redacted] | [redacted] |
| Nivolumab | [redacted] | [redacted] |
| Oteracil | [redacted] | [redacted] |
| Other antineoplastic agents | [redacted] | [redacted] |
| Oxaliplatin | [redacted] | [redacted] |
| Pazopanib | [redacted] | [redacted] |
| Protein kinase inhibitors | [redacted] | [redacted] |

| Anatomical class (ATC Level 1)<br>Pharmacological class (ATC Level 3)<br>WHO Drug Name (Preferred Term)[†] | Lenvatinib (N=157) n (%) | Sorafenib (N=157) n (%) |
|---|---|---|
| Ramucirumab | [redacted] | [redacted] |
| Regorafenib | [redacted] | [redacted] |
| Sorafenib | [redacted] | [redacted] |
| Sunitinib | [redacted] | [redacted] |
| Tivantinib | [redacted] | [redacted] |
| Vorinostat | [redacted] | [redacted] |
| Plant alkaloids and other natural products | [redacted] | [redacted] |
| Brucea javanica | [redacted] | [redacted] |
| Etoposide | [redacted] | [redacted] |
| Paclitaxel | [redacted] | [redacted] |
| Musculoskeletal system | [redacted] | [redacted] |
| Drugs affecting bone structure and mineralisation | [redacted] | [redacted] |
| Zoledronic acid | [redacted] | [redacted] |
| Various | [redacted] | [redacted] |
| All other therapeutic products | [redacted] | [redacted] |
| Folinic acid | [redacted] | [redacted] |
| Investigational drug | [redacted] | [redacted] |
| Investigational drug | [redacted] | [redacted] |
| Protein supplements | [redacted] | [redacted] |
| Protein supplements | [redacted] | [redacted] |
| Unspecified herbal and traditional medicine | [redacted] | [redacted] |
| Coix lacryma-jobi | [redacted] | [redacted] |
| Coix spp. | [redacted] | [redacted] |
| Panax ginseng | [redacted] | [redacted] |
| Unspecified herbal and traditional medicine | [redacted] | [redacted] |

[†]WHODD MAR2016 HD B2 was used to code post-treatment anticancer medications not given for any procedure.
Abbreviations: ATC, anatomical class; WHO, World Health Organization.

A4. **Priority question**: For the Western subgroup, please provide mean (SD) and median (Q1, Q3) number of cycles of treatment for each group.

Table 6 presents the duration of treatment in months for each arm in the Western region. Table 7 presents an approximation of the number of cycles in the same population, calculated using the following equation: number of cycles = number of months*((365/12)/28).

**Table 6: Duration of treatment for the Western region (months) – Safety Analysis Set**

| | Lenvatinib N=155 | Sorafenib N=156 |
|---|---|---|
| Duration of Treatment† (months) | | |
| Mean (SD) | 8.0 (6.88) | 6.8 (6.20) |
| Median | 6.2 | 4.6 |
| Q1, Q3 | 2.9, 11.3 | 2.1, 9.2 |
| Min, max | 0.0, 32.8 | 0.1, 32.8 |

Abbreviations: SD, standard deviation.

**Table 7: Duration of treatment for the Western region (approximation of cycles) – Safety Analysis Set**

| | Lenvatinib N=155 | Sorafenib N=156 |
|---|---|---|
| Duration of Treatment† (months) | | |
| Mean (SD) | 8.7 (7.47) | 7.4 (6.74) |
| Median | 6.7 | 5.0 |
| Q1, Q3 | 3.2, 12.3 | 2.3, 10.0 |
| Min, max | 0.0, 35.6 | 0.1, 35.6 |

Abbreviations: SD, standard deviation. Approximation of number of cycles based on following equation: number of cycles = number of months*((365/12)/28).

A5. **Priority question**: Kudo 2018 states that, “Modifications to sorafenib doses were implemented according to prescribing information in each region”; for each treatment group by geographic region, please provide mean (SD) and median (Q1, Q3) dose intensity, and rates of dose interruption and reduction.

Details of dose intensity and dose interruption/reduction for both treatment arms in REFLECT are presented for the Asia-Pacific region in Table 8 and the Western region in Table 9.

**Table 8: Dose intensity and dose interruptions/reductions in the Asia-Pacific region – Safety Analysis Set**

| | Lenvatinib N=321 | Sorafenib N=319 |
|---|---|---|
| Dose intensity (mg/day/patient) | | |
| Mean (SD) | 9.0 (2.64) | 661.3 (178.35) |
| Median | 8.0 | 776.2 |
| Q1, Q3 | 7.8, 11.9 | 493.0, 800.0 |
| Min, Max | 1.7, 12.1 | 126.3, 800.0 |
| Total number of patients with† | | |
| Dose interruption | 133 (41.4) | 107 (33.5) |

| | | |
|---|---|---|
| Dose reduction | 114 (35.5) | 119 (37.3) |
| Dose interruption or reduction | 178 (55.5) | 165 (51.7) |

†Patients may be counted in multiple categories
Abbreviations: SD, standard deviation.

**Table 9: Dose intensity and dose interruptions/reductions in the Western region – Safety Analysis Set**

| | Lenvatinib N=155 | Sorafenib N=156 |
|---|---|---|
| Dose intensity (mg/day/patient) | | |
| Mean (SD) | 10.2 (9.23) | 669.1 (162.45) |
| Median | 10.1 | 750.5 |
| Q1, Q3 | 8.0, 12.0 | 541.8, 800.0 |
| Min, Max | 3.1, 120.0 | 245.5, 800.0 |
| Total number of patients with† | | |
| Dose interruption | 68 (43.9) | 50 (32.1) |
| Dose reduction | 71 (45.8) | 67 (42.9) |
| Dose interruption or reduction | 99 (63.9) | 89 (57.1) |

†Patients may be counted in multiple categories
Abbreviations: SD, standard deviation.

A6. **Priority question**: Please provide equivalent adverse event data used in the model for the Western subgroup, unadjusted for treatment exposure (i.e. Grade 3 or above occurring in at least 5% of either group, plus diarrhoea, asthenia and fatigue).

Table 10 presents equivalent data for the Western region concerning AEs that were considered in the economic model. As requested, these are unadjusted for treatment exposure.

**Table 10: AEs considered in the economic model (Grade 3 or 4 TEAEs occurring in ≥5% of patients in either treatment arm of REFLECT, plus diarrhoea, asthenia and fatigue) for the Western region – Safety Analysis Set**

| AEs | Number of patients, n (%) | | Average number of episodes per patient | |
|---|---|---|---|---|
| | Lenvatinib (N=155) | Sorafenib (N=156) | Lenvatinib | Sorafenib |
| Hypertension | 36 (23.2) | 18 (11.5) | 1.22 | 1.22 |
| Weight decreased | 18 (11.6) | 11 (7.1) | 1.05 | 1.00 |
| Blood bilirubin increased | 11 (7.1) | 5 (3.2) | 1.18 | 1.00 |
| Proteinuria | 5 (3.2) | 6 (3.8) | 1.00 | 1.00 |
| Gamma-glutamyltransferase increased | 10 (6.5) | 9 (5.8) | 1.00 | 1.00 |
| Platelet count decreased | 2 (1.3) | 1 (0.6) | 1.00 | 1.00 |
| Aspartate aminotransferase increased | 7 (4.5) | 6 (3.8) | 1.00 | 1.33 |

| AEs | Number of patients, n (%) | | Average number of episodes per patient | |
|---|---|---|---|---|
| Diarrhoea | 8 (5.2) | 11 (7.1) | 1.00 | 1.00 |
| Fatigue† | 11 (7.1) | 11 (7.1) | 1.00 | 1.00 |
| Palmar-plantar erthrodysaesthesia syndrome | 6 (3.9) | 14 (9.0) | 1.17 | 1.36 |
| Asthenia | 9 (5.8) | 9 (5.8) | 1.11 | 1.00 |

Abbreviations: AE, adverse event.

A7. Please provide details of the number of patients in each group whose survival follow-up was discontinued by the sponsor (company submission, page 20), and reasons why.

The study protocol allowed for the sponsor to choose to discontinue survival follow-up after completion of the primary study analysis. However, the sponsor did not discontinue the survival follow-up of any patients in the trial. The end of study status for the Full Analysis Set is presented in Table 11 (and Table 14.1.1.6 of the CSR).

**Table 11: Patient disposition: end of study status – Full Analysis Set**

| | Lenvatinib (N=478) n (%) | Sorafenib (N=476) n (%) | Total (N=954) n (%) |
|---|---|---|---|
| On-study† | 109 (22.8) | 107 (22.5) | 216 (22.6) |
| Off-study‡ | 369 (77.2) | 369 (77.5) | 738 (77.4) |
| Reason for off-study | | | |
| Death | 351 (73.4) | 350 (73.5) | 701 (73.5) |
| Lost to follow-up | 5 (1.0) | 11 (2.3) | 16 (1.7) |
| Administrative/other | | | |
| Withdrawal of consent | 13 (2.7) | 8 (1.7) | 21 (2.2) |
| Others | 0 (0.0) | 0 (0.0) | 0 (0.0) |

† On-study refers to patients who were still on study treatment or in survival follow-up as of the cut-off date; ‡off-study refers to patients who were no longer followed up for survival as of the cut-off date.

A8. Please explain the clinical reasoning behind AFP levels being dichotomised at 200 ng/mL, and whether differences between mean and median baseline levels are clinically meaningful.

A cut-off level of 200 ng/mL has been used in a number of previously published studies to determine the prognostic value of AFP. Two multivariate analyses, using data from the sorafenib SHARP study (3) and pooled data from the SHARP and Asia-Pacific studies (4), dichotomised AFP levels at 200 ng/mL and found that AFP >200 ng/mL was a prognostic factor for poor overall survival. Furthermore, the EASL–EORTC guideline for the management of HCC advises the use of >200 ng/mL and/or >400 ng/mL as prognostic factors of poor outcome (5). In REFLECT, mean AFP levels at baseline were substantially higher than median levels in both treatment groups (median 89 ng/mL, mean 17096.5 ng/mL for overall trial population; CSR page 105, table 13). AFP levels can vary substantially from

patient to patient, with values ranging between 0 and 1567470 ng/ml at baseline in REFLECT (CSR page 106, table 13). The substantial difference between median and mean values at baseline reflects a small number of patients with very high baseline AFP levels, which skews the mean value to the far right of the AFP distribution.

A9. Please clarify whether there was an investigator assessment of PFS using RECIST v1.1 in REFLECT and, if there was, please provide the results.

An investigator assessment of PFS using RECIST 1.1 was not performed in the REFLECT study. All efficacy endpoints conducted for the primary analyses in REFLECT (with the exception of OS) were based on tumour response evaluations as determined by the investigator according to mRECIST for HCC for hepatic lesions. mRECIST was used as it more appropriately reflects changes in intrahepatic lesions by measuring only the viable portion of the lesions. Other differences between mRECIST and RECIST 1.1 include: a requirement for interval growth for new atypical hepatic lesions before considering them as unequivocal; and a requirement for cytological evidence of malignancy to consider ascites/effusions malignant given the underlying cirrhosis common in HCC patients (6). Therefore, the sponsor, together with the protocol steering committee, and following established scientific evidence at the time point of study inception, determined that mRECIST was the more appropriate criteria for investigator assessment of response in the REFLECT study.

**Literature searching**

A10. Please provide details of any clinical effectiveness studies that were excluded from the systematic literature review for the reason, “Studies where the primary endpoint was not met”, as listed in Table 1 of Appendix D.1.2.

Several investigational therapies have failed to meet the endpoints of non-inferiority or superiority for OS compared with sorafenib (sunitinib, brivanib, linifanib, sorafenib + erlotinib, and sorafenib + doxorubicin). The following studies concerning these therapies were excluded from the systematic literature review for this reason:


1. Cheng AL, Kang YK, Lin DY, Park JW, Kudo M, Qin S, et al. Sunitinib versus sorafenib in advanced hepatocellular cancer: results of a randomized phase III trial. J Clin Oncol. 2013;31(32):4067-75.

2. Johnson PJ, Qin S, Park JW, Poon RT, Raoul JL, Philip PA, et al. Brivanib versus sorafenib as first-line therapy in patients with unresectable, advanced hepatocellular carcinoma: results from the randomized phase III BRISK-FL study. J Clin Oncol. 2013;31(28):3517-24.

3. Cainap C, Qin S, Huang WT, Chung IJ, Pan H, Cheng Y, et al. Linifanib versus Sorafenib in patients with advanced hepatocellular carcinoma: results of a randomized phase III trial. J Clin Oncol. 2015;33(2):172-9.

4. Zhu AX, Rosmorduc O, Evans TR, Ross PJ, Santoro A, Carrilho FJ, et al. SEARCH: a phase III, randomized, double-blind, placebo-controlled trial of sorafenib plus erlotinib in patients with advanced hepatocellular carcinoma. J Clin Oncol. 2015;33(6):559-66.

5. Abou-Alfa GK, Niedzwieski D, Knox JJ, Kaubisch A, Posey J, Tan BR, et al. Phase III randomized study of sorafenib plus doxorubicin versus sorafenib in patients with advanced hepatocellular carcinoma (HCC): CALGB 80802 (Alliance). Journal of Clinical Oncology. 2016;34(15_suppl):4003-.

## Section B: Clarification on cost-effectiveness data

During preparation of responses to the ERG clarification questions, it was necessary to re-estimate statistical models used in the prediction of outcomes. The original multivariable analyses submitted to NICE contained Child-Pugh *score* as a continuous variable. All analyses used to generate the model results below are based on re-estimated statistical models which include Child-Pugh *class* (A vs B), which was considered more appropriate than the treatment of Child-Pugh as a continuous predictor. In addition, all scenarios presented incorporate corrected mortality costs as detailed in the response to B14. All other aspects of the analyses remain the same. Results of model modifications and scenario analyses are presented below in Table 12.

**Table 12: Results of model modifications and scenarios**

| # | Question | Scenario | Incremental costs | Incremental QALYs | NMB† |
|---|---|---|---|---|---|
| 1 | - | Original base-case | [redacted] | 0.17 | [redacted] |
| 2 | - | Revised base-case‡ | [redacted] | 0.18 | [redacted] |
| 3 | B1 | #2 & Independent assessment of PFS | [redacted] | 0.18 | [redacted] |
| 4 | B1 | # 2 & CGP approach | [redacted] | 0.17 | [redacted] |
| 5 | B1 | #3 & CGP approach | [redacted] | 0.17 | [redacted] |
| 6 | B2 | #7 & Western subgroup | [redacted] | 0.08 | [redacted] |
| 7 | B3 | #2 & adjustment for post-progression therapies & exclude post-progression drug costs | [redacted] | 0.23 | [redacted] |
| 8 | B5 | #2 & Include grade 5 AEs | [redacted] | 0.18 | [redacted] |
| 9 | B6 | #2 & Use on/off treatment utility values | [redacted] | 0.17 | [redacted] |
| 10 | B10 | #2 & Include all post-progression therapies | [redacted] | | [redacted] |
| 11 | B11 | #2 & Exclude post-progression drug costs | [redacted] | 0.18 | [redacted] |
| 12 | B15 | #2 & Drug wastage included | [redacted] | 0.18 | [redacted] |
| 13 | B17 | #2 & Mortality costs excluded | [redacted] | 0.18 | [redacted] |

Abbreviations: AEs, adverse events; CGP, corrected group prognosis; NMB, net monetary benefit; PFS, progression-free survival; QALYs, quality-adjusted life-years.

† NMB willingness-to-pay threshold of £50,000 per QALY
‡ Including adjustment to the treatment of Child-Pugh in models of PFS and OS, and corrected mortality costs as detailed in the response to B14. Please note that both the original and revised base-case use the mean of covariates approach.

**Survival analysis**

B1. **Priority question**: For the covariate adjusted survival analyses for independent imaging review (IIR) PFS and OS using the full population, please provide an additional set of analyses using the corrected group prognosis (CGP) method and include an option in the economic model to alternate between that and the mean of covariates (MoC) method.

The base-case analysis uses investigator assessment of PFS (using FDA censoring rules – see Question A1); this was the pre-specified primary endpoint and forms the basis of the key secondary endpoint submitted as part of regulatory approval. For consistency, scenarios presented here are based on investigator assessment of PFS. Note that results presented in Table 12 demonstrate that the effect of using independent assessment of PFS is minimal.

An option is included in the electronic model to either use the MoC method or the CGP method – this option can be found in Row 23 of the 'Model Parameters' sheet.

The CGP method passes each patient from REFLECT, with the exception of those with missing AFP data (N = 4), through the multivariable PFS and OS models, and averages the twelve survival curves – six parametric distributions for each of PFS and OS – over the 950 patients considered.

This procedure has been conducted in advance, and the averaged survival curves can be found in the 'Appendix Extrapolation (CGP)' sheet. The Visual Basic for Applications (VBA) code used to generate these averaged curves can be found in the module labelled 'CGP'. Two additional averaged curves are included to support a scenario based on independent assessment of PFS (see Columns AL and AM on the 'Appendix Extrapolation (CGP)' sheet). Results using the CGP method are presented in Table 12.

B2. **Priority question**: Please perform adjusted and unadjusted survival analyses for IIR PFS and OS using the Western population subgroup of the REFLECT trial and update the economic model to include this subgroup analysis as an option. For the adjusted analyses, please use both the CGP and MoC methods and include both as options in the economic model.

In the REFLECT study, although the HR for OS (lenvatinib:sorafenib) was higher in the Western region (HR 1.08 [95% CI 0.82, 1.42]) than in the overall population (HR 0.92 [95% CI 0.79, 1.06]) and the Asia-Pacific population (HR 0.86 [95% CI 0.72, 1.02]) , the difference in results between the regions was due to better performance of sorafenib in the Western region, rather than worse performance by lenvatinib. The median OS for lenvatinib in the Western region was 13.6 months (95% CI: 11.5, 17.7) and 13.5 months (95% CI: 11.7, 15.1) in the Asia-Pacific region; 95% CI: 11.7, 15.1). In contrast, the median OS for sorafenib was 14.2 months (95% CI: 11.9, 18.0) in the Western region; longer than that observed with

sorafenib in the Asia-Pacific region (11.0 months; 95% CI: 9.6, 12.5). The different HR for OS in the Western region is likely accounted for by observed imbalances in post-progression treatment and baseline imbalances in this region. A higher percentage of patients in the sorafenib arm compared with the lenvatinib arm received post-progression anticancer therapy with anticancer medications (including investigational agents) and anticancer procedures (including TACE) (45.2% vs 28.0%, respectively). Further, imbalances in important prognostic and predictive factors (AFP ≥200 ng/mL and Hepatitis C aetiology), while observed for the entire study, were more prevalent in the Western region. The number of patients with AFP ≥ 200 ng/mL was 30% greater in the lenvatinib arm than in the sorafenib arm in the Western region. When stratified by geographic region, the HR favoured lenvatinib in both Western and Asia-Pacific regions when adjusted for post-treatment anti-cancer treatment. In the Western region, the adjusted HR (95% CI) was [redacted] compared with 1.08 (0.82, 1.42) in the unadjusted analysis.

A forest plot of HRs for OS for lenvatinib vs sorafenib (presented in Company Submission Form B Appendices, Figure 3) showed that the CIs for the Western and Asia-Pacific regions broadly overlap, and also overlap the point estimate for the overall population. Tests of interaction show no evidence of treatment-effect modifiers by region ([redacted]) for OS.

In contrast, the results of all secondary efficacy endpoints which are not influenced by the effects of post-progression anti-cancer therapy (progression-free survival [PFS], time to progression [TTP], and objective response rate [ORR]) were very consistent for both the lenvatinib and sorafenib arms in both regions (Western and Asia-Pacific) and the overall study population.

Since the treatment results for all subgroups of REFLECT were within the expected random variation, and there was no interaction of treatment and region, we conclude that the OS results for the Western region are, in fact, consistent with those of the overall study.
We therefore believe that the overall trial effect (rather than the subgroup-specific treatment effect) is the most representative treatment effect for patients in the Western, and other, subgroups. Furthermore, REFLECT was not powered to detect statistically significant differences in OS in the Western subgroup. We believe any differential interpretation of treatment effect in subgroups should be undertaken with extreme caution; NICE committees have drawn similar conclusions (i.e. that the central estimate of effect should be accepted for a geographical subgroup), in similar contexts, previously (7).

In order to answer the ERG's request a scenario analysis has been performed based on multivariable statistical analyses estimated using the overall trial population. Specifically, the analysis is based on:

- Multivariable model of PFS (investigator assessed) estimated in the overall population (including region as a covariate)

- Multivariable model of OS estimated in the overall population (including region as a covariate), which further adjusts for the use of post-progression therapies (see also B3)

    - This model was used to reflect the differences in post-progression therapy use observed in REFLECT between the trial arms in the Western population.

- For consistency, the costs of post-progression therapies were removed in this analysis (see also B3)

In order to estimate cost-effectiveness in the Western subgroup, the characteristics of the Western population only are used in the prediction of model outcomes. Specifically, whilst the base-case economic evaluation uses the mean characteristics of the overall population to generate model predictions (using the MoC method), the Western subgroup analysis uses the mean characteristics of the Western subgroup only.

This approach ensures efficient use of the available data within the economic model whilst still allowing prediction of outcomes for the Western population. Utility values specific to the Western population are also used in this scenario analysis (see question B7). Results are presented in Table 12.

B3. **Priority question**: Please provide an additional set of OS parametric models adjusted for post-progression anti-cancer therapies, for the full population and the Western subgroup, and include these as options in the economic model.

An option is included in the economic model in which post-progression therapies are adjusted for in the full population, in addition to previously included covariates. This is the fourth option in the dropdown menu on Row 22 of the 'Model Parameters' sheet. In the scenario in which this analysis is presented – see Row 92 on the 'Sensitivity analysis' sheet – the costs of post-progression therapies are excluded for consistency. Results are presented in Table 12.

Please see Question B2 for the rationale for why further scenarios are not considered for the Western subgroup.

**Data**

B4. **Priority question:** For IIR PFS, OS, and time to treatment discontinuation, please provide the following for both the full population and Western subgroup of the REFLECT trial:

a. individual times for each event/censor;

b. reasons for censoring (depending on the response to A1, please consider providing an alternative IIR PFS analysis that does not censor for "no progression at the time of treatment discontinuation").

Please see separate CSV file titled 'REFLECT_outcomes_ERG_clarification.csv'. Please note that all censoring for TTD was a result of subjects being on treatment at the time of trial cut-off.

**Adverse events**

B5. Please clarify why grade 5 adverse events are not included in the economic model. Also, please provide a scenario analysis that includes grade 5 treatment-emergent adverse events (regardless of incidence level) using appropriate costs.

Grade 5 adverse events were initially not included in the economic model to avoid double counting the costs of mortality. An option is now included in the economic model to include the cost of one hospitalisation for each grade 5 adverse event, regardless of incidence level. This option can be found on Row 38 of the 'Model Parameters' sheet.

The effect of including grade 5 adverse events is presented in Table 12.

**Health-related quality of life**

B6. Given that disutilities associated with adverse events were not explicitly modelled, please consider providing adjusted and unadjusted treatment-specific utilities for the on- and off-treatment periods for those who are progression-free, for both the full population and the Western subgroup in the REFLECT trial. Please also add these as options in the economic model.

An analysis was conducted to provide adjusted and unadjusted treatment-specific utilities for the on- and off-treatment periods for those who are progression-free, for both the full population and the Western subgroup (see tables below). In each analysis, the difference between lenvatinib and sorafenib was not found to be statistically significant (see Table 13, Table 14, Table 15 and Table 16). It was therefore determined that the average values across the two trial arms for each of the on-treatment and off-treatment states were the most appropriate values to use (see Table 17 and Table 18). The values presented in Table 17 were included as an alternative option in the economic model. Results are presented in Table 12.

Please note that that the post-progression utility value remains the same as presented in the base-case economic evaluation.

**Table 13: Utility analysis in the pre-progression, on treatment state (overall population)**

| Effect | Level | Parameter estimate | Std. Error | t-value | p-value | 95% CI | |
|---|---|---|---|---|---|---|---|
| Intercept | | 0.231 | 0.0454 | 5.0944 | 0.0000 | 0.1423 | 0.3207 |
| trtpn | Lenvatinib | 0.008 | 0.0128 | 0.6627 | 0.5077 | -0.0167 | 0.0337 |
| trtpn | Sorafenib | 0.000 | . | . | . | . | . |
| BASE | | 0.662 | 0.0347 | 19.0733 | 0.0000 | 0.5941 | 0.7304 |
| AGE | | -0.001 | 0.0005 | -1.5518 | 0.1211 | -0.0019 | 0.0002 |
| SEX | F | -0.013 | 0.0175 | -0.7444 | 0.4568 | -0.0474 | 0.0213 |
| SEX | M | 0.000 | . | . | . | . | . |
| PRCANPR | No | -0.001 | 0.0147 | -0.0634 | 0.9495 | -0.0297 | 0.0279 |
| PRCANPR | Yes | 0.000 | . | . | . | . | . |
| STRATA1N | 1 | 0.010 | 0.0146 | 0.6530 | 0.5139 | -0.0191 | 0.0382 |
| STRATA1N | 2 | 0.000 | . | . | . | . | . |
| STRATA3N | 0 | 0.013 | 0.0143 | 0.9030 | 0.3668 | -0.0151 | 0.0408 |
| STRATA3N | 1 | 0.000 | . | . | . | . | . |

Abbreviations: CI, confidence interval; Std. Error, standard Error.

**Table 14: Utility analysis in the pre-progression, off treatment state (overall population)**

| Effect | Level | Parameter estimate | Std. Error | t-value | p-value | 95% CI | |
|---|---|---|---|---|---|---|---|
| Intercept | | 0.128 | 0.3617 | 0.3541 | 0.7249 | -0.6004 | 0.8566 |
| trtpn | Lenvatinib | 0.006 | 0.0962 | 0.0663 | 0.9474 | -0.1875 | 0.2002 |
| trtpn | Sorafenib | 0.000 | . | . | . | . | . |
| BASE | | 0.761 | 0.3044 | 2.4997 | 0.0161 | 0.1478 | 1.3741 |
| AGE | | -0.000 | 0.0037 | -0.1015 | 0.9196 | -0.0077 | 0.0070 |
| SEX | F | -0.041 | 0.1783 | -0.2291 | 0.8198 | -0.4000 | 0.3183 |
| SEX | M | 0.000 | . | . | . | . | . |
| PRCANPR | No | 0.034 | 0.0994 | 0.3457 | 0.7312 | -0.1659 | 0.2346 |
| PRCANPR | Yes | 0.000 | . | . | . | . | . |
| STRATA1N | 1 | -0.086 | 0.0941 | -0.9105 | 0.3674 | -0.2752 | 0.1038 |
| STRATA1N | 2 | 0.000 | . | . | . | . | . |
| STRATA3N | 0 | -0.047 | 0.1015 | -0.4676 | 0.6423 | -0.2520 | 0.1570 |
| STRATA3N | 1 | 0.000 | . | . | . | . | . |

Abbreviations: CI, confidence interval; Std. Error, standard error.

**Table 15: Utility analysis in the pre-progression, on treatment state (Western subgroup)**

| Effect | Level | Parameter estimate | Std. Error | t-value | p-value | 95% CI | |
|---|---|---|---|---|---|---|---|

| Effect | Level | Parameter estimate | Std. Error | t-value | p-value | 95% CI | |
|---|---|---|---|---|---|---|---|
| Intercept | | 0.314 | 0.0771 | 4.0669 | 0.0001 | 0.1617 | 0.4656 |
| trtpn | Lenvatinib | 0.004 | 0.0215 | 0.1645 | 0.8695 | -0.0387 | 0.0458 |
| trtpn | Sorafenib | 0.000 | . | . | . | . | . |
| BASE | | 0.557 | 0.0558 | 9.9854 | 0.0000 | 0.4473 | 0.6672 |
| AGE | | -0.001 | 0.0009 | -1.3313 | 0.1844 | -0.0031 | 0.0006 |
| SEX | F | -0.020 | 0.0271 | -0.7413 | 0.4592 | -0.0735 | 0.0333 |
| SEX | M | 0.000 | . | . | . | . | . |
| PRCANPR | No | 0.009 | 0.0225 | 0.3845 | 0.7010 | -0.0357 | 0.0531 |
| PRCANPR | Yes | 0.000 | . | . | . | . | . |
| STRATA1N | 1 | 0.000 | . | . | . | . | . |
| STRATA3N | 2 | 0.051 | 0.0238 | 2.1561 | 0.0321 | 0.0044 | 0.0981 |
| STRATA3N | 0 | 0.000 | . | . | . | . | . |

Abbreviations: CI, confidence interval; Std. Error, standard error.

**Table 16: Utility analysis in the pre-progression, off treatment state (Western subgroup)**

| Effect | Level | Parameter estimate | Std. Error | t-value | p-value | 95% CI | |
|---|---|---|---|---|---|---|---|
| Intercept | | 0.613 | 0.2922 | 2.0981 | 0.0577 | -0.0236 | 1.2495 |
| trtpn | Lenvatinib | -0.150 | 0.0916 | -1.6396 | 0.1270 | -0.3496 | 0.0494 |
| trtpn | Sorafenib | 0.000 | . | . | . | . | . |
| BASE | | 0.088 | 0.2775 | 0.3179 | 0.7560 | -0.5163 | 0.6927 |
| AGE | | -0.001 | 0.0032 | -0.3601 | 0.7250 | -0.0081 | 0.0058 |
| SEX | F | -0.060 | 0.1237 | -0.4835 | 0.6374 | -0.3292 | 0.2096 |
| SEX | M | 0.000 | . | . | . | . | . |
| PRCANPR | No | 0.101 | 0.0820 | 1.2302 | 0.2422 | -0.0778 | 0.2795 |
| PRCANPR | Yes | 0.000 | . | . | . | . | . |
| STRATA1N | 1 | 0.000 | . | . | . | . | . |
| STRATA3N | 2 | 0.245 | 0.1049 | 2.3395 | 0.0374 | 0.0169 | 0.4740 |
| STRATA3N | 0 | 0.000 | . | . | . | . | . |

Abbreviations: CI, confidence interval; Std. Error, standard error.

**Table 17: Utility analysis based on treatment status (overall population)**

| | Statistic | Total study population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| **Baseline** | n | 885 | 443 | 442 |
| | Mean (SE) | 0.833 (0.0067) | 0.830 (0.0100) | 0.837 (0.0090) |
| | Adj Mean (SE)[3] | N/A | 0.789 (0.0114) | 0.796 (0.0115) |
| **Progression-Free Survival** | n | 846 | 417 | 429 |
| | Mean (SE) | 0.747 (0.0079) | 0.752 (0.0105) | 0.742 (0.0117) |

| | Statistic | Total study population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| **(On Treatment)[1]** | Adj Mean (SE)[3] | N/A | 0.748 (0.0116) | 0.739 (0.0114) |
| **Progression-Free Survival (Off Treatment)[2]** | n | 53 | 27 | 26 |
| | Mean (SE) | 0.689 (0.0430) | 0.713 (0.0544) | 0.664 (0.0677) |
| | Adj Mean (SE)[3] | N/A | 0.700 (0.0888) | 0.694 (0.1100) |

[1] The average of all post-baseline pre-progression on-treatment EQ-5D HUI scores among all patients progression free.
[2] The average of all post-baseline pre-progression off-treatment (after last dose of study medication) EQ-5D HUI scores among all patients progression free.
[3] Least-squared means adjusted for prior treatment, age, sex, geographical region, baseline EQ-5D HUI score, and baseline ECOG-PS.
Program: (t_ea_hui_MM_NICE.sas) (22MAR18:12:34:20)
Analysis datasets: adqs, adttdef
Abbreviations: Adj, Adjusted; ECOG PS, Eastern Cooperative Oncology Group Performance Status; EQ-5D, EuroQol Five Dimension Health Survey; SE, Standard Error; N/A, Not Applicable; HUI, Health Utility Index.

**Table 18: Utility analysis based on treatment status (Western population)**

| | Statistic | Western population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| **Baseline** | n | 273 | 138 | 135 |
| | Mean (SE) | 0.780 (0.0125) | 0.778 (0.0176) | 0.781 (0.0179) |
| | Adj Mean (SE)[3] | N/A | 0.757 (0.0199) | 0.763 (0.0203) |
| **Progression-Free Survival (On Treatment)[1]** | n | 266 | 127 | 139 |
| | Mean (SE) | 0.695 (0.0134) | 0.701 (0.0177) | 0.689 (0.0200) |
| | Adj Mean (SE)[3] | N/A | 0.697 (0.0177) | 0.693 (0.0174) |
| **Progression-Free Survival (Off Treatment)[2]** | n | 19 | 11 | 8 |
| | Mean (SE) | 0.714 (0.0471) | 0.654 (0.0667) | 0.797 (0.0561) |
| | Adj Mean (SE)[3] | N/A | 0.606 (0.0606) | 0.757 (0.0938) |

[1] The average of all post-baseline pre-progression on-treatment EQ-5D HUI scores among all patients progression free.
[2] The average of all post-baseline pre-progression off-treatment (after last dose of study medication) EQ-5D HUI scores among all patients progression free.
[3] Least-squared means adjusted for prior treatment, age, sex, geographical region, baseline EQ-5D HUI score, and baseline ECOG-PS.
Program: (t_ea_hui_MM_NICE.sas) (22MAR18:12:34:20)
Analysis datasets: adqs, adttdef
Abbreviations: Adj, Adjusted; ECOG PS, Eastern Cooperative Oncology Group Performance Status; EQ-5D, EuroQol Five Dimension Health Survey; SE, Standard Error; N/A, Not Applicable.

B7. Please provide Table 32 of the company submission (Summary of EQ-5D data) for the Western subgroup.

Please find a version of Table 32 of the company submission for the Western subgroup below (Table 19). These utility values are used in the scenario analysis based on the Western subgroup, as discussed in Question B2.

**Table 19: Utility analysis based on the Western subgroup**

| | | Full Western subgroup | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| Baseline | N | 285 | 144 | 141 |
| | Mean (SE) | 0.777 (0.0122) | 0.773 (0.0173) | 0.781 (0.0173) |
| | Adjusted mean (SE) | N/A | 0.756 (0.0193) | 0.766 (0.0198) |
| Progression-free | N | 268 | 129 | 139 |
| | Mean (SE) | 0.693 (0.0134) | 0.697 (0.0176) | 0.689 (0.0199) |
| | Adjusted mean (SE) | N/A | 0.690 (0.0176) | 0.691 (0.0174) |
| Progressed | N | 229 | 116 | 113 |
| | Mean (SE) | 0.633 (0.0189) | 0.652 (0.0251) | 0.614 (0.0285) |
| | Adjusted mean (SE) | N/A | 0.656 (0.0272) | 0.613 (0.0275) |

Abbreviations: N/A, not applicable; SE, standard error.

B8. Please clarify how covariates for the linear mixed model were chosen. If AFP levels and HCC aetiology were not included as covariates, please justify their exclusion.

The covariates for the linear mixed model were non-systematically pre-specified prior to commencement of the analysis of REFLECT. The economic model does not assume differences in HRQOL between treatment arm (for a given health state).

B9. Please provide further details on the steps taken to adjust mean utility values, including the coefficients and p-values for the covariates included in the linear mixed model.

Please find the coefficients and p-values for each covariate included in the linear mixed model in Table 20.

**Table 20: Linear mixed model outputs**

| Effect | | Estimate | SE | t | P>\|t\| | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|---|---|---|
| **Model: Baseline** | | | | | | | |
| Intercept | | 0.729 | 0.0383 | 19.03 | <.0001 | 0.6538 | 0.8042 |
| Treatment | Lenvatinib | -0.014 | 0.0127 | -1.11 | 0.2679 | 0.0390 | 0.0108 |
| | Sorafenib | 0 | | | | | |
| Region | Asia-Pacific | 0.072 | 0.0141 | 5.11 | <.0001 | 0.0444 | 0.0997 |
| | Western | 0 | | | | | |
| Baseline ECOG-PS | 0 | 0.114 | 0.0134 | 8.52 | <.0001 | 0.0877 | 0.1402 |
| | 1 | 0 | | | | | |
| Age | | -0.0002 | 0.0005 | -0.35 | 0.7284 | -0.0013 | 0.0009 |
| Sex | F | -0.036 | 0.0174 | -2.04 | 0.0412 | -0.0698 | -0.0014 |

| Effect | | Estimate | SE | t | P>\|t\| | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|---|---|---|
| | M | 0 | | | | | |
| Previous Anti-cancer procedure | No | 0.011 | 0.0144 | 0.73 | 0.4642 | -0.0178 | 0.0389 |
| | Yes | 0 | | | | | |
| **Model: Progression-Free Survival** | | | | | | | |
| Intercept | | 0.232 | 0.0455 | 5.10 | <.0001 | 0.1426 | 0.3212 |
| Treatment | Lenvatinib | 0.008 | 0.0128 | 0.63 | 0.5263 | -0.0170 | 0.0333 |
| | Sorafenib | 0 | | | | | |
| Region | Asia-Pacific | 0.010 | 0.0146 | 0.68 | 0.4973 | -0.0187 | 0.0385 |
| | Western | 0 | | | | | |
| Baseline ECOG-PS | 0 | 0.016 | 0.0142 | 1.09 | 0.2763 | -0.0124 | 0.0435 |
| | 1 | 0 | | | | | |
| Age | | -0.001 | 0.0005 | -1.74 | 0.0820 | -0.0020 | 0.0001 |
| Sex | F | -0.013 | 0.0176 | -0.77 | 0.4429 | -0.0480 | 0.0210 |
| | M | 0 | | | | | |
| Previous Anti-cancer procedure | No | -0.001 | 0.0146 | -0.08 | 0.9355 | 0.0299 | 0.0275 |
| | Yes | | | | | | |
| Baseline EQ-5D HUI Score | | 0.665 | 0.0347 | 19.16 | <.0001 | 0.5969 | 0.7332 |
| **Model: Progression** | | | | | | | |
| Intercept | | 0.054 | 0.0755 | 0.72 | 0.4715 | -0.0938 | 0.2025 |
| Treatment | Lenvatinib | 0.009 | 0.0208 | 0.42 | 0.6756 | -0.0322 | 0.0497 |
| | Sorafenib | 0 | | | | | |
| Region | Asia-Pacific | -0.012 | 0.0241 | -0.49 | 0.6224 | -0.0591 | 0.0354 |
| | Western | 0 | | | | | |
| Baseline ECOG-PS | 0 | 0.0300 | 0.0230 | 1.30 | 0.1927 | -0.0152 | 0.0752 |
| | 1 | 0 | | | | | |
| Age | | 0.0001 | 0.0009 | 0.11 | 0.9135 | -0.0017 | 0.0019 |
| Sex | F | -0.066 | 0.0280 | -2.37 | 0.0183 | -0.1214 | -0.0113 |
| | M | 0 | | | | | |
| Previous Anti-cancer procedure | No | 0.007 | 0.0243 | 0.29 | 0.7695 | -0.0405 | 0.0547 |
| | Yes | 0 | | | | | |
| Baseline EQ-5D HUI Score | | 0.741 | 0.0552 | 13.42 | <.0001 | 0.6322 | 0.8489 |

Abbreviations: CI, confidence interval; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; EQ-5D, EuroQol Five Dimension Health Survey; HUI, Health Utility Index; SE, standard error.

**Resource use and costs**

B10. Please provide a scenario, applying appropriate resources and costs, which includes the full list of post-progression therapies and procedures that patients in REFLECT received, including treatments without marketing authorisation in the UK. Please provide results for the full population and Western subgroup and include this as an option in the economic model.

A dropdown menu is now included in the economic model to select to either include only regorafenib and sorafenib as post-progression therapies, or to include all post-progression therapies observed in ≥1% of either arm in REFLECT. This option can be found on Row 35 of the 'Model Parameters' sheet. Results are presented in Table 12.

Post-progression therapies specific to the Western subgroup are not included in the scenario analysis discussed in Question B2; this is because the analysis controls for post-progression therapies, and so it would not be consistent to apply different post-progression drug costs by arm.

Drug therapies observed in ≥1% of patients in either arm of REFLECT include:

- Capecitabine
- Fluorouracil
- Gemcitabine
- Gimeracil with oteracil potassium/tegafur
- Uftoral
- Doxorubicin
- Thalidomide
- Cabozantinib
- Cisplatin
- Nivolumab
- Oxaliplatin
- Protein kinase inhibitors
- Regorafenib
- Sorafenib
- Folinic acid
- Tivantinib
- Investigational drug

Three post-progression drug therapies that were used in ≥1% of patients in either arm in REFLECT were not costed in the economic model. The reasons for each exclusion are presented in Table 21.

**Table 21: Rationale for exclusion of individual post-progression drug therapies**

| Excluded drug | Reason |
|---|---|
| Uftoral | Costs for this therapy are not available from the BNF, MIMS or eMC. |
| Tivantinib | Costs for this therapy are not available from the BNF, MIMS or eMC. |
| Investigational drug | It is unclear which costs should be applied for this drug. |

Abbreviations: BNF, British National Formulary; eMC, electronic Medicines Compendium; MIMS, Monthly Index of Medical Statistics.

Drug costs were taken from the BNF; the cheapest cost per mg was assumed where alternative costs were available. This was considered to be a conservative assumption, given that higher rates of post-progression drug use were observed in the sorafenib arm of REFLECT.

Procedures observed in ≥1% of patients in either arm of REFLECT include:

- High frequency ablation
- Radiotherapy
- Radiotherapy to bone
- Radiotherapy to brain
- Radiotherapy to liver
- Radiotherapy to lymph nodes
- Regional chemotherapy
- Transcatheter arterial chemoembolisation

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Costs for procedures were taken from NHS reference costs.

The proportions of patients used in the model were calculated as the number of patients using each therapy divided by the number of patients experiencing a PFS event (i.e. either progression or death). This is considered appropriate, given that the costs of post-progression therapies are applied in the cycle in which patients leave the progression-free state.

The durations of use of each of the post-progression therapies were generated using patient-level data from REFLECT; the extent of missing data for each of these therapies is presented in Table 22.

**Table 22: Data availability in REFLECT, by post-progression therapy**

| | Lenvatinib | | Sorafenib | | All | |
|---|---|---|---|---|---|---|
| **Missing Pattern** | **Start Date** | **End Date** | **Start Date** | **End Date** | **Start Date** | **End Date** |
| **Capecitabine** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Fluorouracil** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Gemcitabine** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Gimeracil with Oteracil Potassium/Tegafur** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Lenvatinib | | Sorafenib | | All | |
|---|---|---|---|---|---|---|
| **Missing Pattern** | **Start Date** | **End Date** | **Start Date** | **End Date** | **Start Date** | **End Date** |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Uftoral** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Doxorubicin** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Thalidomide** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Cabozantinib** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Cisplatin** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Lenvatinib | | Sorafenib | | All | |
|---|---|---|---|---|---|---|
| **Missing Pattern** | **Start Date** | **End Date** | **Start Date** | **End Date** | **Start Date** | **End Date** |
| **Nivolumab** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Oxaliplatin** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Protein kinase inhibitors** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Regorafenib** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Sorafenib** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Folinic acid** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Lenvatinib | | Sorafenib | | All | |
|---|---|---|---|---|---|---|
| **Missing Pattern** | **Start Date** | **End Date** | **Start Date** | **End Date** | **Start Date** | **End Date** |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Tivantinib** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| **Investigational drug** | | | | | | |
| | N* = [redacted] | | N* = [redacted] | | N* = [redacted] | |
| Completely Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Month and Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Day Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Non-Missing | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

*N is number of subjects who took the respective post-progression anticancer medication in the respective arm and is used as the denominator for calculating the percentages for the corresponding medication and the arm.

B11. Please provide a scenario which excludes post-progression therapy costs (sorafenib and regorafenib). Please provide results for the full population and Western subgroup and include this as an option in the economic model.

A dropdown menu is now included in the economic model to select to either include or exclude the costs of post-progression therapies for the full population. This option can be found on Row 35 of the 'Model Parameters' sheet and the results are presented in Table 12.

Please see Question B2 for the rationale for why further scenarios are not considered for the Western subgroup.

B12. Please clarify why no monitoring or testing costs are applied to patients who receive post-progression therapies (sorafenib or regorafenib).

No monitoring or testing costs were applied to patients who receive post-progression therapies to avoid double counting, given that monitoring and testing costs are included in the post-progression state.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B13. Please clarify why mortality costs and post-progression drug costs are applied using the OS curves (column G in Appendix Transition (LEN) and (SOR)) rather than the health state values (column L in Appendix Transition (LEN) and (SOR)) in which a half cycle correction has been applied. The ERG considers there to be an inconsistent application of the half-cycle correction for each cost category.

The calculation of mortality costs (Column V in the 'Appendix Transition' sheets) is based on the OS curve (Column G) rather than the 'Dead' health state (Column L). A simplifying assumption of the model is that all mortality costs are applied to patients in the cycle in which they die; this requires that we know how many people have died within a particular cycle.

For the first cycle, this is calculated as the proportion alive at time point 0 minus the proportion alive at time point 1; subsequent cycles are calculated as the proportion alive at time point x-1 minus the proportion alive at time point x. Since the values at each time point, as opposed to the values at each cycle midpoint, are required, Column G is used rather than Column L.

Similarly, post-progression therapy costs are assumed – for simplicity – to be applied in the cycle in which patients leave the progression-free state. It is therefore necessary to use the PFS curve (Column F) as opposed to the half-cycle corrected 'progression-free' health state (Column J).

B14. The ERG notes that Georghiou and Bardsley (2014) report end-of-life costs using a 2010/11 cost year rather than a 2013/14 cost year assumed by the company. Please consider inflating this cost appropriately.

This cost has now been inflated from 2010/11 to 2016/17 and updated in the economic model.

B15. Please amend the drug wastage scenario to account for the wastage associated with post-progression drug costs, as the current approach is inconsistent.

In the scenario that accounts for drug wastage, all post-progression drugs are now also associated with an additional 7 days of drug costs. Results are presented in Table 12.

B16. In the drug wastage scenario, please clarify why primary drug wastage costs are informed by the TTD curves rather than the on-treatment health states in which a half-cycle correction has been applied.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

As in Question B13, the costs of drug wastage are assumed to apply in the cycle in which patients discontinue. It is therefore necessary to use the TTD curve (Column E) as opposed to the half-cycle corrected 'on treatment' health state (Column I).

B17. Please explain the difference between the social care costs applied to patients with progressed disease and end-of-life care costs. The ERG is concerned end-of-life care costs have been double counted.

There is a risk that there may be some double counting for all components of end of life care costs, given that each of these aspects of resource use is costed in the progressed health state. However, a scenario is presented in which end of life care costs are excluded; this scenario is associated with a negligible (0.5%) decrease in the net monetary benefit associated with lenvatinib (Table 12).

**Systematic literature review**

B18. Please clarify why TA189 and TA474 were not included as studies reporting relevant resource and cost use data in the literature review.

A single systematic literature review (SLR) was conducted to identify both cost-effectiveness/utility and cost/resource use studies. TA189 and TA474 were identified and reported as cost-effectiveness studies but were not then reported again under cost/resource studies. Both data sources were used to inform resource use inputs for the submission, through use of the resource use survey that was presented in TA189 and updated in TA474 (8).

B19. Please clarify why the NICE submission for regorafenib (reference 69 in company submission) in Table 34 of the company submission was not included in the literature review for studies reporting HRQOL data.

The NICE submission for regorafenib was not identified in the grey literature search in error. However, this omission does not impact on the economic strategy as HRQoL data from the REFLECT study was used to inform the economic model.

B20. Please clarify why the Cochrane library was restricted by intervention (Lenvatinib or Sorafenib) for the HRQOL search.

The Cochrane library HRQoL search was not restricted by intervention; this was included in the submission in error.

B21. Figure 5 in Appendix H does not include the number of records identified via grey literature searches. Please add this information to figure 5.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The existing figure presented within the submission is correct based on our search. As highlighted in question B19, the regorafenib NICE submission was not identified in error.

B22. Please provide excluded studies lists with reasons for exclusion for the cost-effectiveness, resource use and HRQOL searches.

A list of studies excluded at full-text review stage (N=17) from the economic SLR which included cost-effectiveness/utility and costs/resource use studies is presented in Table 23. A list of studies excluded at full-text review stage (N=61) from the HRQoL SLR is presented in Table 24.

**Table 23: Excluded records – economic SLR (N=17)**

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| 1 | Reiss et al, 2017 | Starting dose of sorafenib for the treatment of hepatocellular carcinoma: A retrospective, multi-institutional study | Journal of Clinical Oncology. 2017; 35 (31): 3575-3581 | Outcomes |
| 2 | Venkatachalam et al, 2017 | Cost of treatment-related adverse events (TRAES) in second-line (2L) advanced hepatocellular carcinoma (AHCC): Match adjusted indirect comparison (MAIC) of nivolumab and regorafenib | Value in Health. 2017; 20 (9): A502-A503 | Population |
| 3 | Pollom et al, 2017 | Cost-effectiveness of Stereotactic Body Radiation Therapy versus Radiofrequency Ablation for Hepatocellular Carcinoma: A Markov Modeling Study | Radiology. 2017; 283(2):460-468 | Population |
| 4 | Pollom et al, 2016 | Cost-effectiveness of local therapies for inoperable, localized hepatocellular carcinoma | International Journal of Radiation Oncology. 2016; 96 (2 Supplement 1): E138 | Population |
| 5 | Hung et al, 2017 | Escalating health care expenditures in cancer decedents' last year of life: A decade of evidence from a retrospective population-based cohort study in Taiwan | Oncologist. 2017; 22 (4):460-469 | Population |
| 6 | Tang et al 2016 | Combination Therapy of Radiofrequency Ablation and Transarterial Chemoembolization for Unresectable Hepatocellular Carcinoma | Medicine (United States). 2016; 95 (20): e3754 | Population |
| 7 | Xu et al, 2014 | Hospitalizations and costs associated with hepatitis c and advanced liver | Health Affairs. 2014; 33 | Population |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| | | disease continue to increase | (10):1728-1735) | |
| 8 | Vadot et al, 2013 | Result and cost of hepatic chemoembolisation with drug eluting beads in 21 patients | Diagnostic and Interventional Imaging. 2013; 94 (1):53-59) | Population |
| 9 | Zalesak et al, 2012 | Recent trends in the prevalence of HCV-infected medicare patients and the associated cost burden | Journal of Managed Care Pharmacy. 2012; 18(2):195 | Population |
| 10 | Bernal et al, 2008 | International Atomic Energy Agency-Sponsored Multination Study of Intra-Arterial Rhenium-188-Labeled Lipiodol in the Treatment of Inoperable Hepatocellular Carcinoma: Results With Special Emphasis on Prognostic Value of Dosimetric Study | Seminars in Nuclear Medicine. 2008; 38(2): S40-S45 | Outcomes |
| 11 | Sundram et al, 2004 | Preliminary results of transarterial rhenium-188 HDD lipiodol in the treatment of inoperable primary hepatocellular carcinoma | European Journal of Nuclear Medicine and Molecular Imaging. 2008; 31(2):250-257 | Population |
| 12 | Hidajat et al, 2004 | Repetitive transarterial chemoembolization (rTACE) of hepatocellular carcinoma: Comparisons between an arterial port system and conventional angiographic technique | European Journal of Radiology. 2004; 51(1):6-11) | Population |
| 13 | Kieran et al, 2015 | Hepatitis C in the era of direct-acting antivirals: real-world costs of untreated chronic hepatitis C; a cross-sectional study | BMC Infectious Diseases. 2015; 15:471 | Population |
| 14 | Uka et al, 2007 | Similar effects of recombinant interferon-alpha-2b and natural interferon-alpha when combined with intra-arterial 5-fluorouracil for the treatment of advanced hepatocellular carcinoma | Liver International. 2007; 27(9):1209-16 | Outcomes |
| 15 | Bernal et al, 2007 | Intra-arterial rhenium-188 lipiodol in the treatment of inoperable hepatocellular carcinoma: results of an IAEA-sponsored multination study | International Journal of Radiation Oncology, Biology, Physics. 2007; 69(5):1448-55 | Outcomes |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| 16 | Tsoulfas et al, 2012 | Clinical and financial outcomes from the coordinated treatment of primary and metastatic hepatobiliary malignancies: Early experience of a new center | HPB. 2012; 14:552 | Outcomes |
| 17 | Livraghi et al, 1995 | Hepatocellular carcinoma and cirrhosis in 746 patients: long-term results of percutaneous ethanol injection | Radiology. 1995; 197(1):101-8 | Population |

**Table 24: Excluded studies – HRQoL SLR N=61)**

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| 1 | Zugazagoitia et al, 2013 | Sorafenib for non-selected patient population with advanced hepatocellular carcinoma: efficacy and safety data according to liver function | Clin Transl Oncol 2013; 15(2):146-53 | Outcomes |
| 2 | Dimitroulopoulos et al, 2013 | Demographic profile and outcome of 126 consecutive HCC cirrhotic patients treated with nexavar. A 5 year greek multicenter study | Journal of Clinical and Experimental Hepatology. 2013; 3 (1 SUPPL. 1):S105-S106) | Study design |
| 3 | Zhang et al, 2015 | Sorafenib continuation or discontinuation in patients with unresectable hepatocellular carcinoma after a complete response | Oncotarget. 2015; 6(27):24550-9 | Population |
| 4 | Montella et al, 2013 | Sorafenib in elderly patients with advanced hepatocellular carcinoma: a case series. | Oncology. 2013; 84(5):265-72, | Study design |
| 5 | Tzoracoleftherakis et al, 1999 | Intra-arterial versus systemic chemotherapy for non-operable hepatocellular carcinoma | Hepato-gastroenterology. 1999; 46(26):1122-5 | Outcomes |
| 6 | Xing et al, 2015 | Preservation of quality of life with doxorubicin drug-eluting bead transarterial chemoembolization for unresectable hepatocellular carcinoma: Longitudinal prospective study | Journal of Gastroenterology & Hepatology; 2015 30(7):1167-74 | Population |
| 7 | Aliberti et al, 2017 | Chemoembolization Adopting Polyethylene Glycol Drug-Eluting Embolics Loaded With Doxorubicin for the Treatment of Hepatocellular | American Journal of Roentgenology. 2017; 209(2):430-434 | Outcomes |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| | | Carcinoma | | |
| 8 | Nowak et al, 2008 | Assessment of health-related quality of life and patient benefit as outcome measures for clinical trials in hepatocellular carcinoma | Asia-Pacific Journal of Clinical Oncology. 2008; 4 (1):55-67 | Population |
| 9 | Cebon 2006 | Somatostatin receptor expression, tumour response, and quality of life in patients with advanced hepatocellular carcinoma treated with long-acting octreotide | British Journal of Cancer.2006; 95 (7):853-861 | Population |
| 10 | Dimitroulopoulos et al, 2002 | The role of sandostatin LAR in treating patients with advanced hepatocellular cancer | Hepato-Gastroenterology. 2002; 49(47):1245-50 | Outcomes |
| 11 | Farooqi et al, 2000 | Efficacy of octreotide in cases of inoperable hepatocellular carcinoma: A clinical trial | Journal of the College of Physicians and Surgeons—Pakistan. 2000; 10(7):258-60 | Outcomes |
| 12 | Yuen et al, 2002 | A randomized placebo-controlled study of long-acting octreotide for the treatment of advanced hepatocellular carcinoma | Hepatology. 2002; 36(3):687-91 | Outcomes |
| 13 | Gill et al, 2005 | Treatment outcomes with long acting octreotide in inoperable hepatocellular carcinoma: a local experience and review of literature | The Journal of the Pakistan Medical Association. 2005; 55(4):135-8 | Outcomes |
| 14 | Treiber et al, 2007 | Octreotide alone or in combination with rofecoxib as palliative treatment for advanced hepatocellular cancer | Zeitschrift fur Gastroenterologie. 2007; 45(5):369-77 | Outcomes |
| 15 | Verset et al, 2007 | Efficacy of the combination of long-acting release octreotide and tamoxifen in patients with advanced hepatocellular carcinoma: a randomised multicentre phase III study | Br J Cancer. 2007; 97(5):582-8 | Outcomes |
| 16 | Manesis et al, 1995 | Treatment of hepatocellular carcinoma with combined suppression and inhibition of sex hormones: a randomized, controlled trial | Hepatology. 1995; 21(6):1535-42 | Outcomes |
| 17 | Doffoel et al, 2008 | Multicentre randomised phase III trial comparing Tamoxifen alone or | Eur J Cancer. 2008; 44(4):528- | Population |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| | | with Transarterial Lipiodol Chemoembolisation for unresectable hepatocellular carcinoma in cirrhotic patients (Federation Francophone de Cancerologie Digestive 9402) | 38 | |
| 18 | Li et al, 2013 | Transarterial chemoembolization combined with celecoxib and lanreotide in the treatment of unresectable hepatocellular carcinoma: A prospective randomized controlled trial | Gastroenterology. 2013; 144 (5 SUPPL 1):S967 | Population |
| 19 | Yen et al, 2009 | Phase I/II study of PHY906/capecitabine in advanced hepatocellular carcinoma | Anticancer Research. 2009; 29 (10):4083-4092 | Population |
| 20 | Dollinger et al, 2010 | Thymostimulin versus placebo for palliative treatment of locally advanced or metastasised hepatocellular carcinoma: a phase III clinical trial | BMC Cancer. 2010; 10:457 | Population |
| 21 | Samonakis et al, 2002 | Treatment of hepatocellular carcinoma with long acting somatostatin analogues | Oncology reports. 2002; 9 (4):903-907 | Outcomes |
| 22 | Uchino et al, 1993 | Chemohormonal therapy of unresectable hepatocellular carcinoma | Am J Clin Oncol. 1993; 16(3):206-9 | Outcomes |
| 23 | Bonnetain et al, 2008 | Quality of life as a prognostic factor of overall survival in patients with advanced hepatocellular carcinoma: results from two French clinical trials | Quality of Life Research 2008; 17(6):831-43 | Outcomes |
| 24 | Farooqi et al, 2001 | Efficacy of intrahepatic absolute alcohol in unresectable hepatocellular carcinoma | Journal of the College of Physicians and Surgeons--Pakistan. 2001; 11(6):383-6 | Outcomes |
| 25 | Bronowicki et al, 1994 | Transcatheter oily chemoembolization for hepatocellular carcinoma. A 4-year study of 127 French patients | Cancer. 1994; 74(1):16-24 | Outcomes |
| 26 | Lin et al, 2017 | Cryoablation combined with allogenic natural killer cell immunotherapy improves the curative effect in patients with | Oncotarget. 2017; 8 (47):81967-81977 | Outcomes |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| | | advanced hepatocellular cancer | | |
| 27 | Liu et al, 2014 | Combination of percutaneous radiofrequency ablation with transarterial chemoembolization for hepatocellular carcinoma: Observation of clinical effects | Chinese Journal of Cancer Research. 2014; 26 (4):471-477 | Outcomes |
| 28 | Eltawil et al, 2012 | Quality of life and survival analysis of patients undergoing transarterial chemoembolization for primary hepatic malignancies: a prospective cohort study | HPB. 2012; 14(5):341-50 | Population |
| 29 | Bai et al, 2013 | The promise of clinical interventions for hepatocellular carcinoma from the west to mainland China | Palliative & supportive care. 2013; 11 (6):503-522 | Population |
| 30 | Poon et al, 2004 | Long-term oral branched chain amino acids in patients undergoing chemoembolization for hepatocellular carcinoma: a randomized trial | Alimentary pharmacology & therapeutics. 2004; 19(7):779-88 | Population |
| 31 | Chay et al, 2017 | Coriolus versicolor (Yunzhi) Use as Therapy in Advanced Hepatocellular Carcinoma Patients with Poor Liver Function or Who Are Unfit for Standard Therapy | Journal of Alternative & Complementary Medicine. 2017; 23(8):648-652 | Population |
| 32 | Cowawintaweewat et al, 2006 | Prognostic improvement of patients with advanced liver cancer after active hexose correlated compound (AHCC) treatment | Asian Pacific Journal of Allergy and Immunology. 2006; 24 (1) (pp 33-45) | Outcomes |
| 33 | Yau et al, 2015 | Preliminary efficacy, safety, pharmacokinetics, pharmacodynamics and quality of life study of pegylated recombinant human arginase 1 in patients with advanced hepatocellular carcinoma | Invest New Drugs. 2015; 33(2):496-504 | Outcomes |
| 34 | Salem et al, 2013 | Increased quality of life among hepatocellular carcinoma patients treated with radioembolization, compared with chemoembolization | Clinical Gastroenterology & Hepatology. 2013; 11(10):1358-1365 | Population |
| 35 | Kolligs et al, 2015 | Pilot randomized trial of selective internal radiation therapy vs. chemoembolization in unresectable hepatocellular carcinoma | Liver International. 2015; 35(6):1715-21 | Study design |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| 36 | Chie et al, 2015 | Quality of life changes in patients undergoing treatment for hepatocellular carcinoma | Quality of Life Research. 2015; 24(10):2499-506 | Population |
| 37 | Mutsaers et al, 2016 | Systematic review of patient-reported quality of life following stereotactic ablative body radiation therapy for primary and metastatic liver cancer | International Journal of Radiation Oncology. 2016; 96 (2 Suppl 1):E536-E537) | Population |
| 38 | Rosler et al, 1994 | Superselective radioembolization of hepatocellular carcinoma: 5-year results of a prospective study | NuklearMedizin. 1994; 33 (5): 206-214 | Outcomes |
| 39 | Shun et al, 2008 | Changes in quality of life and its related factors in liver cancer patients receiving stereotactic radiation therapy | Supportive Care in Cancer. 2008; 16(9):1059-65 | Outcomes |
| 40 | Li et al, 2007 | Short and long term efficacy of high intensity focused ultrasound therapy for advanced hepatocellular carcinoma | Journal of Gastroenterology & Hepatology. 2007; 22(12):2148-54 | Outcomes |
| 41 | Palmieri et al, 2012 | Psycological profile and health related quality of life (HRQOL) in patients with hepatocellular carcinoma (HCC) | Journal of Hepatology. 2012; 56:S291 | Population |
| 42 | Butt et al, 2014 | Psychometric properties of a brief, clinically relevant measure of pain in patients with hepatocellular carcinoma | Quality of life research. 2014; 23 (9):2447-2455 | Population |
| 43 | Chie et al, 2017 | Differences in health-related quality of life between European and Asian patients with hepatocellular carcinoma | Asia-Pacific Journal of Clinical Oncology. 2017; 13(5): e304-e311 | Population |
| 44 | Fan et al, 2013 | Health-related quality of life in patients with hepatocellular carcinoma: The mediation effects of illness perceptions and coping | Psycho-Oncology. 2013; 22 (6): 1353-1360) | Population |
| 45 | Fan et al, 2012 | Illness experience in patients with hepatocellular carcinoma: an interpretative phenomenological analysis study | European Journal of Gastroenterology & Hepatology. 2012; 24(2):203-8 | Population |
| 46 | Fielding et al, | Quality of life as a predictor of cancer survival among Chinese | Eur J Cancer. 2007; 43 (11): | Population |

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| | 2007 | liver and lung cancer patient | 1723-1730 | |
| 47 | Gomes et al, 2014 | Health-related quality of life of patients with hepatocellular carcinoma submitted to multiple therapy modalities | European Journal of Surgical Oncology. 2014; 40 (11): S131 | Population |
| 48 | Hansen et al, 2015 | Patients with hepatocellular carcinoma near the end of life: A longitudinal qualitative study of their illness experiences | Cancer Nursing. 2015; 38 (4): E19-E27 | Outcomes |
| 49 | Kaiser et al, 2014 | Important and relevant symptoms including pain concerns in hepatocellular carcinoma (HCC): a patient interview study. | Supportive Care in Cancer. 2014; 22(4):919-26 | Outcomes |
| 50 | Li et al, 2015 | Fear of progression and quality of life in patients with hepatocellular carcinoma. | Psycho-Oncology. 2015; 24:267-268 | Population |
| 51 | Li et al, 2017 | Prognostic values of EORTC QLQ-C30 and QLQ-HCC18 index-scores in patients with hepatocellular carcinoma - clinical application of health-related quality-of-life data | BMC Cancer. 2017; 17 (1) | Population |
| 52 | Meier et al, 2015 | Role functioning is associated with survival in patients with hepatocellular carcinoma | Quality of Life Research. 2015; 24(7):1669-75 | Population |
| 53 | Mikoshiba et al, 2013 | Depressive symptoms after treatment in hepatocellular carcinoma survivors: Prevalence, determinants, and impact on health-related quality of life | Psycho-Oncology.2013; 22 (10): 2347-2353 | Population |
| 54 | Paris et al, 2011 | Assessing symptom burden of patients with advanced hepatoceullular carcinoma | Journal of the American Geriatrics Society. 2011; 59: S10 | Outcomes |
| 55 | Qiao et al, 2012 | Health-related quality of life evaluated by tumor node metastasis staging system in patients with Hepato cellular carcinoma | World Journal of Gastroenterology. 2012;18 (21): 2689-2694 | Study design |
| 56 | Ryu et al, 2010 | Symptom clusters and quality of life in Korean patients with hepatocellular carcinoma | Cancer Nursing. 2010; 33 (1):3-10 | Population |
| 57 | Steel et al, 2007 | Health-related quality of life: Hepatocellular carcinoma, chronic liver disease, and the general | Quality of Life Research. 2007; 16 (2): 203-215 | Population |

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

| | Author, year | Title | Citation | Reason for exclusion |
|---|---|---|---|---|
| | | population | | |
| 58 | Steel et al, 2014 | Health-related quality of life as a prognostic factor in patients with advanced cancer | Cancer. 2014; 120 (23): 3717-3721 | Population |
| 59 | Sun et al, 2008 | Symptom concerns and quality of life in hepatobiliary cancers | Oncology nursing forum. 2008; 35 (3): E45-52 | Population |
| 60 | Wang et al, 2012 | Symptom clusters in Chinese patients with primary liver cancer | Oncology Nursing Forum. 2012; 39(6):E468-79 | Outcomes |
| 61 | Zhang et al, 2012 | A systematic review of health state utilities in patients with advanced hepatocellular carcinoma | Value in Health. 2012; 15 (4): A225) | Study design |

B23. The ERG has identified discrepancies between Table 18 of Appendix J and the economic model, specifically:

- post-progression drug costs for lenvatinib (£[redacted] vs £[redacted]);
- total costs for lenvatinib (£[redacted] vs £[redacted]);
- primary drug costs for sorafenib (£10,582 vs £21,163);
- post-progression drug costs for sorafenib (£1,269 vs 1,864);
- total costs for sorafenib (£53,440 vs £64,617).

Please clarify if the values in the economic model are correct.

We can confirm that incorrect values were transcribed across to the submission document and that the values presented in the economic model are correct, i.e.:

- post-progression drug costs for lenvatinib: £[redacted];
- total costs for lenvatinib: £[redacted]
- primary drug costs for sorafenib: £21,163
- post-progression drug costs for sorafenib: £1,864);
- total costs for sorafenib: £64,617

**Section C: Textual clarifications and additional points**

C1. Please provide the listings and appendices for the CSR.

The full listings and appendices have not been provided, however specific data can be made available on request.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

**References**

1. Food and Drug Administration. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics. 2007. Available from: https://www.fda.gov/downloads/Drugs/Guidances/ucm071590.pdf. Accessed on: October 2017.
2. European Medicines Agency. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man. EMA/CHMP/27994/2008/Rev1 2012.
3. Llovet JM, Pena CE, Lathia CD, Shan M, Meinhardt G, Bruix J, et al. Plasma biomarkers as predictors of outcome in patients with advanced hepatocellular carcinoma. Clin Cancer Res. 2012;18(8):2290-300.
4. Bruix J, Raoul JL, Sherman M, Mazzaferro V, Bolondi L, Craxi A, et al. Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma: subanalyses of a phase III trial. J Hepatol. 2012;57(4):821-9.
5. European Association For The Study Of The Liver, European Organisation For Research And Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. Eur J Cancer. 2012;48(5):599-641.
6. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30(1):52-60.
7. National Institute for Health and Care Excellence. Cabozantinib for previously treated advanced renal cell carcinoma 2017. Available from: https://www.nice.org.uk/guidance/ta463/evidence. Accessed on: March 2018
8. National Institute for Health and Care Excellence. Sorafenib for treating advanced hepatocellular carcinoma. 2017. Available from: https://www.nice.org.uk/guidance/ta474. Accessed on: October 2017.

NICE National Institute for Health and Care Excellence

## Patient organisation submission

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

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 | **Charles Gore** |

| | |
|---|---|
| 2. Name of organisation | The Hepatitis C Trust |
| 3. Job title or position | CEO |
| 4a. Brief description of the organisation (including who funds it). How many members does it have? | The national patient charity for people living with or affected by hepatitis C funded by grant-making trusts, individual donations, some government grants and grants from industry. We have never received any money from makers of cancer drugs. We have over 3,000 members of our patient association. |
| 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? | Through our national helpline which takes about 150 calls a week, our support groups and our work on the ground through our peer community and prison projects and our outreach service. Although HCC is comparatively rare in people with hepatitis C, as more and more people get cured of hepatitis C, people with end stage liver disease, and particularly HCC, are becoming proportionately bigger users of our services. |
| **Living with the condition** | |
| 6. What is it like to live with the condition? What do carers | HCC is a cancer with poor prognosis. It is often diagnosed late – and for some people may be the first sign they have hepatitis C. Once diagnosed liver transplant may an only option but the lack of livers |

| | |
|---|---|
| experience when caring for someone with the condition? | means the wait for a transplant may give the cancer time to grow or proliferate past the point at which a transplant will be considered. I personally supported someone attending our support groups to whom that happened. People also often think that curing their hepatitis C will cure their cancer or that it will remove entirely the risk of cancer, only to find neither is true. It can therefore be agonising mentally both for patients and for carers. Physically it can cause digestive problems, weight loss and pain and may be associated with symptoms from the concomitant decompensated cirrhosis like ascites and aesophogeal varices and bleeding. Generally people feel increasingly unwell. |
| **Current treatment of the condition in the NHS** | |
| 7. What do patients or carers think of current treatments and care available on the NHS? | The current treatment for unresectable HCC is sorafenib which now appears to be better understood and is providing longer survival (in the phase 3 trial comparing lenvatinib vs sorafenib the latter's overall survival was significantly longer than originally achieved in the early days of its use). However it still offers short survival and very short time to progression. |
| 8. Is there an unmet need for patients with this condition? | Unresectable HCC is essentially a fatal condition. There is a significant need to increase survivability as well as time to progression and progression-free survival. |
| **Advantages of the technology** | |
| 9. What do patients or carers think are the advantages of the technology? | It improves survival overall vs no intervention and time to progression and progression-free survival vs the current standard of care sorafenib. As result there appears to be later onset of cancer symptom (role function, pain, diarrhoea, nutrition and body image) deterioration than with sorafenib. Lenvatinib also has a diiferent side effect profile than sorafenib and that choice may be beneficial for patients. There may also be advantages of having a competitor in the market to allow NHSE to obtain better prices |

| | |
|---|---|
| **Disadvantages of the technology** | |
| 10. What do patients or carers think are the disadvantages of the technology? | Side effects |
| **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? | Liver cancer disproportionately affects men (though not women) living in deprived areas in England. It also disproportionately affects Asian and Black people. |

| **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:

- Lenvatinib is non-inferior to sorafenib but offers a different side effect profile
- It offers increased time to progression
- It offers increased progression-free survival
- Competition in the market may lower prices to the NHS (cf hepatitis C drugs)
-

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

## Professional organisation submission

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

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 | **Dr Syed Hyder Hussaini** |
| 2. Name of organisation | **BSG Liver Section** |
| 3. Job title or position | **Consultant Hepatologist** |

| | |
|---|---|
| 4. Are you (please tick all that apply): | ✓ an employee or representative of a healthcare professional organisation that represents clinicians?<br>✓ a specialist in the treatment of people with this condition?<br>✓ a specialist in the clinical evidence base for this condition or technology?<br><br>☐ other (please specify): |
| 5a. Brief description of the organisation (including who funds it). | The British Society of Gastroenterology is an organisation focused on the promotion of gastroenterology within the United Kingdom. It has over three thousand members drawn from the ranks of physicians, surgeons, pathologists, radiologists, scientists, nurses, dietitians, and others interested in the field. Founded in 1937 it has grown from a club to be a major force in British medicine, with representation within the British Royal Colleges and consequently the Department of Health and Government. Internationally it is represented at World and European level. The BSG is a registered charity. It is funded by subscription from members. |
| 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, or prevent progression or disability.) | To improve overall survival from hepatocellular cancer (HCC) that is not amenable to curative therapy |

| | |
|---|---|
| 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.) | Compared with standard of care, a significant treatment response for lenvatinib therapy would be prolongation of: (i) Overall survival (OS); (ii) time to progression of symptoms (TTP), (iii) Progression-free survival (PFS) and (iv) Objective response rate (ORR) – the proportion of patients with reduction in tumour burden of a predefined amount.<br><br>Although OS is the accepted metric for beneficial treatment response for solid tumours, in the assessment of palliative treatment for HCC, this may be unreliable as many patients may die from complications of cirrhosis as opposed to tumour related death and thus confound results.<br><br>TTP & PFS are often used as an endpoint in diseases with very short survival times post therapy such as HCC [3]. PFS differs from TTP in that it is more objective and less prone to bias. PFS is useful when comparing different and or multiple treatment regimens for HCC. The metric of ORR has the advantage that the observed effect is attributable directly to the drug, not the natural history of the disease |
| 8. In your view, is there an unmet need for patients and healthcare professionals in this condition? | There is an unmet need is to improve efficacy of first line systemic therapy for palliative therapy for HCC as measured by improved overall survival and secondary end points such as PFS since this last metric of is a good endpoint for comparison of different treatment modalities in advanced HCC. |
| | |

| What is the expected place of the technology in current practice? | |
|---|---|
| 9. How is the condition currently treated in the NHS? | HCC is a rapidly progressive tumour usually occurring in patients with cirrhosis who often may have decompensated disease. In UK clinical practice only about 30% of patients with hepatocellular carcinoma are suitable for curative therapy such as liver transplantation, local resection or radiofrequency ablation or palliative chemoembolization. Following these treatments about half may need further locoregional therapy or systemic treatment due to treatment failure or recurrence of tumour.<br>Sorafenib, a multi kinase inhibitor, has been the standard therapy for patients with unresectable hepatocellular carcinoma since 2007, when it was shown to prolong overall survival, as verified in the SHARP trial [1] by 2.8 months compared to placebo with a corresponding an extension in time to disease progression of 11.7 weeks. These data were confirmed by a study conducted in the Asia-Pacific region [2]. However, Sorafanib is associated with a number of severe adverse events (such as diarrhoea and hand-foot skin reaction) which are unpredictable and can affect quality of life and treatment leads to tumour regression in only 2-3% [1-2]. Nonetheless, sorafenib is the standard systemic therapy for HCC recommended by international guidelines. Sorafenib is indicated for patients with well-preserved liver function (Child-Pugh score 5-7) and with advanced tumours (BCLC B/C) or those tumours progressing after loco-regional therapies. Other multi-tyrosine kinase inhibitors have been tested in clinical trials either alone or in combination with other novel loco-regional therapies for HCC. However, none have been found to be an effective first line systemic therapy for HCC. |
| ● Are any clinical guidelines used in the treatment of the condition, and if so, which? | AASLD Guidelines for the Treatment of Hepatocellular Carcinoma (2018)<br>https://www.aasld.org/sites/default/files/guideline.../HCC%20Guideline%202018.pdf<br><br>Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update.<br>Hepatol Int. 2017 Jul;11(4):317-370. doi: 10.1007/s12072-017-9799-9. Epub 2017 Jun 15. |

| | |
|---|---|
| | EASL–EORTC Clinical Practice Guidelines: Management of hepatocellular carcinoma (2012)<br>http://www.easl.eu/research/our-contributions/clinical-practice-guidelines/detail/management-of-hepatocellular-carcinoma-easl-eortc-clinical-practice-guidelines |
| ● 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.) | All cases are discussed at joint Hepatobiliary cancer multidisciplinary meetings. The hub is the regional HBP Centre, the spoke a secondary care centre which will often deliver systemic therapy for HCC treatment. Patients are staged by cross-sectional imaging (usually CT/MRI) and clinically assessed for stage of cirrhosis (Child-Pugh score) and performance status (ECOG). These data are incorporated into an overall HCC Barcelona Clinic Liver Cancer staging score (BCLC). Essentially BCLC Stage B/C patients and those that fail curative therapy or chemoembolization are considered for systemic therapy. Care pathways are well defined with no major difference of opinion between professionals regarding management. |
| ● What impact would the technology have on the current pathway of care? | Improve first line systemic therapy |
| 10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? | Yes |

| | |
|---|---|
| ● How does healthcare resource use differ between the technology and current care? | None |
| ● In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.) | Specialist hepatology and oncology clinics |
| ● What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.) | None |
| 11. Do you expect the technology to provide clinically meaningful benefits compared with current care? | Yes |
| ● Do you expect the technology to increase length of life more than current care? | Yes.<br>The REFLECT phase III [4] study compared lenvatinib to sorafenib in over 900 patients with unresectable HCC with a good performance status. The median overall survival with lenvatinib compared with sorafenib was 13.6 months and 12.3 months, respectively (HR, 0.92; 95%; CI, 0.79-1.06). The hazard ratio was estimated with a stratified Cox proportional hazard model.<br>The median progression free survival was 7.4 months with lenvatinib compared with 3.7 months for |

| | |
|---|---|
| | sorafenib (HR, 0.66; 95%; CI, 0.57-0.77; *P* < .00001). The median time to progression was 8.9 months with lenvatinib compared with 3.7 months for sorafenib (HR, 0.63; 95%; CI, 0.53-0.73; *P* < .00001). The objective response rate was 24.1% with lenvatinib versus 9.2% with sorafenib (odds ratio, 3.13%; 95% CI, 2.15-4.56; *P* <.00001). The median treatment duration for lenvatinib was 5.7 months versus 3.7 months for sorafenib.<br>Thus, the REFLECT trial showed that lenvatinib was statistically non-inferior to sorafenib in overall survival for patients with unresectable HCC and superior to sorafenib as measured by progression free survival, time to progression and objective response rate. |
| ● Do you expect the technology to increase health-related quality of life more than current care? | Yes |
| 12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population? | Most appropriate for Child Pugh A to B7, ECOG Performance status 0-1, BCLC Stage B & C<br>Not appropriate for curable stage HCC or advanced disease with decompensation |
| **The use of the technology** | |
| 13. Will the technology be easier or more difficult to use for patients or healthcare | No change than standard of care |

| | |
|---|---|
| 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.) | |
| 14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing? | Stop rule would be symptomatic deterioration on treatment, hepatic decompensation or side effects from therapy. No additional testing required |
| 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 | Yes – survival prolongation |

| | |
|---|---|
| (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? | Improved survival with no current alternative treatments in a condition with an overall short survival time. |
| ● Is the technology a 'step-change' in the management of the condition? | No |
| ● Does the use of the technology address any particular unmet need of the patient population? | Yes – as above |
| 17. How do any side effects or adverse effects of the technology affect the management of the condition | Unknown |

| | |
|---|---|
| and the patient's quality of life? | |
| **Sources of evidence** | |
| 18. Do the clinical trials on the technology reflect current UK clinical practice? | Yes |
| ● 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? | Overall survival at least as good as current standard of care with high probability that outcome may be better as there is an improved progression free survival, time to progression and objective response rate compared to standard of care. |
| ● If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? | Not used |
| ● Are there any adverse effects that were not apparent in clinical trials but have come to light | Common adverse events of lenvatinib were hypertension, diarrhoea, fatigue ,decreased appetite & weight. Drug discontinuations due to adverse events in the REFLECT study [4] were 13% in lenvatinib arm and 9% in the sorafenib arm. In this study dose reductions due to treatment-emergent adverse events (TEAEs) were required for 37% of those in the lenvatinib arm and for 38% of those in the sorafenib group. Grade ≥3 TEAEs were more common with lenvatinib versus sorafenib (57% vs 49%, respectively). |

| | |
|---|---|
| subsequently? | |
| 19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence? | No |
| 20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA474]? | No direct comparator for first line therapy of advanced HCC.<br><br>However, the systemic treatment of patients with HCC is FDA approved with the multikinase inhibitor regorafenib in the second-line setting following sorafenib. Treatment with cabozantinib (tyrosine kinases receptor inhibitor) was recently shown to improve overall survival compared with placebo for patients with previously treated advanced HCC in the phase III CELESTIAL trial and thus is a promising 2nd line therapy for advanced HCC . Finally, the PD-1 inhibitor, nivolumab, has shown promise in a phase II study. |
| 21. How do data on real-world experience compare with the trial data? | No experience |

| | |
|---|---|
| **Equality** | |
| 22a. Are there any potential equality issues that should be taken into account when considering this treatment? | No |
| 22b. Consider whether these issues are different from issues with current care and why. | NA |
| **Key messages** | |

23. In up to 5 bullet points, please summarise the key messages of your submission.

- First line treatment option alternative to sorafenib
- Overall survival at least as good as standard of care for systemic treatment with Sorafanib for HCC
- Clinical improvements in progression free survival, time to progression and objective response rate compared to standard of care
- Increase in frequency of side effects compared to standard of care (sorafenib), but similar treatment discontinuation rate

1. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, de Oliveira AC, et al: Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 2008;359:378-390.

2. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, Luo R, et al: Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol 2009;10:25-34

3. Llovet JM[1], Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, Sherman M, Schwartz M, Lotze M, Talwalkar J, Gores GJ; Panel of Experts in HCC-Design Clinical Trials. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008 May 21;100(10):698-711. doi: 10.1093/jnci/djn134. Epub 2008 May 13.

4. Cheng A-L, Finn RS, Qin S, et al. Phase III trial of lenvatinib (LEN) vs sorafenib (SOR) in first-line treatment of patients (pts) with unresectable hepatocellular carcinoma (uHCC). *J Clin Oncol*. 2017;35 (suppl; abstr 4001)

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

**NHS England submission to NICE for the appraisal of lenvatinib in the treatment of locally advanced/metastatic hepatocellular carcinoma**

1. NHS England notes that the lenvatinib vs sorafenib REFLECT trial recruited patients who were previously untreated with systemic therapy with either histologically or clinically confirmed hepatocellular carcinoma (HCC) and all patients were of Child-Pugh A status and also of ECOG performance score of 0 or 1. If NICE recommends lenvatinib in this indication, NHS England treatment criteria will reflect these (sensible) eligibility criteria. These are set out in the final paragraph of this submission.
2. NHS England notes that dosing in the clinical trial reflected an arbitrary split depending on a patient body weight of 60Kg. Patients of ≥ 60Kg were treated with 12mg daily doses and those with weights < 60Kg received 8mg daily. NHS England notes that all patients in the sorafenib arm had the same starting dose of sorafenib as specified in the sorafenib SPC. It is common for clinicians in England to start sorafenib at a lower dose in some patients and then increase it if sorafenib is sufficiently well tolerated. NHS England presumes that the lenvatinib SPC for HCC will also state this dosing difference according to a weight of < 60Kg or ≥ 60Kg. Trusts will have to be careful in implementation of the starting doses of lenvatinib as these are 24mg daily in differentiated thyroid cancer, 16mg daily in renal cancer and now (presumably) 12 or 8mg daily in HCC.
3. NHS England notes the design of the lenvatinib vs sorafenib RCT in that patients randomised to lenvatinib could switch to sorafenib and not be entered into clinical trials whereas sorafenib patients could continue post progression on sorafenib and receive regorafenib (if available) or enter clinical trials. Use of sorafenib after disease progression on sorafenib is not commissioned by NHS England for rather obvious reasons. As far as NHS England is aware, the only licensed systemic therapy proven to result in a survival benefit after sorafenib is regorafenib and very few patients received regorafenib in either arm of the REFLECT trial (1% in the lenvatinib arm, 5% in the sorafenib arm). Approximately 25-30% of patients in both arms received further systemic therapy with a TKI: this was mainly sorafenib in the lenvatinib arm and was clearly a mixture of TKIs in the sorafenib arm.
4. Despite a significant and clinically meaningful difference in PFS (when considering the disease of HCC) between lenvatinib (7.4 months) and sorafenib (3.7 months), there was no difference in overall survival and NHS England finds this observation difficult to explain when rates of post-progression TKIs were similar between treatment arms and so few received regorafenib. What is clear from consideration of the median survivals durations in the RCT is that there is no clinical evidence to justify a survival gain from lenvatinib over sorafenib in the economic model of almost 3 months.

5. Lenvatinib has a higher response rate of 24% versus 9% for sorafenib. For patients who are symptomatic of bulky disease (eg liver pain or the consequences of compression of large blood vessels) lenvatinib offers a greater chance of inducing a physical reduction in the size of the disease causing symptoms: the QOL tool did show an advantage for lenvatinib in this regard when considering pain.
6. The difference in median PFS for lenvatinib (7.4 mo) vs sorafenib (3.7 mo) is proportionately much greater than the mean treatment durations for lenvatinib (8 mo) vs sorafenib (6 mo). The explanations could be the use of sorafenib post-progression (not commissioned by NHS England) or that some patients did particularly well with sorafenib. There may be other explanations of which NHS England is unaware.
7. There are both similar and different main toxicity profiles between the lenvatinib (hypertension and diarrhoea) and sorafenib (diarrhoea and palmar-plantar skin inflammation). NHS England notes the modestly higher discontinuation rate due to treatment related adverse events in the lenvatinib arm (20% vs 15%). In terms of avoiding a specific toxicity such as palmar plantar syndrome (due to sorafenib) or hypertension (due to lenvatinib), the availability of both drugs would offer patients and clinicians the advantage of choice to avoid specific side-effects.
8. If recommended by NICE as a treatment option for the treatment of previously untreated patients with HCC, NHS England would commission the use of **either** lenvatinib **or** sorafenib as systemic TKI treatment options. It would allow switching from one drug to the other **only** if there were unacceptable side-effects **and** documented evidence that disease progression had not occurred at the time of switching.
9. NHS England notes the patient access scheme (PAS) arrangements for lenvatinib and that these have been incorporated into the economic modelling. There is a commercial access agreement for sorafenib between Bayer and NHS England which will apply to the costing of sorafenib and this information will need to be incorporated into the costing of sorafenib.
10. If the option of lenvatinib is recommended by NICE, NHS England treatment criteria for the use of lenvatinib in HCC are currently planned to be as set out below. Clinicians will have to tick each and every criteria to gain access to lenvatinib.
    1. I confirm that this application is made by and the first cycle of systemic anti-cancer therapy with lenvatinib will be prescribed by a consultant specialist specifically trained and accredited in the use of systemic anti-cancer therapy
    2. I confirm that **one** of the following applies to the patient: the patient has a confirmed histological diagnosis of hepatocellular carcinoma **or** a biopsy is deemed to be very high risk or technically not feasible in the patient and all of the criteria immediately below are met
        - The decision not to biopsy has been made and documented by a specialist HCC multi-disciplinary meeting

- The tumour meets the non-invasive diagnostic criteria of HCC (Journal of Hepatology 2012; 56: 908-943)

3. I confirm that the patient has metastatic disease or locally advanced disease that is ineligible for or failed surgical or locoregional therapies
4. I confirm that the patient has not previously received any systemic therapy for HCC unless the patient had to discontinue sorafenib within 2 months of starting sorafenib because of toxicity (ie there was sorafenib toxicity which could not be managed by dose delay or dose modification ) **and** there has been no disease progression whilst on sorafenib. [NB NHS England will take advice from clinical experts as to whether the time duration should be 2 or 3 months for this rule]
5. I confirm that the patients has Child–Pugh liver function class A
6. I confirm that the patient has an ECOG performance status of either 0 or 1
7. I confirm that lenvatinib is to be used as a single agent
8. I confirm that lenvatinib is to be continued until disease progression or unacceptable toxicity or patient choice to stop treatment
9. I confirm that a formal medical review as to whether treatment with lenvatinib should continue or not and at what dose will be scheduled to occur at least by the start of the second cycle of treatment
10. I confirm that no treatment breaks of more than 6 weeks beyond the expected cycle length are allowed (to allow any toxicity of current therapy to settle or intercurrent comorbidities to improve)
11. I confirm that lenvatinib is to be otherwise used as set out in its Summary of Product Characteristics

[redacted]

NHS England Chair of Chemotherapy Clinical Reference Group and National Clinical Lead for the Cancer Drugs Fund

May 2018

NICE National Institute for Health and Care Excellence

## Clinical expert statement

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

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 Syed Hyder Husaini** |
| 2. Name of organisation | **BSG Liver Section** |

| | |
|---|---|
| 3. Job title or position | **Consultant Hepatologist** |
| 4. Are you (please tick all that apply): | ☒ an employee or representative of a healthcare professional organisation that represents clinicians?<br>☒ a specialist in the treatment of people with this condition?<br>☒ a specialist in the clinical evidence base for this condition or technology?<br>☐ 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<br>☐ no, I disagree with it<br>☐ I agree with some of it, but disagree with some of it<br>☐ 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 rest of this form will be deleted after submission.) | ☒ yes |

| **Topic-specific questions** | |
|---|---|
| 24. Would treatment with lenvatinib be considered in people with ECOG performance status >1 or Child-Pugh grade B liver impairment? | Yes |
| 25. Is best supportive care used in clinical practice given that TA474 recommends sorafenib as a treatment option for treating advanced hepatocellular carcinoma for people with Child-Pugh grade A liver impairment? | Yes |
| **Key messages** | |

26. In up to 5 bullet points, please summarise the key messages of your statement.

- First line treatment option alternative to sorafenib
- Overall survival at least as good as standard of care for systemic treatment with Sorafanib for HCC
- Clinical improvements in progression free survival, time to progression and objective response rate compared to standard of care
- Increase in frequency of side effects compared to standard of care (sorafenib), but similar treatment discontinuation rate

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.

NICE National Institute for Health and Care Excellence

## Clinical expert statement

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

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 | **Professor Tim Meyer** |
| 2. Name of organisation | **UCL Cancer Institute and Royal Free NHS Trust** |

| | |
|---|---|
| 3. Job title or position | **Professor of Experimental Cancer Medicine and Honorary Consultant in Medical Oncology** |
| 4. Are you (please tick all that apply): | ☒ an employee or representative of a healthcare professional organisation that represents clinicians?<br>☒ a specialist in the treatment of people with this condition?<br>☒ a specialist in the clinical evidence base for this condition or technology?<br>☐ 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<br>☐ no, I disagree with it<br>☐ I agree with some of it, but disagree with some of it<br>☒ 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 rest of this form will be deleted after submission.) | ☐ yes |

| 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 for patients with advanced HCC is to improve survival. Secondary aims include reducing tumour burden and delaying tumour 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 survival, I would regard a three month improvement compared with placebo to be clinically relevant in this patient group. Sorafenib, which is the current standard of care, has a response rate of around 2% (by RECIST 1.1) based on the registrational trial (Llovet J, et al NEJM 2008), and the main effect was delay in progression. A competing drug that had an improved response rate would be clinically important, particularly in those patients who are symptomatic due to tumour burden. |
| 9. In your view, is there an unmet need for patients and healthcare professionals in this condition? | Yes. The current standard of care has a low response rate and toxicity that results in up to 25% patients discontinuing therapy. |
| **What is the expected place of the technology in current practice?** | |

| | |
|---|---|
| 10. How is the condition currently treated in the NHS? | The only approved treatment for advanced HCC is sorafenib which is associated with an improvement in median survival of 2-3 months, a response rate of around 2% and a toxicity profile that results in up to 25% patients discontinuing therapy. The median time on treatment is around 4-5 months and treated patients have an expected median survival of around 11 months. |
| • Are any clinical guidelines used in the treatment of the condition, and if so, which? | In Europe, the reference guidelines are published by the European Association for the Study of the Liver (EASL) and the European Society for Medical Oncology (ESMO). These were last updated in 2012 and both have been revised. Publication is expected in mid-2018. |
| • 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.) | The standard of care for HCC is well defined and the Barcelona Clinic Liver Cancer (BCLC) staging and therapeutic algorithm is largely accepted across Europe. |
| • What impact would the technology have on the current pathway of care? | Currently, sorafenib is the only option approved for advanced hepatocellular Carcinoma. The approval of lenvatinib would provide an alternative first-first line therapy. |
| 11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? | Lenvatinib would be an alternative to sorefanib, the current standard of care. |

| | |
|---|---|
| • How does healthcare resource use differ between the technology and current care? | Both sorafenib and lenvatinib are oral agents that are administered in the outpatient setting. Both require outpatient monitoring with clinical review at between at 2 weekly intervals for the first 2 months and 4 weekly thereafter in order to manage toxicity and dose adjustment. Imaging frequency would be 2-3 monthly in both cases. In terms of clinical resource, it is unlikely that there would be significant differences. |
| • In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.) | Treatment should be overseen by an oncologist in secondary care. Because the majority of patients have chronic liver disease in addition to cancer, patients should ideally be treated within specialist HCC clinics jointly attended by an oncologist and a hepatologist. |
| • What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.) | No additional investment would be required to introduce lenvatinib into clinical care. It would simply replace the current standard of care. |
| 12. Do you expect the technology to provide clinically meaningful benefits compared with current care? | The REFLECT trial (Kudo M et al Lancet 2018) compared lenvatinib with sorafenib in an open-label randomised phase III trial. The trial was powered for non-inferiority and met its primary endpoint with a median overall survival of 13.6 and 12.3 months respectively. Secondary endpoints were assessed by blinded independent review and can therefore be considered robust. Both response rate and progression free survival were significantly better for lenvatinib (18.8% versus 6.5% and 7.3 versus 3.6%). There was also a non-significant trend towards delaying deterioration in quality of life in favour of lenvatinib. |
| • Do you expect the technology to increase length of life more than current care? | Although the survival was longer in the lenvatinib treated group compared with the sorafenib treated group, this was not significant, and the trial demonstrated non-inferiority for survival. Therefore, there is no evidence that patients will live longer on lenvatinib. |

| | |
|---|---|
| • Do you expect the technology to increase health-related quality of life more than current care? | During the course of treatment with lenvatinib and sorafenib, the quality of life of patients showed a deterioration that is consistent with disease progression. However the scores for role functioning, pain, diarrhoea, nutrition, body image deteriorated earlier in the sorafenib-treated group which may be the result of earlier disease progression as demonstrated by the differences in progression free survival. |
| 13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population? | Patients in the REFLECT trial had well preserved liver function (Child Pugh Class A) and were performance status <2. Therefore we have no evidence that patients with worse liver function or performance status benefit. The experience with sorafenib has demonstrated that patient with poor liver function and poor performance status do not appear to have meaningful benefit in terms of survival and the same may be true for lenvatinib but we do not have the evidence to support this to date (King J et al, Clin Oncol 2017 (4):256-262).<br>An extensive subgroup analysis was performed in the REFLECT trial and some subgroups had more favourable outcomes with lenvatinib compared with sorafenib. These included those with extrahepatic spread or macroscopic portal vein invasion, AFP ≥200ng/ml and hepatitis B infection. There were no patient subgroups that had better outcomes with sorafenib. |
| **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 | The administration of lenvatinib will be very similar in terms of delivery and monitoring. Hypertension was more common with lenvatinib than sorafenib and this may require increased administration of anti-hypertensives. Conversely, diarrhoea and planter palmer erythema were more common with sorafenib, requiring increased administration of anti-diarrheal medication and topical creams. |

| | |
|---|---|
| treatments needed, additional clinical requirements, factors 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? | Reasons to discontinue treatment will include 1. Patient choice 2. Poor tolerance or 3. Radiological or clinical progression. Most centres will routinely image patients with CT or MRI every 2-3 months and discontinue treatment in the event of progression. |
| 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? | Tumour response and delay in disease progression |
| 17. Do you consider the technology to be innovative in | The increased response rate, delay in disease progression and delay in deterioration of QOL are all important to this patient group. Additionally, there are patient groups that have been shown to derive little |

| | |
|---|---|
| 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? | benefit from sorafenib including those who do not have hepatitis C infection and those with extrahepatic spread (Bruix et al, J Hepatology 2017 Nov;67(5):999-1008). These patients seem to derive greater benefit from lenvatinib. |
| • Is the technology a 'step-change' in the management of the condition? | Sorafenib was approved 10 years ago and multiple trials have failed to demonstrate superiority or non-inferiority for new agents. The trial of lenvatinib is the first to achieve its primary endpoint and therefore represents a step forward in the treatment of advanced HCC. |
| • Does the use of the technology address any particular unmet need of the patient population? | Yes. The subgroup of patients mentioned above that derive little benefit from sorafenib. |
| 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? | The side effect profile is typical for a multikinase inhibitor but, compared with sorafenib, there is a higher incidence of hypertension, decreased appetite, proteinuria, vomiting and hypothyroidism. Conversely, rash, diarrhoea and alopecia is higher for sorafenib. The QOL analysis favours lenvatinib. |
| **Sources of evidence** | |

| | |
|---|---|
| 19. Do the clinical trials on the technology reflect current UK clinical practice? | Yes. Sorafenib was the standard of care control for the REFLECT trial and this is consistent with the standard of care in the UK. |
| • If not, how could the results be extrapolated to the UK setting? | NA |
| • What, in your view, are the most important outcomes, and were they measured in the trials? | Survival, tumour response, progression free survival, QOL and toxicity are the most important outcomes and all were measured in the REFLECT trial using appropriate methodology. |
| • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? | NA |
| • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently? | No |
| 20. Are you aware of any relevant evidence that might | No |

| | |
|---|---|
| not be found by a systematic review of the trial evidence? | |
| 21. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TA474]? | Yes. In the first line setting, there have been two more negative trials published, SARAH and SIRveNIB, both of which failed to show survival benefit for selective internal radiotherapy (SIRT) over sorafenib (Vilgrain V et al, Lancet Oncol. 2017 Dec;18(12):1624-1636, Chow PKH et al, J Clin Oncol. 2018 Mar 2 on line).<br><br>Therefore sorafenib remains the only first line therapy for advanced HCC. |
| 22. How do data on real-world experience compare with the trial data? | Provided patients with preserved liver function and good performance status are treated, the benefits of lenvatinib should be replicated in the real-world setting. The trial excluded patients with main portal vein involvement and those with >50% tumour in the liver and the absolute benefit in these subgroup is unknown. |
| **Equality** | |
| 23a. Are there any potential equality issues that should be taken into account when considering this treatment? | HCC is much more common in men and this was reflect in the trial population in which 85% were male. |

| | |
|---|---|
| 23b. Consider whether these issues are different from issues with current care and why. | No. |
| **Topic-specific questions** | |
| 24. Would treatment with lenvatinib be considered in people with ECOG performance status >1 or Child-Pugh grade B liver impairment? | These patients were excluded from the trial and the benefit and toxicity in these groups is not known. Post marketing surveys of sorafenib have demonstrated that the survival of patient with Child Pugh B and PS >1 is poor but there are no randomised data to determine the absolute difference in survival compared to supportive care King J et al, Clin Oncol 2017 (4):256-262). |
| 25. Is best supportive care used in clinical practice given that TA474 recommends sorafenib as a treatment option for treating advanced hepatocellular carcinoma for people with Child-Pugh grade A liver impairment? | Yes. Those patients who do not receive sorafenib or lenvatinib would receive best supportive care. |

| **Key messages** |
| --- |
| 26. In up to 5 bullet points, please summarise the key messages of your statement.<br>• The survival of patients treated with lenvatinib is non-inferior to those treated with sorafenib.<br>• Lenvatinib was superior to sorafenib with respect to clinically relevant secondary endpoints including tumour response, progression free survival and quality of life.<br>• The side effect profile for lenvatinib is different from sorafenib but this does not seem to affect QOL adversely by comparison.<br>• There are sub-groups of patients in whom sorafenib is relatively ineffective and in whom lenvatinib is more effective, and they may derive particular benefit from lenvatinib.<br>• The REFLECT was the first front-line trial in advanced HCC to meet its endpoint in 10 years and represents a significant advance in the treatment of advanced HCC. |

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.

NICE National Institute for Health and Care Excellence

## Patient expert statement

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

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 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 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 10 pages.

| **About you** | |
|---|---|
| 1.Your name | **Charles Gore** |
| 2. Are you (please tick all that apply): | ☐ a patient with the condition?<br>☐ a carer of a patient with the condition?<br>☒ a patient organisation employee or volunteer? |

| | |
|---|---|
| | ☐ other (please specify): |
| 3. Name of your nominating organisation | The Hepatitis C Trust |
| 4. Did your nominating organisation submit a submission? | ☒ yes, they did<br>☐ no, they didn't<br>☐ I don't know |
| 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<br>☐ no, I disagree with it<br>☐ I agree with some of it, but disagree with some of it<br>☐ 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 rest of this form will be deleted after submission.) | ☒ yes |
| 7. How did you gather the information included in your statement? (please tick all that apply) | ☐ I have personal experience of the condition<br>☐ I have personal experience of the technology being appraised<br>☒ I have other relevant personal experience. Please specify what other experience:<br>☒ I am drawing on others' experiences. Please specify how this information was gathered:<br>I have provided support to people with this condition both here and abroad. In addition we have a national hel0pline that provides support to people with HCC |
| **Living with the condition** | |
| 8. What is it like to live with the condition? What do carers experience when caring for someone with the condition? | Unresectable HCC is rapidly fatal. It is therefore frightening for those with it and induces helplessness in carers, especially if they are young |
| **Current treatment of the condition in the NHS** | |
| 9. What do patients or carers think of current treatments and | Not good enough |

| | |
|---|---|
| care available on the NHS? | |
| 10. Is there an unmet need for patients with this condition? | Yes to prolong life, to prolong progression free survival and ultimately a cure |
| **Advantages of the technology** | |
| 11. What do patients or carers think are the advantages of the technology? | Please see HCT submission |
| **Disadvantages of the technology** | |
| 12. What do patients or carers think are the disadvantages of the technology? | **Please see HCT submission** |
| **Patient population** | |
| 13. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and | **Please see HCT submission** |

| | |
|---|---|
| explain why. | |
| **Equality** | |
| 14. Are there any potential equality issues that should be taken into account when considering this condition and the technology? | Please see HCT submission |
| **Other issues** | |
| 15. Are there any other issues that you would like the committee to consider? | **Please see HCT submission** |
| **Key messages** | |
| 17. In up to 5 bullet points, please summarise the key messages of your statement:<br>• **Please see HCT submission**<br>• <br>• <br>• <br>• | |

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.

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma

## STA REPORT

This report was commissioned by the NIHR HTA Programme as project number 16/56/18

BMJ Technology Assessment Group

**Title:** Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma


| | |
|---|---|
| **Produced by:** | BMJ Technology Assessment Group (BMJ-TAG) |
| **Authors:** | Steven J Edwards, Director of Health Technology Assessment, BMJ-TAG, London<br><br>Kayleigh Kew, Health Technology Assessment Analyst, BMJ-TAG, London<br><br>Peter Cain, Health Economist, BMJ-TAG, London<br><br>Samantha Barton, Senior Health Technology Assessment Analyst, BMJ-TAG, London<br><br>Gemma Marceniuk, Health Economist, BMJ-TAG, London |
| **Correspondence to:** | Steve Edwards, BMJ-TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JP. |
| **Date completed:** | 01/05/2018 |



**Source of funding:** This report was commissioned by the NIHR HTA Programme as project number 16/56/18


**Declared competing interests of the authors:**
No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.


**Acknowledgements:**
The ERG would like to thank Professor Massimo Pinzani (Clinical Hepatologist, Royal Free Hospital) and Dr Pankaj Punia (Consultant Medical Oncologist, University Hospital Birmingham) for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report.




**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:** Edwards SJ, Kew KM, Cain P, Barton S, Marceniuk G. Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma: A Single Technology Appraisal. BMJ Technology Assessment Group, 2018.

**Contributions of authors:**

| | |
|---|---|
| Steve Edwards | Critical appraisal of the company's submission; provided feedback on all versions of the report. Guarantor of the report |
| Kayleigh Kew | Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and drafted the summary, background and clinical results sections |
| Samantha Barton | Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and assisted with drafting the clinical results sections |
| Peter Cain | Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections |
| Gemma Marceniuk | Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections |

All authors read and commented on draft versions of the ERG report.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# TABLE OF ABBREVIATIONS

| Abbreviation | In full |
|---|---|
| AFP | Alpha fetoprotein |
| AIC | Akaike Information Criteria |
| BCLC | Barcelona Clinic Liver Cancer |
| BIC | Bayesian Information Criteria |
| BNF | British National Formulary |
| BOR | Best overall response |
| BSA | Body surface area |
| BSC | Best Supportive Care |
| CEAC | Cost-effectiveness acceptability curve |
| CGP | Corrected group prognosis |
| CHMP | Committee for Medicinal Products for Human Use |
| CI | Confidence interval |
| CONSORT | Consolidated Standards of Reporting Trials |
| CR | Complete response |
| CS | Company's submission |
| CSR | Clinical study report |
| CTCAE | Common Terminology Criteria for Adverse Events |
| EASL | European Association for the Study of the Liver |
| ECOG PS | Eastern Cooperative Oncology Group Performance Status |
| EMA | European Medicines Agency |
| EORTC QLQ-30 | European Organisation for Research and Treatment of Cancer generic cancer quality of life scale |
| EORTC QLQ-HCC18 | EORTC hepatocellular carcinoma-specific module |
| ERG | Evidence Review Group |
| ESDO | European Society of Digestive Oncology |
| ESMO | European Society for Medical Oncology |
| EQ-5D-(3L) | EuroQol-5 Dimensions (3 level version) |
| FACT-G | Functional Assessment of Cancer Therapy-General |
| FAS | Full Analysis Set |
| FDA | Food and Drug Administration |
| FGF (R) | Fibroblast growth factor (receptor) |
| HCC | Hepatocellular carcinoma |
| HCHS | Hospital and community health service |
| HR | Hazard ratio |
| HRQoL | Health-related quality of life |
| HSUV | Health state utility value |
| HTA | Health Technology Appraisal |
| ICER | Incremental cost-effectiveness ratio |
| IIR | Independent imaging review |
| INR | International normalized ratio |
| ITT | Intention-to-treat |
| IV | Intravenous |
| K–M | Kaplan–Meier |
| LY | Life year |
| MOC | Mean of covariates |

| | |
|---|---|
| mRECIST | Modified Response Evaluation Criteria in Solid Tumours |
| MRI | Magnetic resonance imaging |
| NHS | National Health Service |
| NICE | National Institute for Health and Care Excellence |
| NMB | Net monetary benefit |
| NYHA | New York Heart Association |
| OR | Odds ratio |
| ORR | Objective response rate |
| OS | Overall survival |
| PAS | Patient Access Scheme |
| PD | Progressive disease |
| PDGFR | Platelet-derived growth factor receptor |
| PFS | Progression-free survival |
| PH | Proportional hazard |
| PPES | Palmar-plantar erythrodysaesthesia syndrome |
| PR | Partial response |
| PSA | Probabilistic sensitivity analysis |
| PRISMA | Preferred reporting items for systematic reviews and meta-analyses |
| PSSRU | Personal and Social Services Research Unit |
| PY | Patient year |
| Q | Quartile |
| QALY | Quality-adjusted life year |
| QoL | Quality of life |
| RCT | Randomised controlled trial |
| RD | Risk difference |
| RECIST | Response Evaluation Criteria in Solid Tumours |
| RTK | Receptor tyrosine kinase |
| SAE | Serious adverse event |
| SAP | Statistical Analysis Plan |
| SAS | Safety analysis set |
| SD | Standard deviation |
| SLR | Systematic literature review |
| SmPC | Summary of medical product characteristics |
| STA | Single technology appraisal |
| TACE | Transcatheter arterial chemoembolisation |
| TEAE | Treatment-emergent adverse event |
| TTD | Time to discontinuation |
| TTP | Time to progression |
| TSD | Technical Support Document |
| UK | United Kingdom |
| VAS | Visual analogue scale |
| VEGF (R) | Vascular endothelial growth factor (receptor) |

# 1 SUMMARY

## *1.1 Critique of the decision problem in the company's submission*

The company of lenvatinib ([redacted] Eisai Co., Ltd) submitted to the National Institute for Health and Care Excellence (NICE) clinical and economic evidence in support of the effectiveness of lenvatinib in the treatment of adults with unresectable hepatocellular carcinoma (HCC) who have not previously received systemic treatment.

The company provides a reasonable overview of the disease area and current service provision. Briefly, HCC is the most common primary liver cancer in England and is associated with a poor prognosis (5-year survival less than 15%). HCC commonly occurs in the presence of liver cirrhosis, and major risk factors include chronic alcohol consumption, hepatitis B or hepatitis C infection, and non-alcoholic fatty liver disease. Severity and prognosis of HCC is commonly categorised with the Barcelona Clinic Liver Cancer (BCLC) system, which helps to guide choice of treatment. Sorafenib is the only systemic therapy approved by NICE as an option for treating HCC, and is only recommended for patients with advanced disease (BCLC stage C), and well preserved liver function (Child–Pugh grade A).

Lenvatinib, which the company positions as an alternative to sorafenib, inhibits the activity of multiple vascular endothelial growth factor (VEGF) and other tyrosine kinase receptors implicated in tumour growth and spread. Lenvatinib's marketing authorisation covers its use in combination with everolimus for advanced renal cell carcinoma and differentiated thyroid carcinoma but, at the time of writing, the extension for use in HCC had not been approved; the draft indication is for the treatment of adult patients who have received no prior systemic therapy for HCC.

The clinical evidence presented in the company's submission (CS) is derived from REFLECT, an international phase III, open-label randomised controlled trial (RCT). REFLECT enrolled patients with previously untreated, advanced or unresectable HCC, which is in line with the final scope issued by NICE. The comparator (sorafenib) and outcomes in REFLECT were also aligned with the final scope, and are relevant to UK clinical practice. While also listed in the scope, the company does not provide evidence comparing lenvatinib with best supportive care (BSC), which the ERG's clinical experts considered appropriate; BSC is not a relevant comparator because it is only given when a patient refuses, or is considered unfit for, systemic treatment. The Evidence Review Group's (ERG's) clinical experts consider REFLECT to provide evidence that is relevant to the decision problem, but the trial's inclusion criteria mean results may not be generalisable to patients with compromised liver function (Child–Pugh stage B or worse) or poor performance status.

## 1.2 Summary of clinical effectiveness evidence submitted by the company

The company's systematic literature review (SLR) identified the international phase III, open-label RCT, REFLECT, as the only study providing direct evidence for the population, intervention, comparator (sorafenib) and outcomes outlined in the final scope issued by NICE:

- Population: 954 adults with previously untreated, advanced or unresectable HCC. Patients were Child–Pugh class A and primarily BCLC stage B. Around two-thirds were from the Asia–Pacific region (N = 640), and a third from the Western region (N = 314), including 20 patients from the UK;

- Intervention: lenvatinib once daily oral dose of 8 mg for patients weighing less than 60 kg, and 12 mg for patients weighing 60 kg or greater;

- Comparator: sorafenib twice daily oral dose of 400 mg;

- Outcomes: overall survival (OS), progression-free survival (PFS), time to progression (TTP), response rates, adverse effects and health-related quality of life (HRQoL). The primary objective of REFLECT was to show non-inferiority of lenvatinib compared with sorafenib in terms of OS, with an upper 95% CI of 1.08. The company's statistical approach was generally appropriate and well described. However, proportional hazards (PH) assessments conducted by the company suggest the assumption does not hold for OS or PFS, so HRs from Cox PH models should be interpreted with caution.

Potential oversights in the company's SLR with regards to contextual evidence are unlikely to impact the submission given the robustness and maturity of evidence from RELFECT.

The ERG's clinical experts consider REFLECT to provide evidence that is relevant to the decision problem, but highlight that the inclusion criteria mean results may not generalise to those with compromised liver function (Child–Pugh stage B or worse) or Eastern Cooperative Oncology Group (ECOG) performance status of 2 or above.

Lenvatinib met criteria for non-inferiority to sorafenib in the primary analysis of OS (full population, adjusted for stratification factors). At the final data cut-off, 73.4% of the lenvatinib group and 73.5% of the sorafenib group had died, and median OS in the full population was 13.6 months and 12.3 months, respectively (hazard ratio [HR] 0.92, 95% confidence interval [CI]: 0.79 to 1.06).

Median PFS and TTP were consistently longer in the lenvatinib group than the sorafenib group, regardless of the measure used (modified Response Evaluation Criteria in Solid Tumours [mRECIST],

or RECIST v1.1) or outcome assessor (investigator or independent imaging review [IIR]). The primary analyses were investigator-assessed with mRECIST (a validated modified set of criteria for HCC) and indicated median PFS of 7.4 months for lenvatinib and 3.7 months for sorafenib (HR 0.66, 95% CI: 0.57 to 0.77), and corresponding TTP of 8.9 months and 3.7 months, respectively (HR 0.63, 95% CI: 0.53 to 0.73).

The ERG considers rules for censoring in the company's primary analyses of PFS and TTP (based on Food and Drug Administration [FDA] guidance) to favour lenvatinib (see Section 1.5.2). An alternative PFS analysis was provided at the clarification stage based guidance from the European Medicines Agency (EMA); median PFS in both groups was [redacted] impact on the relative treatment effect [redacted].

While TTP is often preferred over PFS for HCC because cirrhosis-related deaths can mask drug benefit in PFS, the similarity of group medians and HRs between TTP and PFS may reflect that patients have well-preserved liver function at baseline in REFLECT.

Objective response rate (ORR) was also higher with lenvatinib than with sorafenib in the primary investigator-assessed analysis based on mRECIST (24.1% vs 9.2%; odds ratio [OR] 3.13, 95% CI: 2.15 to 4.56, p<0.00001). The magnitude of effect varied between the primary analysis and IIR assessments (mRECIST and RECIST v1.1), but the difference was statistically significant and clinically meaningful regardless of the measure chosen.

Across HRQoL scales, there appears to be a reduction in some symptom and treatment-emergent adverse event (TEAE) with lenvatinib compared to sorafenib ([redacted], role functioning, pain, body image), and vice versa ([redacted]), but results [redacted].

Most subgroup and supplementary analyses of OS and PFS were consistent with the primary analyses. Notable results that were inconsistent with the primary analyses included OS in the Western geographical subgroup, for which the HR lay in favour of sorafenib (HR 1.08, 95% CI: 0.82 to 1.42), [redacted]; possible reasons for the difference in effect are discussed with reference to the applicability of the full population to UK clinical practice in Section 1.5.2. While the PH assumption does not hold for PFS, the 95% CI for the HR indicate that the benefit of lenvatinib is not statistically significant in the Western subgroup (HR 0.81, 95% CI: 0.61 to 1.08), those with hepatitis C aetiology (HR 0.78, 95% CI: 0.56, 1.09), and other strata that included a small number of patients (age ≥75 years and female).

Almost all patients in both groups experienced at least one TEAE (98.7% and 99.4% for lenvatinib and sorafenib, respectively), and the proportion considered by the investigators to be related to study treatment was also similar (93.9% and 95.2%). The ERG considered it more appropriate to compare incidence over the full course of treatment, rather than exposure-adjusted rates, because the former is consistent with the effectiveness evidence.

Drug withdrawals due to TEAEs (19.7% vs 14.5% for lenvatinib and sorafenib, respectively), dose modifications due to TEAEs (61.8% vs 55.6%), SAEs (43.1% vs 30.3%) and Grade 3+ TEAEs (75.0% vs 66.5%; 56.7% vs 48.6% judged by the investigators to be related to study treatment) were all more common in the lenvatinib group than the sorafenib group.

AEs of any grade that were notably more frequent in the lenvatinib group than the sorafenib group ([redacted]) were hypertension (42.2% vs 30.3%), decreased appetite (34.0% vs 26.7%), weight decreased (30.9% vs 22.3%), proteinuria (24.6% vs 11.4%), dysphonia (23.7% vs 12.0%), nausea (19.5% vs 14.3%), platelet count decreased (18.3% vs 12.2%), hypothyroidism (16.4% vs 1.7%), vomiting (16.2% vs 7.6%), constipation (16.0% vs 10.9%), ascites (14.3% vs 9.3%), peripheral oedema (13.9% vs 6.9%), and back pain (10.5% vs 6.5%). TEAEs of any grade experienced by notably fewer patients in the lenvatinib group than the sorafenib group, respectively, were diarrhoea (38.7% vs 46.3%), palmar-plantar erythrodysaesthesia syndrome (PPES; 26.9% vs 52.4%), rash (9.7% vs 16.0%), and alopecia (2.9% vs 25.1%).

Grade 3–4 TEAEs that were notably more frequent (95% CI for RRs [redacted]) in the lenvatinib group than the sorafenib group were hypertension (23.3% vs 14.3%), weight decrease (7.6% vs 2.9%) and proteinuria (5.7% vs 1.7%). PPES was the only grade 3–4 TEAE that was less common in the lenvatinib group than the sorafenib group (2.9% vs 11.4%).

## *1.3 Summary of cost effectiveness evidence submitted by the company*

The company submitted a *de novo* economic model in Microsoft® Excel comparing the cost effectiveness of lenvatinib and sorafenib in patients with untreated, advanced or unresectable HCC who have Child–Pugh class A status. The key data source for the model was the REFLECT RCT, which compared lenvatinib and sorafenib in a patient population consisting of a Western subgroup (32.9%) and an Asia–Pacific subgroup (67.1%).

The company used a partitioned survival modelling approach, using PFS and OS data to inform the health states of progression-free disease and death. An additional health state for progressed disease was informed by the difference between PFS and OS. The proportion of patients in each state was estimated in cycles of 28 days over a time horizon of 20 years.

To extrapolate beyond the trial follow-up period, the company fitted parametric survival curves to the PFS and OS data using exponential, Weibull, Gompertz, log-logistic, lognormal and generalised gamma functions. Suitable model fit was judged using the Akaike Information Criterion and Bayesian Information Criterion. Visual fit of the curves compared to the Kaplan–Meier data was not done as the company adjusted the survival models to account for imbalances in covariates.

To select variables for the adjustment of the parametric models, the company firstly sought clinical expert opinion to determine which variables should be considered as prognostic of outcomes. The company then used these covariates to fit a Cox PH model to the OS data with variable selection refined further using backwards stepwise automated variable selection procedure. This approach removes covariates with a non-statistically significant coefficient in a stepwise fashion until the coefficients of all the remaining variables are significant. The resulting set of variables following the selection procedure was then used directly to adjust each of the parametric survival curves for both PFS and OS, regardless of whether the variables still had a statistically significant coefficient or not. The company also provided an additional set of curves with only alpha fetoprotein (AFP) levels included as a covariate, and an additional set of OS curves with a binary post-progression treatment variable included in the adjustment, but these were not used in the company's base case.

Time to treatment discontinuation data were almost complete in the REFLECT trial, with the exception of 4% of patients remaining on treatment at the end of follow-up in the sorafenib group. These data were used directly by the company to inform the proportion of patients who were on treatment at any given time in the model. The company assumed that the 4% remaining on treatment at the end of the trial discontinued treatment immediately after the follow-up period.

Treatment costs were included based on the dosing received in the REFLECT trial. The dosing schedules were in line with the Summary of Product Characteristics (SmPCs) for lenvatinib and sorafenib. The mean doses in the full trial population were used. Disease management resource use and costs were informed by two surveys of clinical experts conducted by the holder of the marketing authorisation for sorafenib. These were conducted for TA189 and an updated survey was conducted for the reconsideration in TA474. The resources use was elicited for both the progression-free state and the progressed disease state, which had a greater resource requirement and, therefore, greater costs. The costs associated with the management of adverse events (AEs) were restricted to grade 3 and 4 events that occurred in at least 5% of patients in either group of the REFLECT trial. Following clarification questions, the company also provided estimates of resource potentially required for the management of grade 5 AEs, and included a scenario analysis to include these costs.

Health-state utility values (HSUVs) were elicited during the REFLECT trial and the full population values were used and were estimated for the progression-free state and the progressed disease state. The

company used a linear mixed model to adjust for imbalances in covariates, but the selection of variables was not done systematically. However, the company determined that there was no significant effect by treatment group or by sub-population and, hence, the company used the full population values to determine the HSUVs.

The results (without the sorafenib patient access scheme applied) of the company's base case analysis showed that lenvatinib was, overall, the cheaper and more effective option, and therefore, it dominated sorafenib. The company provided a range of scenario analyses and sensitivity analyses as well as a probabilistic sensitivity analysis.

## 1.4 Summary of the ERG's critique of cost effectiveness evidence submitted

The company's economic model was generally sound, easy to navigate and included a range of options to perform alternative scenario analyses. However, there were a few issues with the company's methods that, in the opinion of the ERG, make the results of the company's base case analysis potentially unreliable.

The first of these issues is the company's approach used to adjust the parametric survival models. The ERG considers it inappropriate to fit a Cox PH model to the data to select a set of variables to use in the adjustment of the parametric models. The underlying model used when applying an automated variable selection procedure may influence whether a variable is selected or not, and, therefore, the ERG considers it more appropriate for the variable selection to be performed using each of the parametric models individually. The ERG considers the company's use of the OS Cox PH model to also inform variable selection for the PFS parametric models to add further uncertainty to the results. The magnitude of the impact that these issues would have on the model is unclear; however, the ERG considers the company's adjusted models to be preferable to the unadjusted models. The ERG also notes that the results of applying the selection procedure on each parametric model individually may not have any impact at all. The company's approach does add uncertainty to the conclusions, but given that the data are from an RCT with only some imbalances in covariates, that may give some reassurance that the adjustments made are sufficient.

Another issue that the ERG considers to be inappropriate in the company's base case analysis is the choice of parametric curve for the modelling of PFS. Although the ERG agrees with the company's decision to use the same parametric form for each treatment group, the ERG disagrees that the lognormal distribution provides the most suitable fit overall. The company chose the lognormal as it provided the best fit for the lenvatinib group; however, the goodness-of-fit for the sorafenib group was poor. The ERG considers the generalised gamma to be more suitable as it provided a similar fit to the

lognormal curve for the lenvatinib group and provided the best fit to the sorafenib group, and, therefore, overall the ERG considers it a better fitting model.

The issues with PFS censoring described in Section 1.5.2 add further uncertainty to the modelling; however, the impact of this is difficult to quantify. As the company declined a request to provide patient level data, the ERG could not perform further analysis to update its preferred base case. However, the ERG considered a range of scenario analyses to assess the impact of reducing the scale parameter of the generalised gamma function to reduce the relative treatment effect arbitrarily. The results of these analyses, along with plots of the resulting PFS curves for comparison are presented in Section 6.4. The ICERs resulting from varying the reduction in the scale parameter by between 5% and 15%, ranged from £2,085 to £14,024 per QALY, respectively.

The ERG also identified a few other minor issues regarding subsequent treatment costs and AEs included; however, these issues were found to have only a minor impact on the model results or did not impact the base case analysis. Discussion of these issues can be found throughout Section 5 of this report.

## *1.5 ERG commentary on the robustness of evidence submitted by the company*

### 1.5.1 Strengths

***Clinical***

- The submission is based on head-to-head evidence from an international phase III, open-label RCT, which is closely aligned with the NICE final scope, meaning an indirect comparison was not required to derive estimates of relative treatment effectiveness.

- Data for all clinically relevant outcomes are available from REFLECT, and data for key survival outcomes (OS, PFS and TTP) are mature.

- A retrospective IIR of tumour assessments was conducted for PFS, TTP and ORR, which mitigates the risk of bias associated with the open-label design and investigator-assessed outcomes.

- The company's statistical approach was generally appropriate and well described. Where *post-hoc* analyses were conducted to test the robustness of treatment effects, these were justified fully with reference to the statistical analysis plan.

- The company provided additional data and conducted a range of supplementary and subgroup analyses that allowed the ERG to assess the robustness of relative treatment effects (including the impact of subsequent treatments, rules for censoring, and geographical subgroups).

***Economic***

- The company submission was well written with the methods and analyses described succinctly and transparently. The economic model appeared to be well constructed and was easy to use.

- The company provided a range of analyses to assess the impact of alternative approaches and these were provided as options within the economic model.

### 1.5.2 Weaknesses and areas of uncertainty

***Clinical***

- Baseline imbalances that may favour the sorafenib group were noted in the proportion of patients with AFP levels ≥200 ng/ml and in aetiology of HCC (hepatitis B, hepatitis C, or alcohol). Adding the variables as covariates in supplementary analysis of OS had some impact on the relative treatment effect.

- The primary analyses of PFS and TTP were based on investigator assessments and may be prone to bias because REFLECT was an open-label study. Alternative analyses were available from retrospective IIR assessments, and results were similar.

- A version of RECIST modified to evaluate change more accurately in HCC (mRECIST) was used for the primary tumour assessments for PFS, TTP and ORR in REFLECT. mRECIST is endorsed by the European Association for the Study of the Liver (EASL), and the ERG's clinical experts advised that UK centres choose either mRECIST and RECIST v1.1. IIR results were available for comparison, and meaningful differences were only noted for ORR, which did not inform the economic model.

- The ERG was concerned that the majority Asia–Pacific population may reduce generalisability of the full REFLECT population to UK patients. The following differences were noted, but the ERG did not consider there to be sufficient evidence that results from the Western subgroup were more applicable to UK clinical practice than the mixed full population to justify loss of precision by focussing on a subgroup:

    - *Baseline characteristics*: compared with the full population, the Western subgroup were heavier (~13% < 60 kg vs ~30%), had more heart disease (~23% New York Heart

Association class I or II vs ~10%), less underlying cirrhosis (40% vs 50%), less hepatitis B (20–25% vs ~50%), and more hepatitis C, alcohol, other, or unknown aetiology. Hepatitis C aetiology has been linked to increased benefit with sorafenib, so the full population may therefore overestimate the relative effect of lenvatinib for UK patients. However, evidence from separate Western and Asia–Pacific trials for sorafenib in HCC (SHARP and the Asia–Pacific study) suggest that the pattern of baseline differences does not necessarily impact relative treatment effects;

- *Treatment exposure*: Treatment duration was longer in the lenvatinib group than the sorafenib group in the full population and the difference between groups was smaller in the Western subgroup. Time on treatment and dose intensity was somewhat higher in both groups in the Western subgroup than the full population (Table 6), which is likely due to weight differences between the Western and Asia–Pacific subgroups;

- *Adverse events*: comparing Grade 3+ TEAEs with the full population, the Western subgroup experienced more fatigue and asthenia, and less decrease platelet count than the full population, regardless of treatment received. Differences between the incidence of decreased weight, proteinuria and PPES that favoured sorafenib in the full population were smaller in the Western subgroup;

- *Subsequent treatments*: neither the full population nor the Western subgroup were considered reflective of UK clinical practice with regards to the extent and type of subsequent treatments received. There was more imbalance in the Western subgroup, mostly due to patients in the sorafenib group being eligible for clinical trials after discontinuation of the study drug;

- *Clinical outcomes*: Results for OS and PFS in the Western subgroup were both less favourable than results for the full population, and less precise due to the restricted population. As may be expected from pattern of imbalance in subsequent interventions described above, results for the full population and Western subgroup [redacted] when the analyses included a binary variable adjustment for patients who received subsequent anti-cancer interventions.

- PH assessments conducted by the company suggest the assumption does not hold for OS or PFS in REFLECT, so hazard ratios calculated from Cox PH models should be interpreted with caution; independent statistical models were considered for both PFS and OS in the review of cost-effectiveness.

- Censoring in the primary analyses of PFS and TTP was based on guidance from the FDA and included censoring at treatment discontinuation if there was no disease progression. The primary analyses are likely to favour lenvatinib because treatment discontinuation for reasons other than progression (that is, due to TEAEs or patient choice) was more common in the lenvatinib group than the sorafenib group. A sensitivity analysis of PFS provided by the company at the clarification stage, based on EMA guidance, included all progressions and deaths as events; median PFS in both groups [redacted] impact on the relative treatment effect [redacted]. Comparing HRs may be unreliable because the PH assumption does not hold and K–M data were not available so the consequences of using this analysis could not be evaluated in the economic model. The ERG considers the consistency in direction of effect to provide robust evidence for a PFS benefit with lenvatinib compared with sorafenib, although the rules for censoring mean the extent of benefit may be [redacted] by the primary PFS analysis in the economic model.

- In the lenvatinib group, median time on treatment (5.7 months) was shorter than the primary analyses of median PFS and TTP (7.4 to 8.9 months, respectively), which may reflect patients discontinuing for reasons other than disease progression (TEAEs and patient choice); median PFS, TTP and time on treatment were equivalent in the sorafenib group (3.7 months).

- The ERG advises caution in the interpretation of exposure-adjusted TEAE rates, because they are less likely than incidence to capture burden of TEAEs across the full course of treatment. Adjusting for exposure causes inconsistency with the effectiveness analyses, and does not account for the tendency of TEAEs to emerge in the early stages of treatment.

***Economic***

The key areas of weakness identified by the ERG relate to the covariate adjustments for the parametric survival models and the censoring applied to the PFS analysis that potentially overestimates the PFS benefits for lenvatinib compared to sorafenib.

The ERG considers that applying a more appropriate covariate adjustment approach, i.e. selecting variables based on the coefficients estimated for each parametric model individually, may not have a great impact, but it still adds to the uncertainty. The ERG considers that, given the analysis is based on an RCT with only some imbalances in the covariates, this may provide some reassurance that the adjustments made by the company are sufficient.

The censoring applied in the PFS analysis is a key source of uncertainty for which it is difficult to quantify the impact. Given the increased disease management costs incurred for treating patients in the post-progression health state, the potentially overestimated PFS could have an important impact on the difference in costs as well as QALYs. The ERG attempted to assess the potential impact of this by reducing the PFS benefit of lenvatinib by arbitrarily reducing the scale parameter of the generalised gamma curve in the ERG's base case analysis. The results of this are given in Section 6.2.

## 1.6 Summary of exploratory and sensitivity analyses undertaken by the ERG

***Economic***

The ERG made a minor correction to the company's base case analysis to ensure that all costs were incorporated within the half-cycle correction in the model. This had only a small impact on the company's ICER.

The ERG then performed a range of scenario analyses around the company's corrected base case including two key changes that formed the ERG's preferred base case. These were to adjust the OS models for subsequent anti-cancer interventions instead of just applying the costs of them, and the second was to use the generalised gamma function to model PFS (including a correction to prevent the curves for each treatment group crossing).

The impact of the first change increased the OS benefit in favour of lenvatinib because there was a greater use of subsequent treatments in the sorafenib group, causing OS to be overestimated more than for the lenvatinib group. The cost difference was less favourable to lenvatinib after this change because of the removal of the subsequent treatment costs that were greater in the sorafenib group. The results of the ERG's base case are presented in Table A with changes incorporated cumulatively.

The ERG also conducted further scenario analyses around the company's corrected base case; a key one being the ERG's application of a mortality adjustment to the company's estimation of newly progressed. The ERG considered the company's approach, to take the difference in PFS between cycles, to be an overestimate of newly progressed patients as it did not factor in those who leave the PFS state because of death rather than progression. This, therefore, potentially overestimated subsequent treatment costs. This change did not apply to the ERG's base case as subsequent treatment costs were removed.

Table A. ERG base case ICER

| Results per patient | Sorafenib (1) | Lenvatinib (2) | Incremental value (2-1) |
|---|---|---|---|
| **Company's corrected base case** | | | |
| Total costs (£) | £65,574 | [redacted] | [redacted] |
| QALYs | 1.03 | [redacted] | [redacted] |
| ICER | | | Dominant |
| **Post-progression adjustment to OS and removal of post-progression therapy costs** | | | |
| Total costs (£) | £60,243 | [redacted] | [redacted] |
| QALYs | 0.95 | [redacted] | [redacted] |
| **ICER (compared with base case)** | | | Dominant |
| **ICER with all changes incorporated** | | | Dominant |
| **Gamma distribution for PFS (with prevention of curves crossing)** | | | |
| Total costs (£) | £56,237 | [redacted] | [redacted] |
| QALYs | 0.96 | [redacted] | [redacted] |
| **ICER (compared with base case)** | | | Dominant |
| **ICER with all changes incorporated** | | | Dominant |
| **ERG's preferred base case ICER** | | | **Dominant** |
| Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALYs, quality-adjusted life years. | | | |

The ERG also conducted a range of scenario analyses to reduce the PFS benefit for lenvatinib to account for the censoring of treatment discontinuation applied by the company in the PFS analysis. This was done by reducing the scale parameter of the generalised gamma function applied in the ERG's preferred base case analysis. This was done in increments of 5% up to 15%, and resulted in ICERs ranging from £2,085 to £14,024 per QALY, respectively. The ERG also conducted a scenario that applied the generalised gamma function for OS too. This function provided a similarly good fit to the data but with a more conservative extrapolation. This scenario resulted in lenvatinib dominating sorafenib as per the company's base case analysis. The results of each of these analyses can be found in full in Section 6.2.

# 2 BACKGROUND

## 2.1 Critique of company's description of underlying health problems

Section B.1.3 of the company submission (CS) provides an overview of the key aspects of hepatocellular carcinoma (HCC), including prevalence, prognosis and disease management. The final scope issued by the National Institute for Health and Care Excellence (NICE) for this Single Technology Appraisal (STA) indicates the population of interest to be adults with unresectable HCC who have not previously received systemic treatment.[1]

Overall, the Evidence Review Group (ERG) considers the CS to present an appropriate overview of HCC that is relevant to the decision problem. The ERG considers that expansion of the reporting of the system used to stage HCC would aid in understanding the evidence base submitted in support of the clinical effectiveness of lenvatinib in HCC. Together with a synopsis of information from the CS, the ERG provides additional detail on the categorisation of stage of HCC.

As outlined in the CS, with some supplementary information provided by the ERG:

- HCC starts in the main type of liver cell, the hepatocellular cell, and is the most common type of primary liver cancer in England, accounting for 55.3% of male and 28.2% of female primary liver cancer diagnoses between 1998 and 2007;[2]
- In 2015, there were with 2,456 new diagnoses of HCC in England;[3]
- Incidence of liver cancer in the UK has increased by 142% since the early 1990s, and the rate is predicted to increase from 9.5 to 15 cases per 100,000 people between 2014 and 2035;[4]
- HCC prevalence increases with age, and, as noted above, is more common in men than women;
- 70–90% of HCC cases occur in the presence of liver cirrhosis;[5, 6]
- Major risk factors for liver cancer are those that result in chronic hepatic injury, and eventually cirrhosis, including chronic alcohol consumption, infection with hepatitis B or hepatitis C, and non-alcoholic fatty liver disease;
- In the UK, the most common causes of liver cirrhosis are non-alcoholic fatty liver disease (41%; mainly caused by obesity and type 2 diabetes) and alcoholic liver disease (30%): infection with hepatitis B or C is associated with 12–16% of cases of liver cirrhosis in the UK;[7-9]
- Frequently, the early stages of liver cancer are symptomless because the liver can function well even if only a small portion of the organ is working;[10]

- Symptoms associated with advanced stages of liver cancer include weight loss, yellowing of the skin and whites of the eyes, loss of appetite, feeling full after eating small amounts, sickness, and itching;[10]

- Suspected cases of liver cancer are investigated with blood tests and an ultrasound scan;[11]

- On confirmation of HCC, stage of disease is typically assessed using the Barcelona Clinic Liver Cancer (BCLC) system, which considers tumour burden, liver function and performance status (Table 1);[12]

- Factors captured by the BCLC system are thought to affect prognosis of HCC; increased tumour size, higher number of HCC nodules, vascular invasion of tumours, poor performance status and liver function, and elevated serum alpha-fetoprotein levels are typically associated with poorer prognosis;[13-15]

- Tumour size and whether the cancer has spread to other sites are frequently determined using X-ray, computed tomography scan or magnetic resonance imaging.[11]

As part of the discussion of treatment options available for HCC, the CS outlines the five categorisations of the BCLC staging system, and the factors that contribute to BCLC stage. To aid interpretation of the figure available in the CS, the ERG presents an overview of the specific components forming each BCLC stage (Table 1).

Table 1. Overview of the Barcelona Clinic Liver Cancer staging system for liver cancer[12]

| BCLC stage | Tumour size | Performance status score[a] | Child–Pugh class |
|---|---|---|---|
| 0 (very early HCC) | Single tumour ≤2 cm | 0 | A |
| A (early HCC) | • Single tumour ≤5 cm; or<br>• Up to 3 tumours, all of which are ≤3 cm in size | 0 | A or B |
| B (intermediate HCC) | Many tumours in the liver | 0 | A or B |
| C (advanced HCC) | Cancer spread to blood vessels, lymph nodes or other organs | 1 or 2 | A or B |
| D (terminal HCC) | – | 3 or 4 | C |

[a] Performance status is determined by the ECOG/WHO system. Scores are as follows:[16]
- 0 = fully active and able to carry on all pre-disease performance without restriction;
- 1 = restricted in a physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature;
- 2 = ambulatory and capable of all selfcare but unable to carry out any work activities; up and about more than 50% of waking hours;
- 3 = capable of only limited selfcare; confirmed to bed or chair more than 50% of waking hours;
- 4 = completely disabled; cannot carry on any selfcare; totally confined to bed or chair;
- 5= dead.

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; ECOG, Eastern Cooperative Oncology Group; HCC, hepatocellular carcinoma; PS, performance status; WHO, World Health Organization.

Prognosis of HCC is poor, with 5-year survival rates for liver cancer in England being less than 15%[17] Survival rates for liver cancer based on BCLC stage are presented in Table 2.

Table 2. Survival rates for liver cancer based on Barcelona Clinic Liver Cancer staging system[12]

| BCLC stage | Overall survival | |
|---|---|---|
| | Without treatment (median) | With treatment |
| 0 | >3 years | 70–90% at 5 years or more |
| A | 3 years | 50–70% at 5 years or more |
| B | 16 months | 20 months (median) |
| C | 4–8 months | 6–11 months (median) |
| D | <4 months | N/R |

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; N/R, not reported.

Although the CS describes the relationship between Child–Pugh grade and BCLC stage, details on estimation of Child–Pugh score are not available. A description of the components considered in the Child–Pugh system might be useful to the Appraisal Committee. The Child–Pugh system assesses how well the liver is functioning and is used to assess the prognosis of people with liver disease. The tool is based on five clinical measures of liver function:[18]

- total bilirubin level;
- serum albumin level;
- time taken for blood to clot (prothrombin time);
- presence of ascites (collection of fluid in the abdomen);
- brain function (hepatic encephalopathy).

Each function is scored between 1 and 3, with a score of 3 indicating the worst outcome for that measure of liver performance: overall score ranges between 5 and 15. Based on the total score of the five individual parameters, people are categorised as Child–Pugh class A, B, or C, where:

- class A denotes that the liver is working normally (score of 5 or 6);
- class B indicates mild to moderate illness (score of 7–9);
- class C equates to severe liver damage (score of 10–15).

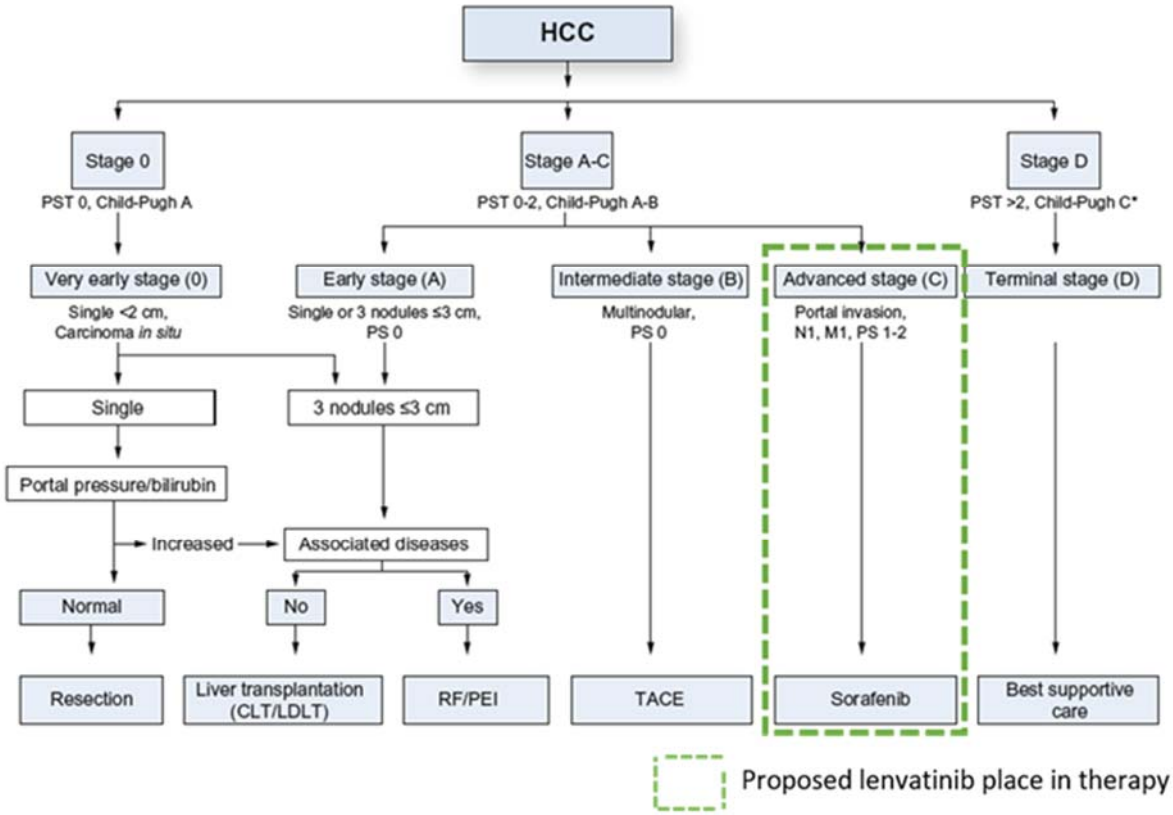
One-year survival for those categorised as Child–Pugh Class A is about 95%, compared with 80% and 44% for those assigned to Classes B and C, respectively.[19]

## 2.2 Critique of company's overview of current service provision

The CS provides a reasonable overview of current service provision for the management of untreated advanced or unresectable HCC. Within the CS, the company refers to guidance on treatment of HCC from the European Association for the Study of the Liver–European Organisation for Research and Treatment of Cancer (EASL–EORTC) and European Society for Medical Oncology–European Society of Digestive Oncology (ESMO–ESDO).[20, 21] The ERG identified UK-specific guidance on management of HCC published in 2003,[22] which does not encompass current treatment options available for HCC.

Within the guidelines from the EASL–EORTC and ESMO–ESDO, choice of therapy is determined by BCLC stage of disease. Treatments for early stage HCC are given with curative intent and include surgery (hepatic resection), and percutaneous radiofrequency/thermal ablation in patients with well-preserved liver function, or liver transplantation for those with impaired liver function (Figure 1). However, for those with an advanced stage of disease, treatment is predominantly given with the goal of managing pain and symptoms and improving quality of life, rather than cure. Options include interventional procedures such as transcatheter arterial chemoembolisation (TACE), selective internal radiation therapy, and external beam radiotherapy.[23] At the time of writing, the only systemic therapy recommended by NICE for treatment of HCC is sorafenib, which is approved for the treatment of advanced HCC but only for those with Child–Pugh grade A liver impairment.[24] A recent publication presents a revised treatment algorithm for HCC substituting sorafenib for 'systemic therapy' as the treatment option for people with advanced stage disease (Figure 1).[25] Sorafenib has yet to be included in guidance from the EASL–EORTC or ESMO–ESDO: updated guidelines from the EASL incorporating sorafenib, lenvatinib and other systemic therapies in the management of HCC are in press at the time of writing.[26] As indicated in Figure 1, the company anticipates that lenvatinib will be used for people who are also eligible for treatment with sorafenib.

Figure 1. BCLC staging and treatment recommendations for HCC with company's proposed position of lenvatinib in treatment pathway (reproduced from company submission, Figure 1 [pg. 14])




Abbreviations: BCLC, Barcelona Clinic Liver Cancer; CLT, cadaveric liver transplantation; EASL-EORTC; European Association for the Study of the Liver-European Organisation for Research and Treatment of Cancer; HCC, hepatocellular carcinoma; LDLT, living donor liver transplantation; PEI, percutaneous ethanol injection; RF, radiofrequency ablation; TACE, transcatheter arterial chemoembolisation.

# 3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

The company provided a summary of the decision problem issued by the National Institute for Health and Care Excellence (NICE), together with the rationale for any deviation from the final scope[1] (Table 3).

Table 3. Summary of decision problem as outlined in the company's submission (reproduced from CS Table 1, pg. 9)

| | **Final scope issued by NICE** | **Decision problem addressed in the submission** | **Rationale if different from the scope** |
|---|---|---|---|
| Population | Adults with unresectable hepatocellular carcinoma who have not previously received systemic treatment | Adults with untreated advanced or unresectable HCC and Child–Pugh Class A liver function. This is consistent with the pivotal lenvatinib RCT (REFLECT) population which included some patients with BCLC stage B (those who were considered ineligible for TACE), with most patients having BCLC stage C disease. | This population is consistent with that of REFLECT, and the lenvatinib licensed indication.*<br>The population is also consistent with the NICE recommended population for the use of sorafenib in HCC, the SHARP study which was the evidence base for this recommendation, and with UK clinical practice as specified in the sorafenib NICE recommendation (i.e. predominantly BCLC stage C (advanced) disease, predominantly good liver function (Child–Pugh Class A) and good ECOG performance status (0–2)). |
| Intervention | Lenvatinib | As per scope | NA |
| Comparator(s) | •Sorafenib<br>•Best supportive care | •Sorafenib | BSC is not considered to be a relevant comparator due to the small numbers of patients in the population defined above (<5% according to a UK clinical expert [see section B.3.3.3 for details]) that would receive this as an alternative to sorafenib. Feedback from UK clinical experts is that in the defined patient population (adults with untreated advanced or unresectable HCC and Child–Pugh Class A liver function), almost all patients would be eligible to receive systemic therapy |
| Outcomes | •Overall survival<br>•Progression-free survival<br>•Time to progression<br>•Response rates<br>•Adverse effects of treatment<br>•Health-related quality of life | As per scope | NA |

Abbreviations used in table: BCLC, Barcelona Clinic Liver Cancer; BSC, best supportive care; ECOG, Eastern Cooperative Oncology Group; EMA, European Medicines Agency; HCC, hepatocellular carcinoma; NA, not applicable; NICE, National Institute for Health and Care Excellence; RCT, randomised controlled trial; SmPC, summary of medical product characteristics; TACE, transcatheter arterial chemoembolisation

The current draft SmPC is presented in Appendix C. The population addressed in this submission and detailed in this table is based on anticipated changes to the licensed indication requested by the EMA rapporteur that have not yet been incorporated into the draft SmPC.

## 3.1 Population

The company submitted clinical effectiveness evidence from the REFLECT phase III randomised controlled trial (RCT).[27] Supporting clinical effectiveness evidence from a phase I/II dose finding study was also submitted,[28] but the ERG has focused its critique on REFLECT because it was the only study used to inform the economic model.

The inclusion criteria for REFLECT match the population outlined in the final scope issued by NICE,[1] that is, adults with unresectable hepatocellular carcinoma (HCC) who have not previously received systemic treatment. The company propose that lenvatinib will be used as an alternative to sorafenib as first-line therapy for patients with advanced or unresectable HCC, Child–Pugh class A liver function, Barcelona Clinic Liver Cancer (BCLC) Stage C disease, or patients with BCLC Stage B disease who were not eligible for transcatheter arterial chemoembolisation (TACE). The ERG notes that patients with BCLC Stage B HCC have multiple liver tumours and are thus not eligible for resection (Figure 1). As such, the ERG considers the inclusion of patients with BCLC Stage B HCC in line with the NICE final scope, given that systemic therapy would be considered if a patient was considered ineligible for the recommended treatment (i.e. TACE; Figure 1).

The ERG's clinical experts considered the key eligibility criteria for REFLECT broadly in line with patient eligibility considerations for systemic therapy in UK clinical practice:

- histologically, cytologically or clinically confirmed HCC according to American Association for the Study of Liver Diseases criteria;
- one or more measurable target lesions based on modified Response Evaluation Criteria for Solid Tumours (mRECIST);
- BCLC stage B (intermediate) or C (advanced), Child–Pugh class A (normal liver function) and Eastern Cooperative Oncology Group (ECOG) 0 or 1 (good performance status);
- survival expectation of 12 weeks or longer after starting study drug;
- less than 50% liver occupation and no invasion into the bile duct or main portal branch;

However, it was noted that REFLECT excluded patients with compromised liver function (Child–Pugh stage B or worse), and those with ECOG performance status of 2 or above, meaning the results of REFLECT may not be generalisable to those patients.

In the company submission (CS), baseline demographic and disease characteristics were provided for the lenvatinib and sorafenib groups of the full REFLECT population (Table 53 and Table 54). The ERG

notes that baseline characteristics were well balanced between groups in the full population, and discusses imbalances highlighted by the company in Section 4.2.2 (namely, proportion of patients with alpha-fetoprotein [AFP] levels ≥ 200 ng/ml and hepatitis C aetiology).

Briefly, in both groups of REFLECT, mean age was 61, around 85% were male, two thirds were from Asia–Pacific regions and a third were from Western regions; ~20% had macroscopic portal vein invasion, 60% had extrahepatic spread, and around half had underlying cirrhosis. Approximately 20% of both groups were BCLC stage B (intermediate) who were ineligible for TACE, and the trial was limited to patients with good ECOG performance status (with 63% having a score of 0). Most patients were Child–Pugh class A (i.e. normal liver function), having a score of 5 (76%) or 6 (23%), and only a small number fell outside the eligibility criteria having Child–Pugh class B liver function (1%). In the full population, hepatitis B was the most common factor of carcinogenesis (52.5% and 47.9% of the lenvatinib and sorafenib groups, respectively), followed by hepatitis C (19.0% and 26.5%, respectively).

Characteristics split by geographic region were also provided in the submission appendices (reproduced in Table 55). The ERG noted marked differences between the Asia–Pacific and Western geographical regions, which may suggest the full population is not generalisable to the UK population, and that the Western subgroup (n = 314; 33%) may be more appropriate for UK decision-making. Compared with the full population, fewer patients in the Western subgroup fell in the lower weight group (~13% < 60 kg compared with ~30% in the full population), more had heart disease (~23% New York Heart Association [NYHA] classification I or II compared with ~10%), fewer had underlying cirrhosis (40% compared with 50%), and hepatitis B was less commonly a factor of carcinogenesis (20–25% compared with ~50%), with more Western patients having hepatitis C, alcohol other, or unknown aetiology (see Table 55). The ERG notes that weight group relates to the dose of lenvatinib (8 mg daily for those < 60 kg and 12 mg for those ≥ 60 kg), and that patients with Hepatitis C aetiology may derive more benefit from sorafenib than those with other aetiologies.[29] There is some evidence that the increased benefit from sorafenib for those with Hepatitis C aetiology is not common across TKIs,[30] which, given the higher percentage of Hepatitis C aetiology in the Western subgroup compared with the full population, may result in an overestimation of lenvatinib benefit for UK patients if evidence from the full population is used. The ERG also notes that the Western subgroup were older, had higher baseline ammonia levels, worse Child–Pugh scores, and shorter time since first diagnosis (Table 55).

In relation to regional differences, the ERG notes that the evidence supporting the NICE appraisal of sorafenib for treating advanced HCC was based on a Western trial population (the SHARP trial[31]), and that a separate trial run in the Asia–Pacific regions[32] was not considered relevant to the decision problem.[24] The sorafenib trials compared the same dose of sorafenib (400 mg twice daily) with placebo and there was a similar pattern of differences between the Western and Asia–Pacific populations as highlighted above between the REFLECT subgroups. The ERG notes the similarity of hazard ratios for

the key outcomes of overall survival (OS) and time to progression (TTP) between the two sorafenib trials, suggesting relative treatment effect was not affected by the pattern of baseline differences between regions. However, median time-to-event was longer for both outcomes in both groups in the Western trial[31] compared with the Asia–Pacific trial[32] (Table 4).

Table 4. Outcomes from the Western and Asia–Pacific trials of sorafenib for advanced hepatocellular carcinoma

| | Western (SHARP) [31] | | Asia–Pacific[32] | |
|---|---|---|---|---|
| | Sorafenib (N = 299) | Placebo (N = 303) | Sorafenib (N = 226) | Placebo (N = 76) |
| Median overall survival (months) | 10.7 | 7.9 | 6.5 | 2.8 |
| Hazard ratio (95% CI) | 0.69 (0.55 to 0.87) | | 0.68 (0.50 to 0.93) | |
| Median time to progression (months) | 5.5 | 2.8 | 2.8 | 1.4 |
| Hazard ratio (95% CI) | 0.58 (0.45 to 0.74) | | 0.57 (0.42 to 0.79) | |
| Abbreviations: CI, confidence interval; N, number. | | | | |

The CS also included a table of previous anticancer procedures and radiotherapy received by the lenvatinib and sorafenib groups before entry into REFLECT (CS Table 8). While it was not available separately for the Western region, the ERG's clinical experts considered the types of treatment and proportions of people who had received them generally reflective of what patients receive in the UK.

In summary, the ERG considers the population of REFLECT appropriate to provide evidence that is relevant to the decision problem. It was noted that REFLECT does not provide evidence for the use of lenvatinib for those with compromised liver function (Child–Pugh stage B or worse), or those with ECOG performance status of 2 or above. The ERG highlights baseline differences between the Western and Asia–Pacific subgroups of REFLECT and therefore considers it important to explore subgroup analyses by region for the clinical effectiveness outcomes.

## *3.2 Intervention*

Lenvatinib ([redacted], Eisai Co., Ltd) inhibits the activity of vascular endothelial growth factor (VEGF) receptors (VEGFR1 [FLT1], VEGFR2 [KDR], VEGFR3 [FLT4]) and other tyrosine kinase receptors implicated in angiogenesis, tumour growth, and cancer progression, including FGFR1, 2, 3, and 4, PDGFRα, KIT, and RET. The company's overview of the technology being appraised is reproduced in Table 5.

Table 5. Technology being appraised (reproduced from CS Table 2)

| | |
|---|---|
| UK approved name and brand name | UK approved name: Lenvatinib<br>Brand name: [redacted] |
| Mechanism of action | Lenvatinib is a RTK inhibitor that inhibits the activity of the VEGF receptors VEGFR1 (FLT1), VEGFR2 (KDR), and VEGFR3 (FLT4). Lenvatinib also inhibits other RTKs that have been implicated in angiogenesis, tumour growth, and cancer progression, including the FGF receptors FGFR1, 2, 3, and 4, PDGFRα, KIT, and RET. |
| Marketing authorisation/CE mark status | A regulatory submission was made to the EMA on 24th July 2017.<br>CHMP positive opinion is expected in [redacted] with marketing authorisation expected to be granted by the European Commission by [redacted] |
| Indications and any restriction(s) as described in the summary of product characteristics (SmPC) | The current draft indication† for lenvatinib is for the treatment of adult patients who have received no prior systemic therapy for HCC. |
| Method of administration and dosage | The recommended daily dose of lenvatinib is 8 mg (two 4 mg capsules) given orally QD for patients with a body weight of < 60 kg and 12 mg (three 4 mg capsules) orally QD for patients with a body weight of ≥ 60 kg. The daily dose is to be modified, as needed, according to the dose/toxicity management plan. |
| Additional tests or investigations | None. |
| List price and average cost of a course of treatment | £1,437.00 per pack of 30 x 4 mg capsules.<br>The average cost of a course of treatment (including PAS) is [redacted] |
| Patient access scheme (if applicable) | There is a simple PAS agreed with the Department of Health and the PAS price is incorporated in the submission. |
| Abbreviations: CHMP, Committee for Medicinal Products for Human Use; EMA, European Medicines Agency; FGF(R), fibroblast growth factor (receptor); HCC, hepatocellular carcinoma; PAS, Patient Access Scheme; PDGFRα, platelet derived growth factor receptor alpha; QD, once daily; RTK, receptor tyrosine kinase; SmPC, summary of medical product characteristics; VEGF(R), vascular endothelial growth factor (receptor).<br>[redacted];<br>†The current draft SmPC is presented in Appendix C. The population addressed in this submission and detailed in this table is based on anticipated changes to the licensed indication requested by the EMA rapporteur that have not yet been incorporated into the draft SmPC. | |

Lenvatinib currently has marketing authorisation as Kisplyx®, to be used in combination with everolimus, for advanced renal cell carcinoma previously treated with a VEGF inhibitor, and as Lenvima® for differentiated thyroid carcinoma. At the time of writing, lenvatinib has not received marketing authorisation for use in HCC but the company states that the draft indication is for the treatment of adult patients who have received no prior systemic therapy for HCC.

The intervention in REFLECT matches the NICE final scope, and is in line with the recommended dose and administration of lenvatinib, that is, once daily oral dose of 8 mg (two 4 mg capsules) for patients weighing less than 60 kg and 12 mg (three 4 mg capsules) for patients weighing 60 kg or greater. Given the weight group imbalance between the Western and Asia–Pacific subgroups, the ERG asked to see subgroup data for dose intensity data at the clarification stage. Mean and median dose of lenvatinib group were 10.2 and 10.1 in the Western subgroup compared with 9.4 and 8.9 in the full population, which the ERG did not consider sufficiently different to impact relative treatment effects. As outlined in Table 5, the daily dose in REFLECT could be interrupted and then reduced according to the dose/toxicity management plan and, once reduced, could not be increased. The CS states that patients discontinued treatment at the time of objectively documented disease progression (which the company

confirmed at the clarification stage was according to mRECIST assessed by the investigators), development of unacceptable toxicity, patient request, or withdrawal of consent.

After discontinuation of lenvatinib in REFLECT, patients received a variety of subsequent anti-cancer medications and procedures, a full breakdown of which was provided by the company at the clarification stage. The ERG notes that 25.5% of patients in the lenvatinib group received subsequent anti-cancer procedures [redacted], which are available to patients on the NHS, although a much smaller proportion of patients received any procedure in the Western subgroup (7.0%). Furthermore, around a third of patients (32.6%) of patients received anti-cancer medication after discontinuation of lenvatinib, [redacted]
The proportion of patients receiving subsequent medications) was [redacted] in the Western subgroup [redacted]. The ERG notes that regorafenib (TA514[33]) recently received a negative recommendation for previously treated advanced HCC and, while the NICE website lists several planned technology appraisals for the indication (ID1243, ID13114, ID1025), no drugs are currently approved for this use in the NHS.

The ERG considered the administration of lenvatinib to be in line with its anticipated use in UK clinical practice, but the extent and type of subsequent anti-cancer procedures and medication received in REFLECT do not reflect UK clinical practice.

## *3.3 Comparators*

The NICE final scope lists sorafenib and best supportive care (BSC) as comparators, and the CS only provides evidence for lenvatinib compared with sorafenib. The company state that BSC is not considered a relevant comparator for lenvatinib, "*due to the small numbers of patients in the population defined above…that would receive this as an alternative to sorafenib*". The ERG consulted clinical experts who agreed that sorafenib is the only relevant comparator for lenvatinib because the reasons for choosing BSC over systemic therapy would likely apply equally to sorafenib and lenvatinib; the population considered eligible for systemic therapy is restricted to those with Child–Pugh Class A liver function who would almost all be eligible for sorafenib or lenvatinib, and BSC would only be given if a patient chose not to receive systemic therapy.

Like lenvatinib, sorafenib (Nexavar®, Bayer) inhibits the activity of multiple tyrosine kinase receptors implicated in angiogenesis, tumour growth, and cancer progression. Sorafenib is the only systemic therapy recommended by NICE for the treatment of advanced HCC, and the ERG's clinical experts confirmed that it is the only drug available for patients with advanced HCC in the NHS. Sorafenib has been available through the Cancer Drugs Fund for patients with advanced HCC since 2007 (TA189),[34]

and has been available for routine commissioning since September 2017, following a CDF rapid reconsideration (TA474).[24]

Sorafenib was the comparator for lenvatinib in the REFLECT trial and was given at a starting dose of 400 mg orally, twice daily, in line with the recommended dose and administration described in its summary of medical product characteristics (SmPC). The ERG notes from the journal publication of REFLECT that the dose of sorafenib was modified according to regional prescribing information, and thus sought to compare dose intensity and time on treatment between the Western population and full population of the trial (see Table 6). While some differences were noted between the Western subgroup and full population , the ERG did not consider the extent of difference across treatment groups likely to impact relative treatment effects. As with lenvatinib, sorafenib was discontinued at objectively documented disease progression, development of unacceptable toxicity, patient request, or withdrawal of consent.

After discontinuation of sorafenib in REFLECT, patients received a variety of subsequent anti-cancer medications and procedures, a full breakdown of which was provided by the company at the clarification stage. As with the lenvatinib group, around a quarter of sorafenib-treated patients (27.3%) received subsequent anti-cancer procedures [redacted], which are available to patients on the NHS, and the proportion in the Western subgroup was much smaller (11.5%). Similarly, patients had access to a range of subsequent anti-cancer medications (38.7%) after discontinuation of sorafenib in REFLECT, most of which were [redacted]The proportion of patients receiving subsequent systemic medications or investigational agents was [redacted]in the Western subgroup [redacted]; the extent of subsequent treatment imbalance between groups and the potential impact on relative treatment effects is discussed in Section 4.2.1.

The ERG considers the administration and management of sorafenib in the full population to be a reasonable reflection of how it is used in clinical practice, and that it is the only relevant comparator for lenvatinib in the population of interest. However, as described above, the extent and type of subsequent anti-cancer procedures and medication received in REFLECT do not reflect UK clinical practice because there are currently no drugs approved for use after first-line systemic therapies for advanced HCC in the NHS.

## *3.4 Outcomes*

The company presents direct evidence for lenvatinib compared with sorafenib for all outcomes listed in the final scope issued by NICE, namely:

- Overall survival (OS);
- Progression-free survival (PFS);
- Time to progression (TTP);
- Response rates;
- Adverse effects of treatment;
- Health-related quality of life (HRQoL).

OS, the primary outcome of REFLECT, was defined as the time from randomisation to death from any cause. Patients were censored if they were alive at data cut-off or at the last date they were known to be alive for those who were lost-to-follow-up (CS, pg. 25), which the ERG considers appropriate. OS and all secondary efficacy outcomes were measured in all randomised participants.

PFS was defined as time from randomisation to the date of first documentation of disease progression, or death, whichever occurred first. TTP was defined similarly to PFS but deaths were censored in the TTP analysis. The ERG notes that PFS is sometimes considered unreliable in HCC research because the benefit of drugs in delaying cancer progression can be masked by deaths from the natural history of cirrhosis.[35] Consequently, given the general preference of regulatory and technology assessment agencies for PFS as a cancer endpoint, the ERG consider it appropriate to consider both outcomes in this STA.

Investigator-assessed PFS and TTP were based on 8-weekly tumour response evaluations according to modified Response Evaluation Criteria in Solid Tumours for HCC (mRECIST).[36] Standard RECIST has been challenged in HCC because effective treatments (e.g. sorafenib and TACE) often induce direct tumour necrosis without substantially changing overall tumour size. mRECIST, which was proposed by a panel of experts and is endorsed by the European Association for the Study of the Liver (EASL), measures tumour change by contrast enhancement in the arterial phase, and incorporates novel concepts to assess lymph node involvement, new lesions, and ascites.[36-38]

An alternative PFS analysis was presented in the CS based on retrospective assessments of 8-weekly scans by blinded independent imaging review (IIR) according to mRECIST and RECIST v1.1, but there was not a corresponding investigator assessment according to RECIST v1.1. The ERG's clinical experts considered mRECIST a valid measure, and were aware that radiologists tend to choose either mRECIST and RECIST for tumour assessments in UK centres. Where multiple assessments were available for comparison, the ERG has assessed variation in the results, but considers mRECIST a valid measure for

HCC. The ERG considers IIR-assessed PFS likely to be less biased that the investigator-assessed PFS due to the open-label design of REFLECT.

Censoring rules for PFS and TTP analyses were informed by guidance from the US Food and Drug Administration (FDA),[39] and included reasons that may be considered informative (e.g. not have progression at the time of treatment discontinuation, or starting a new anti-cancer treatment; Table 10). EMA guidance[40] states that informative censoring, that is for reasons potentially related with the actual survival time, may lead to incorrect conclusions about the extent of the treatment difference. The ERG was concerned that the censoring may bias results in favour of lenvatinib because discontinuation due to adverse events and patient choice were more frequent in the lenvatinib group than the sorafenib group, meaning more lenvatinib events could be missed by premature censoring. The company provided a supplementary analysis at the clarification stage where all events were counted in the analysis, and only patients with missing assessments or no progression were censored.

For response rates, results were presented using standard definitions of objective response rate (ORR; the number of people achieving a complete or partial response) and best overall response (BOR; of complete response, partial response, stable disease or progressive disease) based on investigator assessment using mRECIST. The ERG notes that the IIR PFS poster presentation also includes IIR results (mRECIST and RECIST) for TTP, ORR and BOR.[41]

HRQoL was assessed in REFLECT at baseline, 4-weekly during treatment, and at the end of treatment visit with the European Organisation for Research and Treatment generic cancer quality of life scale (EORTC QLQ-C30), the EORTC HCC-specific module (EORTC QLQ-HCC18), and the EuroQol 5 dimensions 3 levels measure of overall health status (EQ 5D-3L). The ERG considers the HRQoL measures sufficient to assess HRQoL, and to provide relevant evidence for use in the economic model.

AEs were assessed until 30 days after the last dose of study treatment in all patients who received at least one dose of the study drug, and were recorded regardless of whether the event was deemed by the investigator to be related to treatment. The ERG considered this appropriate given the open-label design of REFLECT, as the assessment of causal relationships may be subject to bias. All TEAEs were graded on a 5-point severity scale according to Common Terminology Criteria for Adverse Events (CTCAE).

The ERG considers the outcomes presented in the submission clinically relevant and in line with the decision problem outlined in the NICE final scope. Where there was a risk of investigator bias associated with the open-label design for PFS, TTP and response, results from alternative blinded analyses were available.

# 4 CLINICAL EFFECTIVENESS

## *4.1 Critique of the methods of review*

The company conducted a systematic literature review (SLR) to identify efficacy and safety evidence from randomised controlled trials (RCTs) of lenvatinib and sorafenib for adults with unresectable hepatocellular carcinoma (HCC) who had not previously received systemic treatment.

REFLECT provided appropriate evidence to inform estimates of clinical effectiveness and safety in the economic model for lenvatinib versus sorafenib, the only relevant comparator (Section 3.3). The company presents REFLECT as the only available RCT directly comparing lenvatinib with sorafenib and so deemed it unnecessary to incorporate evidence from additional RCTs. The Evidence Review Group (ERG) provides a brief critique of the SLR process, which was detailed in Appendix D of the company's submission (CS), to underscore the ERG's opinion that no relevant evidence was overlooked.

### 4.1.1 Searches

The company searched MEDLINE from 1950, EMBASE from 1947 and the Cochrane Library from inception, all via the OVID platform. All databases were searched up to 11 November 2017. Searches of the Cochrane Library included the Cochrane Database of Systematic Reviews (CDSR); Database of Abstracts of Reviews of Effects (DARE); Cochrane Central Register of Controlled Trials (CENTRAL); Cochrane Methodology Register (CMR); Health Technology Assessment Database (HTA) and NHS Economic Evaluation Database (NHSEED). The ERG notes that the omission of MEDLINE In-Process may mean newer records that have not yet been indexed may have been overlooked. The electronic database searches were supplemented by hand searches of clinical trial databases, conference proceedings, and grey literature (full list provided in CS Appendix D.1.1.5), for which terms, strategies and dates were not detailed.

The company provided search terms and strategies for each of the three electronic databases (CS Appendix D.1.1.4). All three strategies listed population/disease, intervention, comparator, and study design terms separately but the ERG notes the strategies provided did not show how the terms were combined. The ERG considered the population, intervention and comparator terms appropriate for the decision problem, and study design terms were appropriate to find RCTs. Due to time constraints, the ERG was unable to replicate the company's search and appraisal of identified abstracts.

The ERG is satisfied that the searches were sufficiently sensitive to identify the direct clinical effectiveness evidence to support the submission for lenvatinib, that is, the REFLECT study; the ERG's clinical experts were unaware of other relevant evidence that should have been incorporated. However, the ERG could not be certain that all recent publications of REFLECT data had been identified due to

the date of searches, omission of MEDLINE In-Process, and lack of detail in the reporting of the additional searches.

### 4.1.2 Inclusion criteria

The ERG considers the eligibility criteria for the clinical evidence SLR broadly appropriate to identify evidence relevant to the decision problem (Table 1, CS Appendix D.1.2), but relevant contextual evidence may have been overlooked by narrow eligibility criteria for the population and study design.

The SLR sought to identify RCTs of treatment naïve adults with advanced, unresectable, and/or metastatic HCC. While this closely matches the decision problem, pretreated population studies may have been useful to inform the modelling of subsequent treatments and procedures, particularly any that have been conducted in a UK population. Similarly, while the restriction to RCTs is justified to identify studies directly relevant to the decision problem, excluding other study designs (e.g. observational studies and retrospective cohorts) may have overlooked important contextual evidence to assess the generalisability of results from REFLECT to UK patients, and the long-term extrapolation of survival outcomes.

The ERG considers it reasonable that studies were excluded if they did not include at least one outcome of interest, that is, those listed in the final scope issued by the National Institute for Health and Care Excellence (NICE).[1] However, four studies of investigational therapies that otherwise met the inclusion criteria were excluded because the primary endpoint was not met, which the ERG does not consider an appropriate exclusion criteria (doxorubicin,[42] linifanib,[43] sunitinib,[44] brivanib[45]) details provided by the company at the clarification stage). Given that REFLECT provided mature data for the key outcomes and closely matched the decision problem, the ERG did not review the studies for additional evidence regarding sorafenib.

### 4.1.3 Critique of data extraction

The company presented a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram to illustrate the study selection process (CS Appendix D.2; reproduced in Figure 22, Appendix 10.3). The flow diagram details the number of records identified through the electronic databases, but does not give details of records retrieved and sifted through the additional searches. The ERG noted no discrepancies between the numbers reported in the figure and the lists of included and excluded records provided (CS Appendix D.1.2.1).

Sixteen records relating to 10 studies are listed as included (CS Appendix D.1.2.1, Table 2). Five records relate to the REFLECT study, which is the only study able to address the decision problem, and the role of the remaining 11 records (relating to nine studies) is not described. Table 51 (Appendix 10.1) provides a summary of all 10 studies included by the company, plus six additional studies highlighted

by the ERG, with a comment regarding their potential roles for this single technology appraisal (STA). REFLECT is the only study of lenvatinib, and the other 15 studies compared sorafenib with other interventions.[31, 42-55] The populations included in the 15 sorafenib studies vary in comparability to REFLECT with regards to Child–Pugh eligibility, but across them provide a useful evidence base for sorafenib with which to compare evidence from REFLECT.

In summary, the ERG is satisfied that no directly relevant studies were omitted from the analysis of clinical effectiveness evidence, but identified some potential oversights in the identification and incorporation of relevant contextual evidence. Nonetheless, the likely impact on the submission is low given the robustness of direct evidence from REFLECT and maturity of data for the key effectiveness outcomes.

### 4.1.4 Quality assessment

The company provided a quality assessment of REFLECT using criteria suggested in the NICE template for company submission of evidence to the STA process. The ERG's independent validation of the company's assessment is shown together with the company's assessment in Appendix 10.2 (Table 52). The ERG considers REFLECT to be of high methodological quality (low risk of selection, reporting and attrition biases) but highlights the following issues, which may introduce bias for some outcomes (see Table 52 for more details):

- the open-label design of REFLECT may cause detection bias for investigator-assessed progression-free survival (PFS), time to progression (TTP), objective response rate (ORR), and treatment-emergent adverse events (TEAEs), and performance bias for quality of life outcomes;

- censoring rules in the primary analyses of PFS and TTP may favour lenvatinib because more lenvatinib-treated patients discontinued treatment for reasons other than progression, which was a reason for censoring in these analyses. The ERG notes that the primary analyses were based on FDA rules for censoring[39] and were accepted by the European Medicines Agency (EMA), and results based on alternative rules for censoring were provided by the company at the clarification stage (see Section 4.2.3).

### 4.1.5 Evidence Synthesis

No evidence synthesis was required because clinical effectiveness and safety evidence was derived from a single head-to-head RCT.

### 4.1.6 Summary statement

The ERG is satisfied that the company's SLR process followed recommended methodological practices and was sufficiently robust to identify direct clinical effectiveness evidence. The ERG agrees that REFLECT is the only study that provides comparative evidence of lenvatinib versus the only relevant comparator, sorafenib. REFLECT studied the population of interest, and provides evidence for all outcomes listed in the NICE final scope, meaning the incorporation of evidence from additional RCTs was not necessary.

While the ERG identified potential oversights in the SLR searches and eligibility criteria with regards to potentially relevant contextual evidence (e.g. to validate results in the sorafenib arm, or inform extrapolation), the likely impact on the submission is low given the robustness of direct evidence from REFLECT and maturity of data for the key effectiveness outcomes.

With regards to study quality, REFLECT is an international, open-label phase III RCT, which the ERG judged to be generally low risk of bias (e.g. selection, reporting and attrition biases; Table 52). The open-label design may introduce detection bias for investigator-assessed PFS, TTP and ORR, and those based on independent imaging review (IIR) assessments are considered likely to be less biased. Censoring rules in the primary analyses of PFS and TTP may favour lenvatinib because more lenvatinib-treated patients discontinued treatment for reasons other than progression, but analyses using alternative rules were provided by the company for comparison at the clarification stage.

## *4.2 Critique of trials of the technology of interest, their analysis and interpretation*

Clinical effectiveness evidence presented in the CS is based solely on REFLECT, an international phase III, open-label RCT. The primary objective of REFLECT was to compare overall survival (OS) in patients treated with lenvatinib versus sorafenib for patients with untreated, unresectable HCC. Patients had to be classified as Barcelona Clinic Liver Cancer (BCLC) stage B or C and Child–Pugh Class A. The population, intervention, comparator and outcomes in REFLECT closely match the NICE final scope (see Section 3).

The company include a brief description of one other study, E7080-J081-202,[28] which supports the rationale for the dosing strategy used in REFLECT. E7080-J081-202 was a phase I/II dose finding study in Japan and South Korea (CS, pgs. 17–18), and did not provide comparative evidence for the economic model. As such, the ERG has not provided a summary of E7080-J081-202 and focuses solely on REFLECT in the following sections. The ERG did not identify any further clinical effectiveness studies it believes should have included in the submission for lenvatinib.

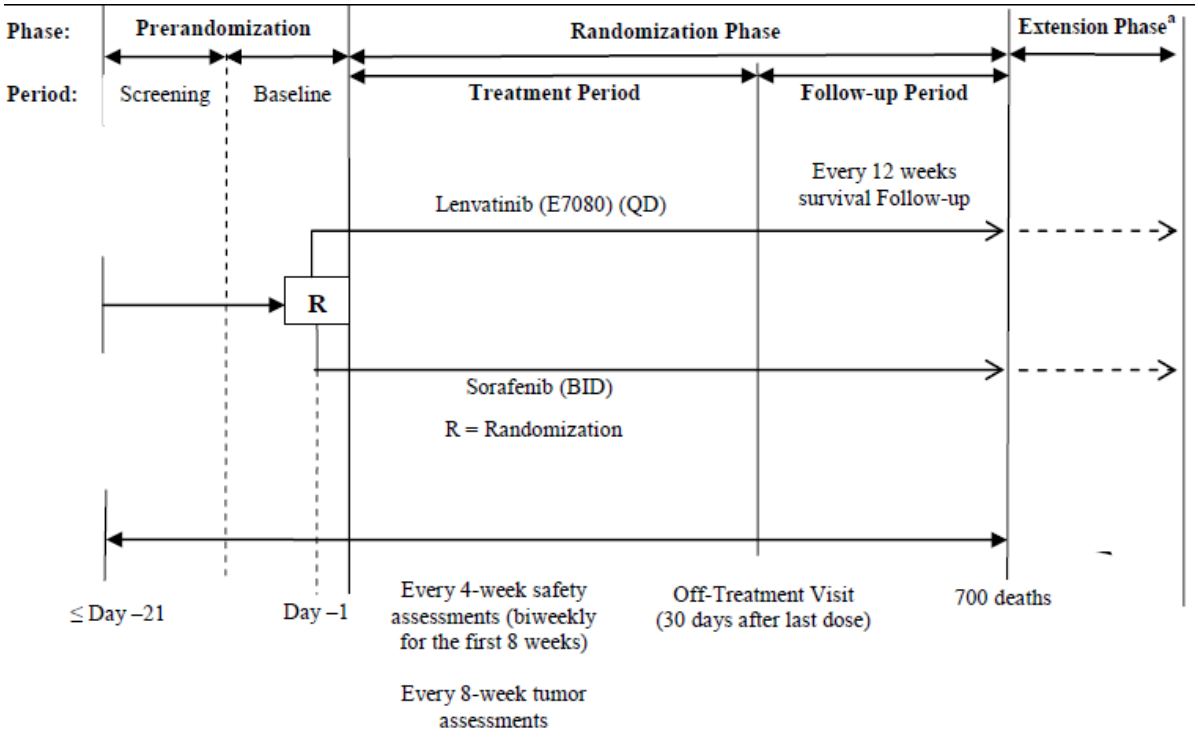
### 4.2.1 Trial conduct

REFLECT is an international, phase III, open-label RCT that recruited adults with previously untreated, histologically, cytologically or clinically confirmed (according to the American Association of the Study of Liver Diseases criteria) BCLC Stage B or C HCC. Patients had Child–Pugh class A liver function, good performance status (Eastern Cooperative Oncology Group [ECOG] 0 or 1), at least one measurable target lesion (according to a modified Response Evaluation Criteria in Solid Tumours [mRECIST]) and adequate liver, bone marrow, renal and pancreatic function. It was noted that REFLECT excluded patients with compromised liver function (Child–Pugh stage B or worse), and those with ECOG performance status of 2 or above, meaning the results of REFLECT may not be generalisable to those patients. Further eligibility criteria relating to survival expectation, liver occupation, and abnormal screening tests are listed in Table 5 of the CS (summarised in Section 3.1).

REFLECT recruited 954 patients at 183 sites across 30 countries in Asia, Europe and North America. Around two thirds of the population (67.1%) were recruited from countries in the Asia–Pacific region (China, Hong Kong, Japan, Republic of Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand) and a third (32.9%) from Western countries (Belgium, France, Germany, Italy, Israel, Poland, Russia, Spain, UK, Canada, USA). Six UK centres recruited 20 patients (2.1%). The ERG requested additional information at the clarification stage to assess the appropriateness of basing the STA on the full population of REFLECT compared with the Western subgroup, which is discussed throughout the following sections (e.g. baseline characteristics, treatment exposure, clinical effectiveness, adverse events and subsequent therapies).

The study was conducted in three phases, as shown in Figure 2. Patients entered the randomisation phase once eligibility was confirmed in the prerandomisation phase; the randomisation phase began when the first patient was randomised and ended at the primary data cut (13 November 2016). Patients remained in the treatment period of the randomisation phase until progression or discontinuation of the drug for any other reason. Patients then entered the follow-up period of the randomisation phase, in which patients were monitored every 12 weeks for OS and subsequent anti-cancer therapies, unless consent was withdrawn. Patients who had not progressed and remained on study treatment at the end of the randomisation phase could continue the same treatment in the extension phase (Figure 2).

Figure 2. Overview of REFLECT study design (reproduced from CS, Figure 2)




[a] Extension Phase also included a Treatment Period and Follow-up Period. All patients still on treatment at the end of the Randomisation Phase entered the Extension Phase and continued on the same study treatment they received in the Randomisation Phase. Abbreviations: QD once daily; BID, twice daily.

During the randomisation phase, study treatment was administered in 28-day cycles and 8-weekly tumour assessments were conducted by study investigators according to mRECIST or the start of another anti-cancer therapy. Retrospective IIR assessments of scans were also conducted according to mRECIST and RECIST v1.1, although there was no investigator assessment according to RECIST v.1. A brief description of mRECIST, and how it compares to RECIST version 1.1, is provided in Section 3.4; the ERG considers both to be validated measures of progression and tumour response in HCC.

Medications that patients were, and were not, permitted to receive for management of HCC or adverse effects during study treatment were the same for lenvatinib and sorafenib (CS, pg. 24). Study treatment continued until investigator-assessed disease progression, unacceptable toxicity, patient request, or withdrawal of consent, at which point an off-treatment visit occurred within 30 days of a patient's final dose. The CS stated that patients who discontinued study treatment for reasons other than disease progression were followed in the randomisation phase until disease progression or the start of another anti-cancer therapy, and then entered the extension phase for survival follow-up. Primary analyses of PFS and TTP censored patients if there was no disease progression at the time treatment was discontinued (see Section 4.2.3), but the company provided a supplementary analysis at the clarification stage including all observed events (see Section 4.3.2).

Randomisation to lenvatinib and sorafenib was in a 1:1 ratio based on a computer-generated scheme reviewed by an independent statistician. The allocation process was concealed by using an Interactive

Voice Response System. Stratification factors, which were prespecified in the statistical analysis plan (SAP), were geographic region (Asia–Pacific or Western), presence of macroscopic portal vein invasion or extrahepatic spread (or both), ECOG performance status (0 or 1), and body weight (<60 kg; ≥60 kg). The same factors were included as covariates in the primary analyses of all clinical outcomes.

Patients randomised to the lenvatinib group (N = 478) received 8 mg once daily if their baseline body weight was <60 kg, and 12 mg once daily if body weight was ≥60 kg, given as two or three 4 mg capsules, respectively. Patients randomised to the sorafenib group (N = 476) received 400 mg twice daily, given as two sets of two 200 mg tablets. Both treatments were orally self-administered and given in continuous 28-day cycles. The company explain that the study was open label because, in addition to the difference in formulation, dose modification for toxicity was different for lenvatinib and sorafenib. As such, the number of placebos required to permit dose reductions and maintain blinding would be unfeasible and could lead to dosing errors.

The Consolidated Standards of Reporting Trials (CONSORT) diagram provided in Appendix D.2 of the CS (and reproduced in Appendix 10.1) indicates that, at the primary data cut on 13 November 2016, 27 of 478 patients randomised to receive lenvatinib and 25 of 476 patients randomised to sorafenib were still taking study treatment. Within those who had discontinued study treatment (N = 451 patients in both groups), radiological progression was a more common reason for discontinuation in the sorafenib group (72.9%) than the lenvatinib group (65.1%), and adverse events and subject choice were more common in the lenvatinib group (13.2% and 5.9%, respectively) than the sorafenib group (9.0% and 3.2%, respectively). The number of patients discontinuing the study drug for other reasons (clinical progression, loss to follow-up, withdrawal of consent and other) were similar between groups.

Lenvatinib interruptions, dose reductions, or discontinuations for toxicities were predefined and provided in the CS (CS Appendix L.1), whereas adjustment of sorafenib dose was in accordance with local prescribing information. If the dose of either drug was reduced due to toxicity, it could not be increased. The ERG requested mean and median dose intensity, duration of treatment, and reduction and interruption data split by geographic region at the clarification stage, which is presented together with the full population data in Table 6. Information provided by the company indicates median time on treatment was somewhat longer in both groups in the Western subgroup than the safety analysis set (SAS; all patients who received at least one dose of the study drug). Treatment duration was longer (mean and median) in the lenvatinib group than the sorafenib group in the SAS and Western subgroup although the difference between groups was smaller in the latter. While some differences were noted between the Western subgroup and full population, the ERG does not consider the extent of differences likely to impact relative treatment effects, and that data for the full population are a reasonable reflection of the number of cycles of drugs that might be required in UK clinical practice.

Table 6. Summary of treatment exposure in REFLECT (safety analysis set and by region)

| | Lenvatinib | | Sorafenib | |
|---|---|---|---|---|
| | **Full population (SAS) N = 476** | **Western subgroup (N = 155)** | **Full population (SAS) N = 475** | **Western subgroup (N = 156)** |
| Months of treatment | | | | |
| Mean (SD) | 8.2 (7.04) | 8.0 (6.88) | 6.0 (6.47) | 6.8 (6.20) |
| Median | 5.7 | 6.2 | 3.7 | 4.6 |
| Q1, Q3 | 2.9, 11.1 | 2.9, 11.3 | 1.8, 7.4 | 2.1, 9.2 |
| Min, max | 0.0, 35.0 | 0.0, 32.8 | 0.1, 38.7 | 0.1, 32.8 |
| TEAE leading to drug interruption | 248 (52.1) | 68 (43.9) | 193 (40.6) | 50 (32.1) |
| TEAE leading to dose reduction | 184 (38.7) | 71 (45.8) | 185 (38.9) | 67 (42.9) |
| TEAE leading to drug withdrawal | 94 (19.7) | - | 69 (14.5) | - |
| Dose intensity (mg/day/patient) | | | | |
| Mean (SD) | 9.4 (5.71) | 10.2 (9.23) | 663.8 (173.15) | 669.1 (162.45) |
| Median | 8.9 | 10.1 | 771.4 | 750.5 |
| Q1, Q3 | 7.9, 12.0 | 8.0, 12.0 | 514.6, 800.0 | 541.8, 800.0 |
| Min, max | 1.7, 12.0* | 3.1, 12.0* | 126.3, 800.0 | 245.5,800.0 |

Abbreviations: CS, company submission; ERG, evidence review group; SAS, safety analysis set; SD, standard deviation; TEAE, treatment emergent adverse effect.
Maximum lenvatinib dose reported as 120.0 in CS Table 19 and Table 9 of the company's response to clarification – ERG has assumed this is a typographical error and has corrected to 12.0 in both cases. Data compiled from Table 19 and Table 20 of the CS, and Tables 6 and 9 of the company's response to clarification.

The ERG notes that median duration of treatment in the sorafenib arm reflects TTP and PFS (all 3.7 months), which would be anticipated given the stopping rule for each treatment (objectively documented disease progression or unacceptable toxicity). However, median time on treatment in the lenvatinib group (5.7 months) was shorter than median PFS and TTP (7.3 to 8.9 months, depending on the assessor and version of RECIST used; see Sections 4.3.2 and 4.3.3). The ERG surmises that the discrepancies may relate to imbalances in reasons for discontinuation and how patients were censored for PFS and TTP, which is discussed with additional information provided by the company at the clarification stage in Section 4.2.3.

Crossover to the alternative treatment at progression was not permitted in the protocol, but patients received various off-study subsequent therapies after discontinuation of the study drug. The ERG requested a detailed breakdown of post-study treatments at the clarification stage, for the full population and Western subgroup, which are presented in Table 7.

The number of patients receiving any subsequent anti-cancer procedure or medication was higher in the sorafenib group (51.1%) than the lenvatinib group (43.1%), and the imbalance was more pronounced in the Western subgroup (45.2% vs 28.0%, respectively). The company states that the imbalance is partly because patients in the sorafenib group were eligible for subsequent treatment in trials that did not permit prior use of lenvatinib, and that OS is likely to be biased as a result (CS, pg. 66). The company provided alternative OS analyses adjusting for subsequent therapies to explore the effect on OS, for the full population and separately for the geographical subgroups (see 4.3.1). As discussed in Section 3, the

extent and type of subsequent anti-cancer procedures and medication received in REFLECT do not reflect UK clinical practice because there are currently no drugs approved for use after first-line systemic therapies for advanced HCC in the National Health Service (NHS).

Table 7. Summary of subsequent procedures and medications (≥1% of patiets in either group; adapted from CS, Table 12 and company's response to clarification, Tables 2, 3, 4, 5)

| Subsequent interventions received during survival follow-up, n (%) | Lenvatinib | | Sorafenib | |
|---|---|---|---|---|
| | Full population | Western | Full population | Western |
| **Any anti-cancer procedure or medication** | 206 (43.1) | 44 (28.0) | 243 (51.1) | 71 (45.2) |
| **Any anti-cancer procedure** | 122 (25.1) | 11 (7.0) | 130 (27.3) | 18 (11.5) |
| High frequency ablation | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to bone | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to brain | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to liver | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to lymph nodes | [redacted] | [redacted] | [redacted] | [redacted] |
| Regional chemotherapy | [redacted] | [redacted] | [redacted] | [redacted] |
| Transcatheter arterial chemoembolisation | [redacted] | [redacted] | [redacted] | [redacted] |
| **Any anti-cancer medication** | 156 (32.6) | [redacted] (26.1) | 184 (38.7) | 61 (38.9) |
| Any antineoplastic and immunomodulating agent | [redacted] | [redacted] | [redacted] | [redacted] |
| Antimetabolites* | [redacted] | [redacted] | [redacted] | [redacted] |
| Cytotoxic antibiotics and related substances<br>Doxorubicin | [redacted] | [redacted] | [redacted] | [redacted] |
| Immunostimulants (interferon) | [redacted] | [redacted] | [redacted] | [redacted] |
| Other neoplastic agents<br>Cabozantinib<br>Cisplatin<br>Ipilimumab<br>Nivolumab<br>Oxaliplatin<br>Protein kinase inhibitors<br>Regorafenib<br>Sorafenib<br>Sunitinib<br>Tivantinib | [redacted] | [redacted] | [redacted] | [redacted] |
| Various<br>Folinic acid<br>Investigational drugs | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: n, number of patients.
*Antimetabolites received by more than 1% of patients in either group were capecitabine, flurouracil, gemcitabine and uftoral.



Outcomes measured in REFLECT were considered clinically relevant by the ERG's clinical experts, and in line with the decision problem outlined in the NICE final scope[1] (Section 3.4). Full details of the definitions used, reasons for censoring, and methods of analysis are provided in Table 9, but briefly:

- OS, the primary outcome of REFLECT, was defined as the time from randomisation to death from any cause. After discontinuation of the study drug, patients were followed-up every 12 weeks for survival and anti-cancer therapies;

- Investigator-assessed PFS, based on 8-weekly tumour assessments using mRECIST, was defined as time from randomisation to the date of first documentation of disease progression, or death, whichever occurred first;

- TTP was defined similarly to PFS, but deaths were censored;

- ORR was defined as the number of people achieving a complete or partial response based on investigator assessment by mRECIST, and best overall response (BOR) as the best of complete response, partial response, stable disease or progressive disease during the study;

- PFS, TTP, ORR and BOR were also assessed by a retrospective IIR, according to mRECIST and RECIST v1.1;

- Health-related quality of life (HRQoL) was assessed in REFLECT at baseline, 4-weekly during treatment, and at the end of treatment visit with the following questionnaires:

    - European Organisation for Research and Treatment generic cancer quality of life scale (EORTC QLQ-C30);

    - EORTC HCC-specific module (EORTC QLQ-HCC18);

    - EuroQol 5 dimensions 3 levels measure of overall health status (EQ 5D-3L);

- TEAEs were assessed until 30 days after the last dose of study treatment in all patients who received at least one dose of the study drug, and were recorded regardless of whether the event was deemed by the investigator to be causally related to treatment. All TEAEs were graded on a 5-point severity scale according to Common Terminology Criteria for Adverse Events (CTCAE).

In summary, REFLECT is a large, international, phase III, open-label RCT that recruited 954 adults with previously untreated HCC. The study was designed to show non-inferiority of lenvatinib compared with sorafenib for OS, and key secondary outcomes matched the NICE final scope (PFS, TTP, ORR, HRQoL and TEAEs). Around two-thirds of the population were from the Asia–Pacific region (N = 640), and a third from the Western region (N = 314), of which 20 patients were from the UK. The ERG's clinical experts consider the REFLECT eligibility criteria generally reflective of a UK population with untreated HCC who would be eligible for lenvatinib should it be approved. However,

it was noted that results may not generalise to those with compromised liver function (Child–Pugh stage B or worse) or ECOG performance status of 2 or above.

The ERG considers REFLECT to be a well-conducted study, which provides robust and mature data for key effectiveness outcomes. Investigator-assessed outcomes are likely to be at high risk of bias, but alternative analyses based on IIR assessments are available for PFS, TTP and ORR. Where the ERG was uncertain about trial conduct or considered there a risk of bias, for example with rules for censoring for PFS and TTP, the company provided adequate explanation and supplementary analyses at the clarification stage for the ERG to assess the potential impact on relative treatment effects.

Data about the Western subgroup were available for comparison against data on the full population to assess which is likely to provide the most appropriate evidence base. Some differences in the administration and management of drugs were noted between the full population and Western subgroup (e.g. in time on treatment and dose intensity in the lenvatinib group), which were considered unlikely to impact relative treatment effects. Patients in the full population and Western subgroup of REFLECT received a variety of subsequent therapies that do not reflect UK clinical practice, which requires consideration in the interpretation of OS.

### 4.2.2 Baseline characteristics

As previously described (Section 3.1), the ERG's clinical experts considered baseline and disease characteristics in REFLECT generally reflective of patients in England with untreated HCC, and mostly well balanced between groups in the full population. However, differences were noted between geographical subgroups (Western and Asia–Pacific) within each randomised treatment group. Baseline characteristics tables for the full population (Table 53 and Table 54) and Western and Asia–Pacific subgroups (Table 55) are provided in Appendix 10.4. Table 8 gives an overview of both, focussing on characteristics highlighted by the company or the ERG's clinical experts as being prognostic indicators.

Baseline characteristics generally were similar for the lenvatinib and sorafenib groups in the full population of REFLECT, and the ERG's clinical experts considered the population generally reflective of a UK population who would be eligible for sorafenib or lenvatinib should it be approved, despite only 20 patients being recruited from UK centres (2.1%). Groups were balanced for mean age (61 years), sex (~85% male), proportion with macroscopic portal vein invasion or extrahepatic spread (~70%), BCLC stage (~20% stage B who were ineligible for TACE, and 80% stage C), ECOG performance status (63%), and Child–Pugh score (99% 5 or 6, i.e. class A, see Section 2). Around half of the patients in each group were recorded as having underlying cirrhosis, although the company highlights this is likely to be an underestimate as cirrhosis was only verified when needed to confirm the clinical diagnosis of HCC. A small number of patients in both groups had Child–Pugh class B liver function (1%), thus falling outside the study's eligibility criteria.

Table 8. Abbreviated demographic and disease baseline characteristics of the REFLECT population – full population and by region

| | | Lenvatinib | | | Sorafenib | | |
|---|---|---|---|---|---|---|---|
| | | **Full population N = 478** | **Asia–Pacific N = 321** | **Western N = 157** | **Full population N = 476** | **Asia–Pacific N = 319** | **Western N = 157** |
| Mean age (SD), years | | 61.3 (11.69) | 60.0 (11.76) | 63.8 (11.15) | 61.2 (12.01) | 60.2 (11.87) | 63.3 (12.06) |
| Male n (%) | | 405 (84.7) | 277 (86.3) | 128 (81.5) | 401 (84.2) | 269 (84.3) | 132 (84.1) |
| Body weight <60 kg (i.e. lower 8 mg dose) | | 153 (32.0) | 132 (41.1) | 21 (13.4) | 146 (30.7) | 127 (39.8) | 19 (12.1) |
| ECOG PS, n (%) | 0 | 304 (63.6) | 206 (64.2) | 98 (62.4) | 301 (63.2) | 204 (63.9) | 97 (61.8) |
| | 1 | 174 (36.4) | 115 (35.8) | 59 (37.6) | 175 (36.8) | 115 (36.1) | 60 (38.2) |
| Macroscopic portal vein invasion, extrahepatic spread, or both, n (%) | | 329 (68.8) | 220 (68.5) | 109 (69.4) | 336 (70.6) | 221 (69.3) | 115 (73.2) |
| Underlying cirrhosis[†], n (%) | | 243 (50.8) | 180 (56.1) | 63 (40.1) | 231 (48.5) | 169 (53.0) | 62 (39.5) |
| BCLC stage B (intermediate disease), n (%) | | 104 (21.8) | 70 (21.8) | 34 (21.7) | 92 (19.3) | 65 (20.4) | 27 (17.2) |
| Number of involved sites per patient, n (%) | 1 | 207 (43.3) | 144 (44.9) | 63 (40.1) | 207 (43.5) | 145 (45.5) | 62 (39.5) |
| | 2 | 167 (34.9) | 110 (34.3) | 57 (36.3) | 183 (38.4) | 112 (35.1) | 71 (45.2) |
| | ≥3 | 103 (21.5) | 67 (20.9) | 36 (22.9) | 86 (18.1) | 62 (19.4) | 24 (15.3) |
| Factor of carcinogenesis, n (%) | Hepatitis B | 251 (52.5) | 212 (66.0) | 39 (24.8) | 228 (47.9) | 197 (61.8) | 31 (19.7) |
| | Hepatitis C | 91 (19.0) | 50 (15.6) | 41 (26.1) | 126 (26.5) | 70 (21.9) | 56 (35.7) |
| | Alcohol | 36 (7.5) | 17 (5.3) | 19 (12.1) | 21 (4.4) | 8 (2.5) | 13 (8.3) |
| | Other or unknown | 100 (20.9) | 42 (13.1) | 58 (36.9) | 101 (21.2) | 44 (13.8) | 57 (36.3) |
| Baseline serum AFP ≥200 ng/mL, n (%) | | 222 (46.4) | 157 (48.9) | 65 (41.4) | 187 (39.3) | 137 (42.9) | 50 (31.8) |

Abbreviations: AFP, alpha-fetoprotein; BCLC, Barcelona Clinic Liver Centre; ECOG PS, Eastern Cooperative Oncology Group performance status; SD, standard deviation.

The company highlights two imbalances between the lenvatinib and sorafenib groups in the full population that might affect the relative treatment effect: proportion of patients with alpha-fetoprotein (AFP) levels ≥200 ng/ml and aetiology of HCC (hepatitis B, hepatitis C, or alcohol). Neither variable was listed as a randomisation stratification factor in the SAP, so the company performed covariate analyses to evaluate whether the imbalances impacted OS in the full population (Appendix L.2; discussed in Section 4.2.3 and with results for OS in Section 4.3.1).

The ERG's clinical experts agreed that elevated AFP is a prognostic indicator, but did not consider there to be a clinical rationale for dichotomising at 200 ng/ml. As such, while a higher proportion of patients in the lenvatinib arm had AFP ≥200 ng/ml (46.4%) than the sorafenib arm (39.3%), this may not be a clinically meaningful distinction (Table 8).

Hepatitis B was the most common factor of carcinogenesis in both groups in the full population (52.5% and 47.9% of the lenvatinib and sorafenib groups, respectively), followed by hepatitis C (19.0% and 26.5%, respectively). There is some evidence the higher proportion of patients with hepatitis C aetiology in the sorafenib arm could bias results because patients may derive more benefit from sorafenib than those with other aetiologies,[29] which has not been shown for other TKIs.[30]

The ERG and its clinical experts noted some key differences between the Asia–Pacific and Western geographical regions. Compared with the full population, fewer patients in the Western subgroup fell into the lower weight group (~13% <60 kg compared with ~30% in the full population), more had heart disease (~23% New York Heart Association [NYHA] classification I or II compared with ~10%), fewer had underlying cirrhosis (40% compared with 50%), and more Western patients had hepatitis C, alcohol, or other aetiologies as underlying causes of carcinogenesis (Table 8). The ERG also notes that the Western subgroup was older, had worse Child–Pugh scores, higher baseline ammonia levels, and shorter time since first diagnosis (Table 55).

While differences are noted between the full population of REFLECT and the Western subgroup, evidence from separate Western and Asia–Pacific trials for sorafenib in HCC (SHARP[31] and the Asia–Pacific study[32]) suggest that expected baseline differences based on geographical location but do not necessarily impact relative treatment effects (Table 4). The ERG does not consider there to be sufficient evidence to justify the inevitable loss of precision on effect estimates by focusing on the Western subgroup rather than the full population of REFLECT, in terms of comparability to a UK population. Where possible, the ERG provides a comparison of clinical effectiveness results from the Western subgroup and full population.

### 4.2.3 Description and critique of statistical approach used

The company presents a summary of statistical analyses employed in REFLECT in Table 9 of the CS. The ERG presents the information in Table 9 with additional information about the definitions, time points and population used for each outcome.

Table 9. Summary of REFLECT efficacy outcomes included in the CS

| Outcome | Definition | Measurements | Analysis |
|---|---|---|---|
| **OS (primary outcome)** | Time from randomisation to date of death (any cause). Assessed 12-weekly. | NA | Non-inferiority based on 2-sided 95% CI of HR estimated with stratified Cox proportional hazards model (margin 1.08). Superiority tested with stratified log-rank (declared if 2-sided p-value 0.05).<br>Primary analysis at 700 events. |
| **PFS** | Time from randomisation to date of first-documented PD or death (any cause). | 8-weekly tumour assessments by:<br>1) Investigator with mRECIST<br>2) IIR with mRECIST<br>3) IIR with RECIST V1.1 | Stratified log-rank with 2-sided α=0.05. HR from Cox proportional hazards model with 2-sided 95% CI, median PFS, and cumulative probability of PFS at 6-month intervals. |
| **TTP** | Time from date of randomisation to first documented PD. Deaths censored. | | |
| **ORR** | Proportion of patients with BOR of complete response partial response. | | Cochran-Mantel-Haenszel chi-square test with stratification factors as strata, tested at 2-sided α=0.05 to produce an odds ratio and 95% CI. |
| **HRQoL** | Completed at baseline, day 1 of each 28-day treatment cycle, and at the off-treatment visit. | Generic: EQ-5D-3L<br>Generic cancer: EORTC QLQ-C30<br>HCC-specific: EORTC QLQ-HCC18 | Relative effects investigated cross-sectionally and longitudinally. |
| **Adverse effects** | Any untoward medical occurrence regardless causal relationship to the study drug. SAEs defined as life-threatening, requiring hospitalisation, or leading to persistent or significant disability, congenital defect, or death. | CTCAE version 4.0 by grades 1 to 5<br>Recorded for 30 days after last dose of study treatment. | Incidence (number of patients with at least one) and episodes per patient-year to account for treatment exposure. Relative risk and risk difference with 95% CI presented in CS Appendix L.5. |

Abbreviations: BOR, best overall response; CI, confidence interval; IIR, independent imaging review; HR, hazard ratio; HRQoL, health-related quality of life; NA, not applicable; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; TTP, time to progression.

The sample size of REFLECT was determined by the number of events required to detect the non-inferiority and superiority of lenvatinib to sorafenib in OS, which was based on the following:

- Expected median OS of sorafenib of approximately 10 months;
- Target hazard ratio (HR) of 0.80 and a lenvatinib OS improvement of 2.5 months, yielding 97% power to declare non-inferiority at an upper 95% confidence interval (CI) margin of 1.08 and 82% power to declare superiority;

- Overall false-positive rate with a 2-sided alpha of 0.05;
- 5% of patients would have major protocol deviations and would be included in the FAS but not the per protocol set;
- Two interim futility analyses at approximately 30% and 70% of the target events.

The above assumptions led to a calculation of approximately 700 required deaths at the time of primary analysis, and a randomised sample size of approximately 940 patients. The ERG notes that median OS for sorafenib in the REFELCT primary analysis was longer than expected (12.3 vs 10 months), and the target HR was not met in the primary or supportive analyses of OS, but considers the assumptions generally appropriate.

For all efficacy outcomes, the primary analysis used the FAS, that is, all randomised patients. Secondary per protocol analyses were conducted including data only for patients who had received at least one dose of study medication and had no major protocol deviations. TEAE analyses used the safety population consisting of all patients who had received at least one dose of study medication, which was all but two patients in the lenvatinib group and one patient in the sorafenib group. The ERG considers the populations for analysis appropriate.

The company provides a summary of statistical methods for each outcome in CS, Table 9, which outlines a fixed sequence of analyses for the efficacy endpoints. Two interim futility analyses (at 30% and 70% of events) and the final analysis of OS were based on non-inferiority using a 2-sided 95% CI of the HR estimated using a Cox proportional hazards (PH) model and including treatment group as a factor. Non-inferiority was declared if the upper limit of the CI was less than 1.08, at which point superiority hypotheses were tested for OS with stratified log-rank tests at an alpha level of 0.05 (2-sided). Differences between groups for PFS and TTP were also tested using stratified log-rank tests at an alpha level of 0.05 (2-sided). ORR between groups was conducted with the Cochran–Mantel–Haenszel chi-square test at a 2-sided alpha level of 0.05, to produce an odds ratio (OR) and 2-sided 95% CIs.

The following stratification factors were prespecified for randomisation and were included as covariates for all stratified log-rank tests and stratified Cox proportional hazard models for the primary endpoint (OS) and key secondary endpoints (PFS, TTP and ORR):

- region (Asia–Pacific; Western);
- macroscopic portal vein invasion and/or extrahepatic spread (yes; no);
- ECOG performance status (0; 1);

- body weight (<60 kg; ≥60 kg).

AFP levels (<200 ng/mL, ≥200 ng/mL) and HCC aetiology (hepatitis B, hepatitis C, or alcohol) are also explored as covariates in the model, which was allowed in the SAP. After consulting clinical experts, the ERG considered the stratification factors and supportive analyses appropriate.

The company presents formal assessments of PH for OS and PFS, based on visual inspection of log-cumulative hazard plots and PH global tests (Schoenfeld residual test). For PFS, there was statistical deviation from the PH assumption based on the PH global test, and the cumulative hazards plot indicated a change in the trend of the hazard. As such, analyses for the economic model were based on independent statistical models for each treatment arm. For OS, while the company found no statistical deviation from the PH assumption, independent statistical models were used because the cumulative hazard plots were neither entirely straight nor parallel, and for consistency with the PFS analysis (see CS, Section B.3.3.1). As such, the ERG interprets HRs with caution for PFS and OS. The PH assumption was not tested for TTP but, given the similarity of the outcome definition and results to PFS, the ERG assumes it does not hold for TTP.

Censoring rules for PFS and TTP analyses were informed by guidance from the FDA,[39] and included reasons that may be considered informative (e.g. not having progression at the time of treatment discontinuation, or starting a new anti-cancer treatment; Table 10). EMA guidance[40] states that informative censoring, that is, for reasons potentially related with the actual survival time, may lead to incorrect conclusions about the extent of the treatment difference. The ERG was concerned that the censoring may bias results in favour of lenvatinib because discontinuation due to TEAEs and patient choice were more frequent in the lenvatinib group than the sorafenib group (Appendix 10.1, Figure 4, and, Section 4.2.1), meaning more lenvatinib events could be missed by premature censoring. At the clarification stage, the company provided a supplementary analysis where all events were counted in the analysis, and only patients with missing assessments or no progression were censored (Table 10); results of the supplementary analysis of PFS are presented alongside the primary analysis in Section 4.3.2.

Table 10. Reasons for censoring in the primary PFS and TTP analyses

| | Lenvatinib N = 478 | Sorafenib N = 476 |
|---|---|---|
| **Progression-free survival primary analysis (investigator-assessed)** | | |
| Patients with events, n (%) | [redacted] | [redacted] |
| Censored patients, n (%) | [redacted] | [redacted] |
| No post baseline tumour assessment | [redacted] | [redacted] |
| Death or progression after missing assessment | [redacted] | [redacted] |
| New anti-cancer treatment started | [redacted] | [redacted] |
| No progression at the time of data cut-off | [redacted] | [redacted] |
| No progression at the time of treatment discontinuation | [redacted] | [redacted] |
| **Time to progression primary analysis (investigator-assessed)** | | |
| Patients with events, n (%) | [redacted] | [redacted] |
| Censored patients, n (%) | [redacted] | [redacted] |
| No post baseline tumour assessment* | [redacted] | [redacted] |
| Death or progression after missing assessment | [redacted] | [redacted] |
| New anti-cancer treatment started | [redacted] | [redacted] |
| No progression at the time of data cut-off | [redacted] | [redacted] |
| No progression at the time of treatment discontinuation | [redacted] | [redacted] |
| **Supplementary analysis of investigator-assessed PFS provided at the clarification stage** | | |
| Patients with events, n (%) | [redacted] | [redacted] |
| Censored patients, n (%) | [redacted] | [redacted] |
| No baseline tumour assessment | [redacted] | [redacted] |
| No post-baseline tumour assessment | [redacted] | [redacted] |
| No progression at the time of data cut-off | [redacted] | [redacted] |
| No progression at the time of consent withdrawal | [redacted] | [redacted] |

Abbreviations: n, number of patients; PFS, progression-free survival; TTP, time to progression.

Details of the statistical analyses conducted for the HRQoL outcomes were not provided in CS Table 9, and the CSR only stated that relative treatment effects were investigated in cross-sectional and longitudinal analyses based on the 4-weekly assessments. The CSR references a separate report containing full details of the patients reported outcomes in REFLECT which was not made available to the ERG.

TEAEs were presented as the number of patients experiencing each event, and as the number of episodes per patient year to control for differences in treatment exposure. The ERG notes the imbalance in time on treatment in the discussion of TEAEs, but considers it appropriate to capture the burden of adverse events over the course of treatment received, that is, from which effectiveness was derived. Furthermore, adjustment for exposure does not account for the tendency of TEAEs to occur in the early stages of treatment; where reported in the draft SmPC, [redacted]. Relative risk and risk difference between treatments with 95% CI for each TEAE were presented in CS (Appendix L.5) to assess the statistical significance of differences between treatment groups.

Subgroup analyses were planned for the primary outcome, OS, and the secondary outcomes PFS, TTP, and ORR. Subgroup analyses were based on randomisation stratification factors, age group (≤65, ≥65 to <75 years and ≥75 years), sex and aetiology (hepatitis B, hepatitis C, and alcohol), and results were presented for each stratum as a HR and 95% CI. While not prespecified, additional subgroup analyses were also conducted for AFP at baseline (<200 ng/mL, ≥200 ng/mL), BCLC stage (B, C), and subsequent anti-cancer therapy (yes, no), which was allowed in the SAP. The ERG considers the planned and additional subgroups appropriate and discusses the results for each in Section 4.3.6.

The ERG considers the company's statistical approach appropriate and well described. Where *post-hoc* analyses were conducted to test the robustness of treatment effects, these were justified fully with reference to the SAP. In cases where the ERG was uncertain about methods or considered there a risk of bias, for example with rules for censoring for PFS and TTP, the company provided adequate explanation and supplementary analyses at the clarification stage for the ERG to assess the potential impact on relative treatment effects.

### 4.2.4 Summary statement

The company's SLR identified the REFLECT study as the only RCT that provides evidence directly relevant to the decision problem. The ERG is satisfied that the company's SLR process was sufficiently robust to confirm that no other RCTs were eligible; REFLECT studied the population, intervention and comparator of interest, and provides evidence for all outcomes listed in the NICE final scope, meaning the incorporation of evidence from additional RCTs was not necessary. The ERG identified potential oversights in the SLR searches and eligibility criteria with regards to potentially relevant contextual evidence (e.g. to validate results in the sorafenib arm, or inform extrapolation), but the likely impact on the submission is low given the robustness of direct evidence from REFLECT and maturity of data for the key effectiveness outcomes.

REFLECT is a large, international, phase III, open-label RCT that recruited 954 adults with previously untreated HCC. The study was designed to show non-inferiority of lenvatinib compared with sorafenib for OS, and key secondary outcomes matched the NICE final scope (PFS, TTP, ORR, HRQoL and TEAEs). Around two-thirds of the population were from the Asia–Pacific region (N = 640), and a third from the Western region (N = 314), of which 20 patients were from the UK. The ERG's clinical experts consider REFLECT to provide evidence that is relevant to the decision problem. However, it was noted that results may not generalise to those with compromised liver function (Child–Pugh stage B or worse) or ECOG performance status of 2 or above.

The ERG considers REFLECT to be a well-conducted study, which provides robust and mature data for key effectiveness outcomes. Investigator-assessed outcomes are likely to be at high risk of bias, but alternative analyses based on IIR assessments are available for PFS, TTP and ORR. Where the ERG

was uncertain about trial conduct or considered there a risk of bias, for example with rules for censoring for PFS and TTP, the company provided adequate explanation and supplementary analyses at the clarification stage for the ERG to assess the potential impact on relative treatment effects.

The ERG requested additional information at the clarification stage to assess whether the Western subgroup or full population forms the most appropriate evidence base for this STA. Some differences between the full population and Western subgroup were noted, particularly in terms of baseline characteristics, but evidence from separate Western and Asia–Pacific trials for sorafenib in HCC (SHARP[31] and the Asia–Pacific study[32]) suggest that this does not necessarily impact relative treatment effects (Table 4). Neither the full population nor the Western subgroup were considered reflective of UK clinical practice with regards to the extend and type of subsequent treatments received, and there was more imbalance in the Western subgroup. Overall, the ERG does not consider there to be sufficient evidence to justify the inevitable loss of precision by focusing on the Western subgroup.

The ERG considers the company's statistical approach appropriate and well described and, where post-hoc analyses were conducted to test the robustness of treatment effects, these were justified fully with reference to the SAP. In cases where the ERG was uncertain about methods or considered there a risk of bias, for example with rules for censoring for PFS and TTP, the company provided adequate explanation and supplementary analyses at the clarification stage for the ERG to assess the potential impact on relative treatment effects.



## *4.3 Clinical effectiveness results*

### 4.3.1 Overall survival

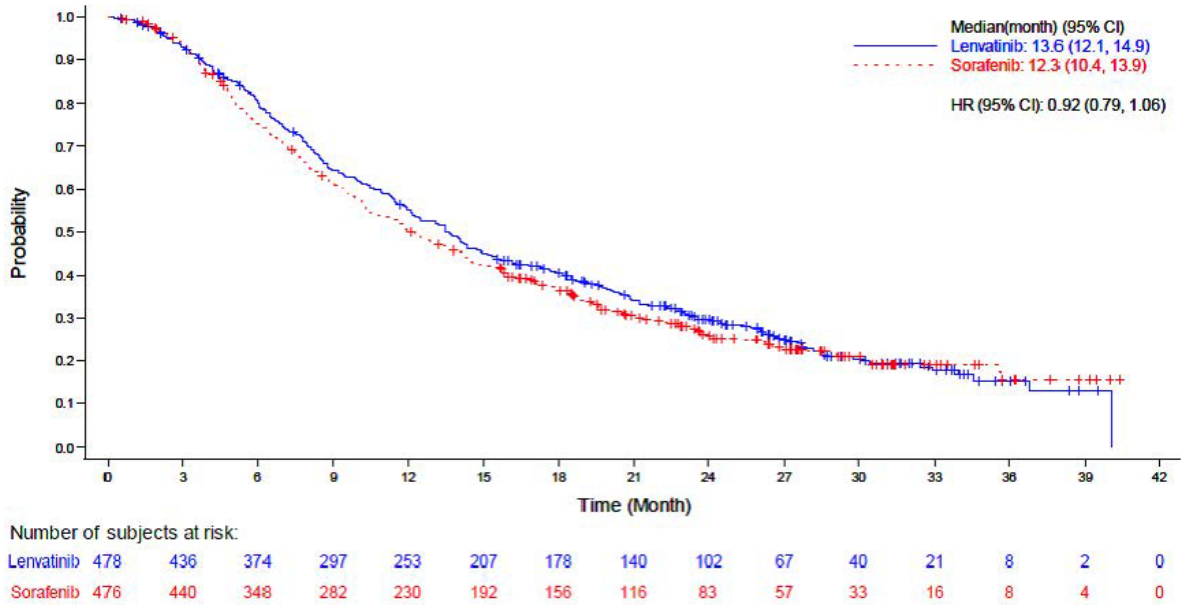
At the primary data cut-off (13 November 2016), 351 patients in the lenvatinib group and 350 (73.4%) in the sorafenib group (73.5%) had died, at which point median survival follow-up was 27.7 months in the lenvatinib group and 27.2 months in the sorafenib group (Table 11). In the primary analysis, median OS was 13.6 months for patients treated with lenvatinib and 12.3 months for patients treated with sorafenib; the HR for lenvatinib versus sorafenib is 0.92 (95% CI: 0.79 to 1.06; Figure 3). The upper 95% CI falls below the predefined limit to declare non-inferiority (Section 4.2.3), but criteria for superiority were not met. While the company found no statistical deviation from the PH assumption for OS, the cumulative hazard plots were neither entirely straight nor parallel, so the ERG interprets the HRs with caution.

Table 11. Overall survival in REFLECT (primary and supportive analyses; adapted from CS, Table 11)

| | Lenvatinib N = 478 | Sorafenib N = 476 |
|---|---|---|
| Deaths, n (%) | 351 (73.4) | 350 (73.5) |
| Censored patients, n (%) | 127 (26.6) | 126 (26.5) |
| Loss to follow-up | 5 (1.0) | 11 (2.3) |
| Withdrawal of consent | 13 (2.7) | 8 (1.7) |
| Alive | 109 (22.8) | 107 (22.5) |
| Median OS (95% CI), months | 13.6 (12.1, 14.9) | 12.3 (10.4, 13.9) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.92 (0.79 to 1.06) | |
| Per protocol population* | 0.91 (0.78 to 1.06) | |
| AFP added as a covariate | 0.86 (0.74 to 1.00)** | |
| HCC aetiology added as a covariate | 0.86 (0.72 to 1.01) | |
| OS rate, % (95% CI) 6 months | [redacted] | [redacted] |
| 12 months | | |
| 18 months | | |
| 14 months | | |
| Median survival follow-up (95% CI), months | 27.7 (26.4 to 29.4) | 27.2 (25.9 to 28.4) |

Abbreviations: AFP, alpha fetoprotein; CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; HCC, hepatocellular carcinoma; OS, overall survival.
†OS rate and 95% CI were calculated using Kaplan–Meier product-limit method and Greenwood Formula; ‡Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor. Efron method was used for ties. §Stratified by region (Region 1: Asia–Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg).
* The per protocol population consisted of all patients who received at least one dose of study medication and had no major protocol violations population, randomisation stratification factors.
** Results for the supportive analyses were reported with 3 decimal places in the submission and the difference between lenvatinib and sorafenib was statistically significant when adjusted for the baseline imbalance in AFP (p = 0.0342).

Figure 3. Kaplan–Meier curve for overall survival (reproduced from CS, Figure 3)




Results from prespecified supportive analyses are also shown in Table 11. OS within the per protocol population was consistent with the FAS (HR 0.91, 95% CI: 0.78 to 1.06; CS, pg. 38), and covariate analyses to adjust for individual baseline factors that were not prespecified for the primary analysis did

not substantially impact the point estimate or conclusions (CS Appendix L.2). The HR moved in favour of lenvatinib when OS was adjusted for the baseline imbalance in AFP levels (HR 0.856, 95% CI: 0.736 to 0.995, p =0.0342), and for the baseline imbalance in HCC aetiologies (HR 0.855, 95% CI: 0.721 to 1.013), but the difference between groups was only statistically significant when the AFP adjustment was included.

A breakdown of subsequent therapies was provided by the company at the clarification stage for the full population and Western subgroup (see Section 4.2.1). Patients in both groups received a variety of subsequent anti-cancer treatments, whereas the ERG's clinical experts advised that patients who receive sorafenib for advanced HCC in UK clinical practice do not receive subsequent systemic therapy. Equally, subsequent treatment would not be available for patients should lenvatinib be approved for use on the NHS. As such, results for the full population and Western subgroup are likely to overestimate absolute OS in both treatment groups. Moreover, more patients in the sorafenib group received anti-cancer therapy during the survival follow-up, which may bias OS in favour of sorafenib (Table 7). The company conducted *post-hoc* analyses to adjust for subsequent anti-cancer therapy in the overall population, which [redacted] the HR for the full population and accompanying 95% CI [redacted]. However, the ERG notes that the adjustment is based only on the categorisation of patients into those who received any subsequent anti-cancer procedure or medication and those who received neither, so does not control for any imbalances in the types of treatment received in each group.

Results for the Western subgroup were available with and without adjustment for subsequent anti-cancer therapies (Table 12). The HR for the Western population lay in favour of sorafenib without adjustment for subsequent anti-cancer therapies, and in favour of [redacted] with the adjustment, but [redacted] (Table 12). For the Asia–Pacific subgroup, the HR lay in favour of lenvatinib [redacted]; the 95% CI crossed 1 without adjustment, [redacted] (Table 12). Results show that the relative treatment effect is sensitive to geographical region, and that differences in the extent of imbalance in subsequent anti-cancer therapies between the two regions means the adjustment [redacted]

Table 12. Overall survival in the Western and Asia–Pacific regions adjusted by use of subsequent anti-cancer treatments (adapted from CS, Table 13)

| | Stratified Cox Model Hazard Ratio (95% CI)* | |
|---|---|---|
| | **Stratification factors only** | **With subsequent treatment adjustment** |
| Full population | 0.92 (0.79 to 1.06) | [redacted] |
| Western (n = 314) | 1.08 (0.82 to 1.42) | [redacted] |
| Asia–Pacific (n = 640) | 0.86 (0.72 to 1.02) | [redacted] |

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status.
*Hazard ratio is for lenvatinib versus sorafenib, based on a Cox model including treatment group as a factor. The Efron method used for correction of tied events. Stratified by region (Region 1: Asia–Pacific; Region 2: Western), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1), and body weight (<60 kg, ≥60 kg); †Status of subsequent anti-cancer therapy (yes/no) was used as an additional covariate factor.

The ERG agrees with the company that REFLECT demonstrates non-inferiority of lenvatinib compared with sorafenib for OS based on the predefined threshold (upper 95% CI 1.08). Subsequent therapies are likely to have biased the relative treatment effect in favour of [redacted], though less so in the full population than the Western subgroup where there was more of an imbalance. The ERG finds it reassuring that results for OS in the Western subgroup are [redacted] to the full population when both are adjusted for subsequent treatments. The ERG thus considers the company's choice to use OS for the full population in the economic model appropriate.

### 4.3.2 Progression-free survival

At the primary data cut-off (13 November 2016), median follow-up for PFS was [redacted] in the lenvatinib group and [redacted] in the sorafenib group (Table 13). Median PFS from the primary analysis (investigator-assessed) was 7.4 months for patients treated with lenvatinib and 3.7 months for patients treated with sorafenib; the HR for lenvatinib versus sorafenib is 0.66 (95% CI: 0.57 to 0.77; Figure 4), meeting predefined criteria for superiority. The ERG notes that there was evidence of deviation from the PH assumption for this outcome, based on a formal assessment presented by the company (CS Section B.3.3.1.1.3), so the HR should be interpreted with caution.

Table 13. Investigator- and IIR-assessed progression-free survival in REFLECT

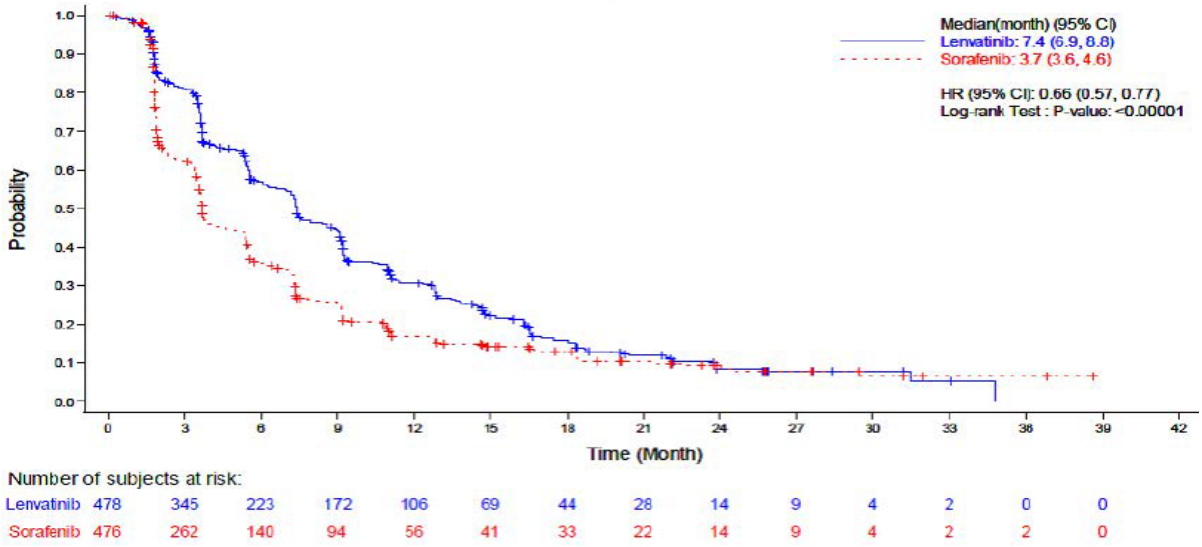
| | Lenvatinib (N = 478) | Sorafenib (N = 476) |
|---|---|---|
| **Investigator-assessed with mRECIST (primary analysis)** | | |
| Patients with events, n (%)<br>Progressive disease<br>Death | [redacted] | [redacted] |
| Median (95% CI), months* | 7.4 (6.9 to 8.8) | 3.7 (3.6 to 4.6) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.66 (0.57 to 0.77) | |
| PFS rate, % (95% CI)† 6 months<br>12 months<br>18 months<br>24 months | [redacted] | [redacted] |
| Median follow-up (95% CI), months | [redacted] | [redacted] |
| ***Post-hoc* analysis - investigator-assessed (mRECIST) with amended censoring rules** | | |
| Patients with events, n (%)<br>Progressive disease<br>Death | [redacted] | [redacted] |
| Median (95% CI), months* | [redacted] | [redacted] |
| Stratified Cox model hazard ratio (95% CI)‡,§ | [redacted] | |
| **IIR-assessed with mRECIST[41]** | | |
| Median PFS (95% CI), months* | 7.3 (5.6 to 7.5) | 3.6 (3.6 to 3.7) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.64 (0.55 to 0.77) | |
| **IIR-assessed with RECIST v1.1[41]** | | |
| Median (95% CI), months* | 7.3 (5.6 to 7.5) | 3.6 (3.6 to 3.9) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.65 (0.56 to 0.77) | |

Abbreviations: CI, confidence interval; IIR, independent imaging review; n = number of patients; mRECIST, modified Response Evaluation Criteria for Solid Tumours; PFS, progression-free survival; RECIST v1.1, Response Evaluation Criteria for Solid Tumours, version 1.1

* 95% confidence intervals are estimated with a generalised Brookmeyer and Crowley method; †PFS rate and 95% CI were calculated using the Kaplan–Meier product-limit method and the Greenwood Formula; ‡Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor; §Stratified by region (Region 1: Asia–Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg)



Figure 4. Kaplan–Meier curve for investigator-assessed progression-free survival (reproduced from CS, Figure 4)



Results from a retrospective IIR process using mRECIST and RECIST version 1.1 were also reported in the CS from a 2018 poster presentation[41] (Table 13). The respective Kaplan–Meier curves were provided in the CS Appendix and are reproduced in Appendix 10.5 (Figure 23). Results were similar to the primary PFS endpoint assessed by study investigators.

As described in Section 4.2.3, patients were censored in the primary analysis if they did not have progression at the time of treatment discontinuation, which the ERG did not consider appropriate given the imbalance between groups in reasons for treatment discontinuation (Section 4.2.1). The ERG was concerned that the censoring was informative (that is, potentially related to outcome), and that more lenvatinib events may be missed by premature censoring because discontinuation due to TEAEs and patient choice was more common in the lenvatinib group than the sorafenib group. The ERG asked for justification at the clarification stage, and the company provided a supplementary analysis where all events were counted in the analysis, and only patients with missing assessments or no progression were censored. Results from the supplementary analysis (Table 13), showed [redacted].

Comparison of HRs is suboptimal given that the PH assumption does not hold for PFS, but no Kaplan–Meier data were provided for the supplementary analysis to compare with that of the primary analysis. As such, it is difficult to assess the extent of potential overestimation of lenvatinib benefit caused by censoring, or the impact on the ICER should the alternative analysis of PFS be used in the economic model.

The ERG considered the consistency in direction of effect to provide robust evidence for a PFS benefit with lenvatinib compared with sorafenib, although the extent of benefit may be [redacted] in the economic model due to rules for censoring that potentially favour [redacted]

### 4.3.3 Time to progression

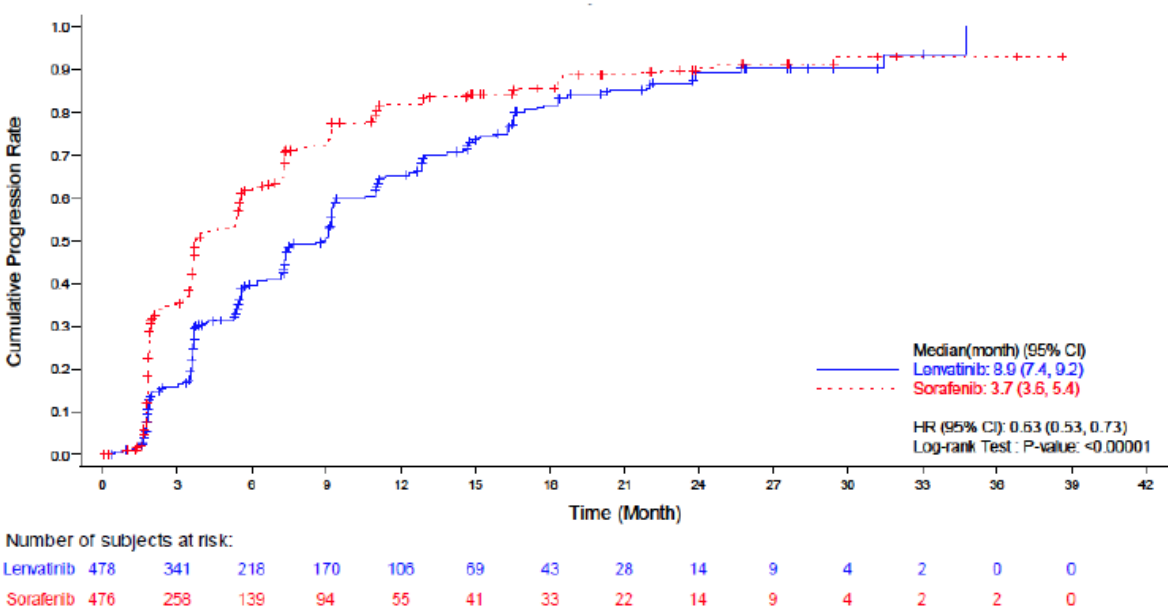
At the primary data cut-off (13 November 2016), median follow-up for TTP was [redacted] in the lenvatinib group and [redacted] in the sorafenib group (Table 14). Median TTP from the primary analysis (investigator-assessed), was 8.9 months for patients treated with lenvatinib and 3.7 months for patients treated with sorafenib; the HR for lenvatinib versus sorafenib is 0.63 (95% CI: 0.53 to 0.73; Figure 5), meeting predefined criteria for superiority. A formal assessment of PH was not conducted by the company, but the ERG advises caution in interpreting the HR given the similarity of the outcome to PFS for which the PH assumption does not hold.

Table 14. Investigator- and IIR-assessed ime to progression in REFLECT

| | Lenvatinib (N = 478) | Sorafenib (N = 476) |
|---|---|---|
| **Investigator-assessed mRECIST** | | |
| Patients with events, n (%) (PD only, death not an event) | [redacted] | [redacted] |
| Median TTP (95% CI), months† | 8.9 (7.4 to 9.2) | 3.7 (3.6 to 5.4) |
| Stratified Cox model hazard ratio (95% CI)§** | 0.63 (0.53 to 0.73) | |
| Cumulative progression rate, % (95% CI) ‡<br>6 months<br>12 months<br>18 months<br>24 months | [redacted] | [redacted] |
| Median follow-up (95% CI), months† | [redacted] | [redacted] |
| **IIR-assessed mRECIST[41]** | | |
| Median TTP (95% CI), months† | 7.4 (7.2 to 9.1) | 3.7 (3.6 to 3.9) |
| Stratified Cox model hazard ratio (95% CI)§** | 0.60 (0.51 to 0.71) | |
| **IIR-assessed RECIST v1.1[41]** | | |
| Median TTP (95% CI), months† | 7.4 (7.3 to 9.1) | 3.7 (3.6 to 5.4) |
| Stratified Cox model hazard ratio (95% CI)§** | 0.61 (0.51 to 0.72) | |

Abbreviations: CI, confidence interval; IIR, independent imaging review; n = number of patients; TTP, time to progression. *Deaths were not counted as progression events in this analysis; †95% confidence intervals are estimated with a generalised Brookmeyer and Crowley method; ‡Cumulative progression rate was calculated using the Kaplan–Meier product-limit method and Greenwood Formula; §Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor. Efron method was used for ties; **Stratified by region (Region 1: Asia–Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg).

Figure 5. Kaplan–Meier curve for time to progression (reproduced from CS, Figure 5)




Abbreviations: CS, company submission; CI, confidence interval; HR, hazard ratio.

Though not reported in the CS, results from a retrospective IIR process using mRECIST and RECIST v1.1 were available in the same poster presentation cited for the IIR PFS results[41] (Table 14). Median TTP was shorter in the lenvatinib group when assessed by IIR but the HR indicated a similar relative effect compared with sorafenib.

As discussed in Section 3.4, the ERG notes that TTP is often preferred in HCC research because the benefit of drugs in delaying cancer progression can be masked by deaths from the natural history of cirrhosis when PFS is used.[35] The similarity of group medians and HRs for TTP and PFS, particularly when comparing the IIR-assessed endpoints, suggests PFS was not confounded as such in REFLECT, which may reflect that patients had well-preserved liver function at study entry.

As discussed for PFS, the ERG noted for TTP the same issue with censoring patients who did not have progression at the time of treatment discontinuation (Table 10), which could favour the lenvatinib group (see 4.2.3). However, as TTP was not used in the economic model, the ERG did not consider it a priority for the company to provide an alternative analysis for this outcome. Naïve comparisons of the median TTP for the sorafenib group in REFLECT with the sorafenib regulatory trials for HCC[31, 32] show TTP was shorter than observed in the Western trial population of SHARP (5.5 months), and longer than observed in the Asia–Pacific trial (2.8 months). The ERG could not compare reasons for censoring between the sorafenib trials and REFLECT because the data were redacted in the committee papers for TA474.[24]

### 4.3.4 Response rates

The primary analysis of ORR based on investigator assessments using mRECIST showed a statistically significant benefit of lenvatinib compared with sorafenib (Table 15). In the lenvatinib group, 115 patients had a best overall response of complete or partial response (24.1%) compared with 44 patients (9.2%) in the sorafenib group (OR 3.13, 95% CI: 2.15 to 4.56). Results using the per protocol set were similar to the primary analysis [redacted]

As for PFS and TTP, results from a retrospective IIR process based on mRECIST and RECIST v1.1 were available from a conference abstract.[41] ORR according to mRECIST was considerably higher in the lenvatinib arm by IIR than judged by investigators (40.6% vs 24.1%, respectively), whereas the rates in the sorafenib arm were similar (12.4% vs 9.2%). Consequently, the magnitude of effect of lenvatinib versus sorafenib for ORR was much larger by IIR (OR 5.01, 95% CI: 3.59 to 7.01) than based on investigator assessments. The difference between investigator assessments in the lenvatinib arm can be seen in Table 15, which shows a larger proportion of patient were judged to have partial response by IIR, whereas investigators judged more patients to have stable disease; rates of complete response, progressive disease and unknown/unevaluable were similar between the IIR and investigator assessments.

Table 15. Response in REFLECT full analysis set (adapted from CS, Table 17 and Lencioni 2018[41])

| | Lenvatinib N = 478 | Sorafenib N = 476 |
|---|---|---|
| **Investigator-assessed response mRECIST** | | |
| **ORR (complete + partial response)** | 115 (24.1) | 44 (9.2) |
| **Odds ratio (95% CI)** | 3.13 (2.15 to 4.56) | |
| Best overall response | | |
| Complete response | 6 (1.3) | 2 (0.4) |
| Partial response | 109 (22.8) | 42 (8.8) |
| Stable disease | 246 (51.5) | 244 (51.3) |
| Durable stable disease* | 167 (34.9) | 139 (29.2) |
| Progressive disease | 71 (14.9) | 147 (30.9) |
| Unknown/not evaluable | 46 (9.6) | 41 (8.6) |
| **Independent imaging review mRECIST** | | |
| ORR (complete + partial response) | 194 (40.6) | 59 (12.4) |
| Odds ratio (95% CI) | 5.01 (3.59 to 7.01) | |
| Best overall response | | |
| Complete response | 10 (2) | 4 (1) |
| Partial response | 184 (38) | 55 (12) |
| Stable disease | 159 (33) | 219 (46) |
| Durable stable disease* | 84 (18) | 90 (19) |
| Progressive disease | 79 (17) | 152 (32) |
| Unknown/not evaluable | 46 (10) | 46 (10) |
| **Independent imaging review RECIST v1.1** | | |
| ORR (complete + partial response) | 90 (18.8) | 31 (6.5) |
| Odds ratio (95% CI) | 3.34 (2.17 to 5.14) | |
| Best overall response | | |
| Complete response | 2 (<1) | 1 (<1) |
| Partial response | 88 (18) | 30 (6) |
| Stable disease | 258 (54) | 250 (53) |
| Durable stable disease* | 163 (34) | 118 (25) |
| Progressive disease | 84 (18) | 152 (32) |
| Unknown/not evaluable | 46 (10) | 43 (9) |

Abbreviations: mRECIST, modified Response Evaluation Criteria in Solid Tumours; ORR, objective response rate; RECIST v1.1, Response Evaluation Criteria in Solid Tumours, version 1.1.
*≥ 23 weeks after randomisation. Odds ratio and 95% CI calculated using the Cochran-Mantel-Haenszel method.

Results based on IIR assessments conducted according to RECIST v1.1 (OR 3.34, 95% CI: 2.17 to 5.14) were closer to the investigator-assessed mRECIST results than when mRECIST results were compared using the two different assessors (Table 15). While it has been shown that ORR tends to be higher with mRECIST than with RECIST v1.1,[38] it is not clear why the variation between measures and assessors seen in the lenvatinib arm is not apparent in the sorafenib arm. Explanations for the pattern of results could include the sensitivity of each measure to the mechanism of action of each drug, detection bias, and familiarity of the assessors with the two measures. Nevertheless, the lenvatinib benefit over sorafenib is statistically significant and clinically meaningful regardless of the measure chosen. ORR was not included in the economic model.

### 4.3.5 Health-related quality of life

HRQoL data from REFLECT had not been published at the time this report was written, meaning the CS mainly included narrative summaries of selected results. HRQoL was assessed with three separate measures (Table 16) administered at baseline, day one of each treatment cycle and at the off-treatment visit. Results from the EORTC scales are discussed here. EQ-5D-3L data were incorporated in the economic model and are discussed in 5.4.8.

Table 16. Summary of HRQoL measures used in REFLECT

| Scale | Scope | Scoring |
|---|---|---|
| EORTC QLQ-HCC18 | HCC specific | 18 items across 8 scales.<br>1 ("not at all") to 4 ("very much"). |
| EORTC QLQ-C30 | Cancer generic | 30 items across 5 functional and 9 symptom domains<br>1 ("not at all") to 4 ("very much") or 1 ("very poor") to 7 ("excellent").<br>Summary score (1 to 100) and 1 to 7 global health status score (higher scores on both indicate better HRQoL). |
| EQ-5D-3L | Generic | 5 domains rated 1 ("no problems") to 3 ("extreme problems"): mobility, self-Care, usual activities, pain/discomfort, and anxiety/depression. Scores give a single health utility index.<br>EQ-VAS from 0 to 100 (worst to best imaginable health state). |

Abbreviations: EORTC-QLQ-HCC18, European Organisation for Research and Treatment of Cancer quality of life scale – hepatocellular carcinoma 18 items; EORTC QLQ-C30, EORTC QLQ cancer 30 items; EQ-5D-3L, EuroQol 5-dimensions, 3-levels; EQ-VAS, EuroQol visual analogue scale; HRQoL, health-related quality of life.

Completion of questionnaires remained above 90% throughout the Randomisation Phase, but the company considered attrition after cycle 18 too high to yield meaningful cross-sectional results (CS, Table 18). Most patients in both groups had discontinued treatment by the data cut-off (92.4% of the lenvatinib group and 94.7% of the sorafenib group) but a table of baseline, cross-sectional, and end of treatment scores for each domain and scale was not presented for comparison between groups. The ERG notes from the CSR that full HRQoL findings from REFLECT were presented in a separate report, which was not provided to the ERG. The following results are based on narrative and graphical summaries in the CS and CSR:

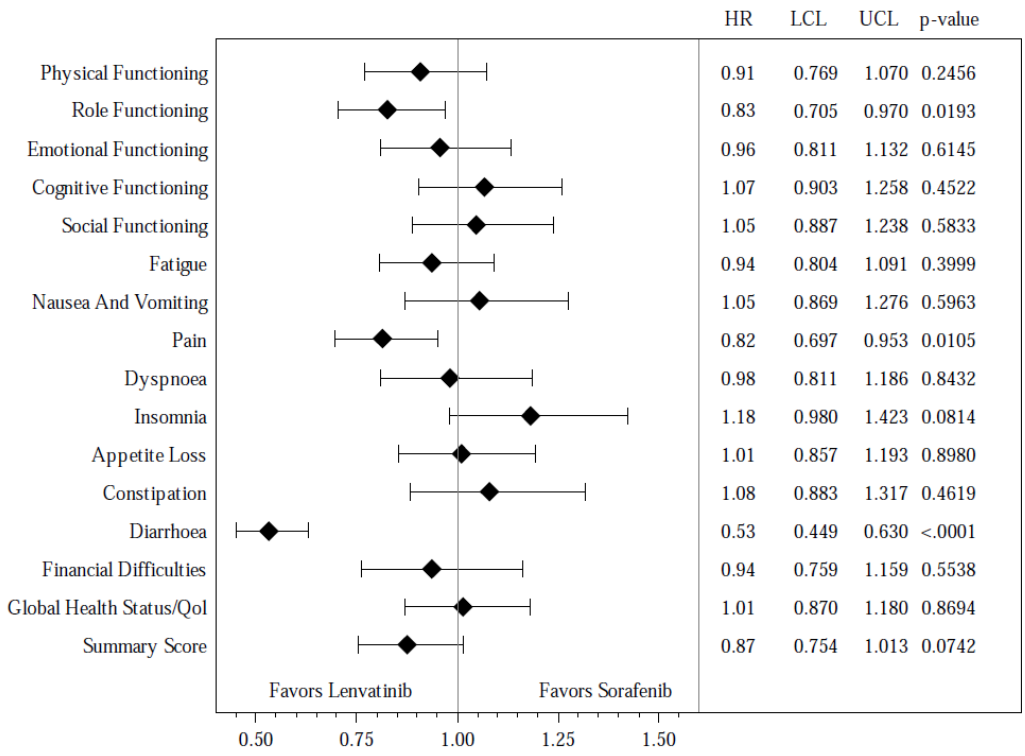
- [redacted];
- [redacted]

[redacted]

- Differences between groups were not statistically significant for the health utility index (HUI) or Visual Analogue Scale (VAS) of the EQ-5D-3L at Cycles 3, 6, 9, 12, 15, or 18 ($p>0.05$).

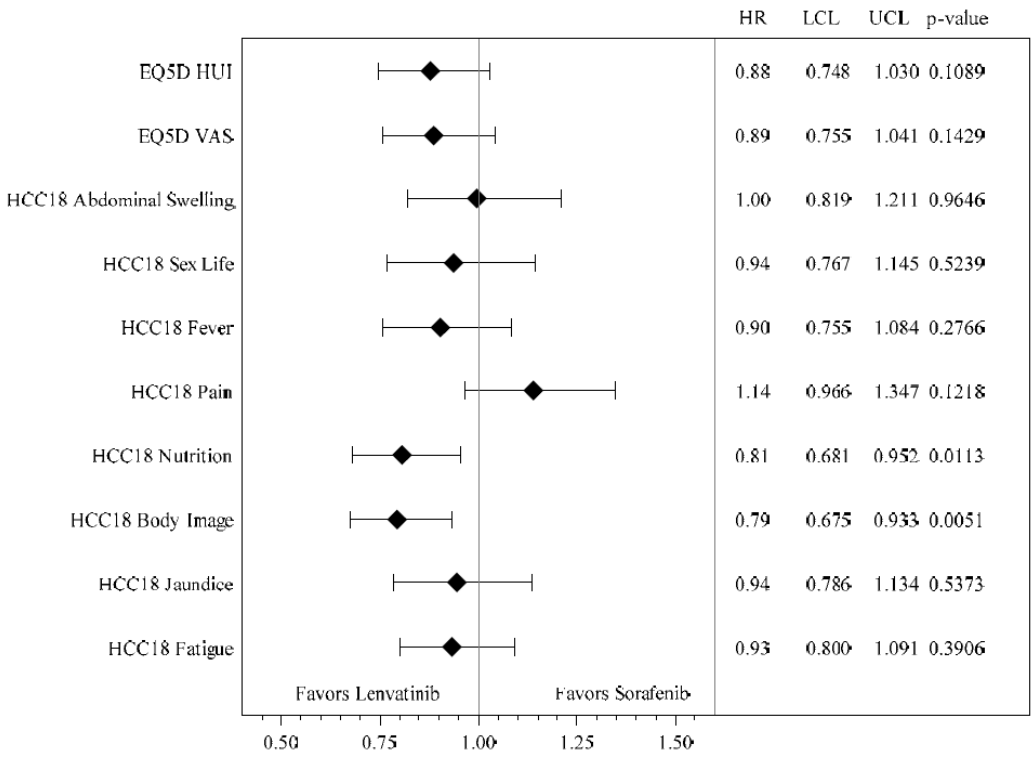
- HRs indicate longer time to clinically meaningful worsening for lenvatinib than sorafenib for the Role Functioning, Pain and Diarrhoea domains of the QLQ-C30 (Figure 6), and for the Body Image and Nutrition domains of the HCC18 (Figure 7).

Figure 6. Time to clinically meaningful worsening on the EORTC QLQ-C30 in REFLECT (reproduced from CS, Figure 6)




Abbreviations: EORTC, European Organisation for Research and Treatment of Cancer; HR, hazard ratio; LCL, lower confidence level; QLQ, Quality of Life Questionnaire; QoL, quality of life; UCL, upper confidence level.

Figure 7. Time to clinically meaningful worsening on the EORTC QLQ-HCC18 and EQ-5D (reproduced from CS, Figure 7)




Abbreviations: EORTC, European Organisation for Research and Treatment of Cancer; HCC, hepatocellular carcinoma; HR, hazard ratio; HUI, Health Utilities Index; LCL, lower confidence level; QoL, quality of life; QLQ, Quality of Life Questionnaire; UCL, upper confidence level; VAS, Visual Analogue Scale.

The ERG considers there to be some quality of life benefits associated with lenvatinib compared with sorafenib, and vice versa, in relation to specific symptoms and TEAEs. However, results provided in the CS and in the CSR mostly indicate [redacted] quality of life [redacted], although complete results for all scale domains across timepoints were not available.

## 4.3.6 Subgroup analyses

### *4.3.6.1 Baseline characteristics*

Subgroup analyses were prespecified for OS and PFS for the four randomisation stratification factors (region, macroscopic portal vein invasion and/or extrahepatic spread, ECOG performance status and body weight), and age group (≤65, ≥65 to <75 years and ≥75 years), sex and aetiology of HCC (HBV, HCV and alcohol). Additional subgroup analyses that were allowed in the SAP[56] were conducted for baseline AFP level (<200 ng/mL, ≥200 ng/mL, BCLC stage (B, C)). Results for all prespecified and additional subgroups were presented in CS Appendix E and are reproduced in Appendix 10.6 (Table 56 and Table 57).

For OS, HRs for lenvatinib versus sorafenib were generally consistent with the primary analysis, ranging from 0.84 (female, ≥65 to <75 years and ≥75 years) to 1.08 (Western region), with varying 95% CIs. The company highlights that the HR for the Western subgroup is driven by longer OS in the sorafenib group rather than shorter OS for lenvatinib in that region, which is likely to be at least partly explained by differences in subsequent anti-cancer therapies and procedures (see results of adjusted analyses in the full population and by region with results for OS in Section 4.3.1).

For PFS, HRs for all subgroup strata for baseline characteristics were in favour of lenvatinib and most can be considered consistent with results of the full population primary analysis. Most results showed a statistically significant benefit of lenvatinib compared with sorafenib, except the Western subgroup (HR 0.81, 95% CI 0.61 to 1.08), hepatitis C aetiology (HR 0.78, 95% CI: 0.56 to 1.09), and strata including only a small number of patients (i.e. age ≥75 years and female).

#### *4.3.6.2 Subsequent anti-cancer therapies*

The company conducted *post-hoc* subgroup analyses to explore the difference in absolute OS and PFS within groups for patients who received subsequent anti-cancer interventions. Separate subgroup analyses were conducted for:

- Patients who received either a subsequent anti-cancer procedure or medication compared with those who had received neither.
- Patients who had subsequent anti-cancer procedures compared with those who had not;
- Patients who received subsequent anti-cancer medications compared with those who had not;

In both groups, OS was substantially longer for patients who had received subsequent anti-cancer interventions (procedure or medication; Table 17) than those who had not. In the lenvatinib arm, median OS was 19.5 months (95% CI: 15.7 to 23.0) for patients who received subsequent anti-cancer interventions compared with 10.5 months (95% CI: 8.6 to 12.2) for those who had not, and the equivalent OS in the sorafenib group was 17.0 months (95% CI: 14.2 to 18.8) and 7.9 months (95% CI: 6.6 to 9.7), respectively. Furthermore, relative treatment effects were consistently more favourable for lenvatinib for patients who received subsequent anti-cancer interventions compared with those who did not, although there was overlap across 95% CIs. Results indicate that the use of subsequent anti-cancer therapy impacts on OS in REFLECT, which is discussed with results for *post-hoc* analyses to adjust to subsequent imbalance between groups in Section 4.3.1.

Table 17. OS subgroup analyses by subsequent anti-cancer procedures and medications (adapted from CS Appendix E, Table 5)

| Subsequent anti-cancer intervention | | Lenvatinib (N = 478) | | | Sorafenib (N = 476) | | | HR lenvatinib vs sorafenib (95% CI) |
|---|---|---|---|---|---|---|---|---|
| | | N | Died | Median (95% CI) | N | Died | Median (95% CI) | |
| Any | Yes | 206 | 143 | 19.5 (15.7, 23.0) | 243 | 175 | 17.0 (14.2, 18.8) | 0.84 (0.67, 1.06) |
| | No | 272 | 208 | 10.5 (8.6, 12.2) | 233 | 175 | 7.9 (6.6, 9.7) | 0.91 (0.74, 1.11) |
| Procedure | Yes | 99 | 63 | 23.0 (18.2, 27.7) | 112 | 82 | 19.6 (17.0, 22.9) | 0.71 (0.51, 1.01) |
| | No | 379 | 288 | 11.6 (10.1, 13.4) | 364 | 268 | 10.1 (8.7, 11.8) | 0.94 (0.79, 1.11) |
| Medication | Yes | 156 | 110 | 20.8 (15.1, 23.5) | 184 | 132 | 17.0 (14.4, 19.1) | 0.87 (0.67, 1.14) |
| | No | 322 | 241 | 11.5 (10.2, 13.2) | 292 | 218 | 9.1 (7.7, 10.3) | 0.90 (0.75, 1.09) |

Abbreviations: CI, confidence interval; HR, hazard ratio; N, number of patients; OS, overall survival.
Median estimated with the Kaplan–Meier product-limit method; 95% CI constructed with a generalised Brookmeyer and Crowley method; HR based on a Cox regression model including treatment group as a factor. Efron method used for correction of tied events; subsequent anti-cancer procedures excluded radiotherapy; subsequent anti-cancer medication excluded medications from procedures.

The same subgroups were explored for PFS which the ERG considered less clinically relevant than the OS results. Generally, subsequent therapies should not impact PFS because they are received after progression has occurred. In some cases, subsequent treatment could have commenced prior to progression if the study treatment was terminated due to TEAEs or patient choice, although these patients were censored in the primary analysis of PFS. The subgroup analysis could suggest subgroup differences between treatments if patients on one treatment were more likely to be healthy enough at progression to receive subsequent treatment, although this cannot be inferred with any certainty from the data. The ERG considers the results difficult to interpret, and presents the subgroup results for reference only (Table 18).

Table 18. PFS subgroup analyses by subsequent anti-cancer procedures and medication (adapted from CS Appendix E, Table 6)

| Subsequent anti-cancer intervention | | Lenvatinib (N = 478) | | | Sorafenib (N = 476) | | | HR lenvatinib vs sorafenib (95% CI) |
|---|---|---|---|---|---|---|---|---|
| | | N | Ev | Median (95% CI) | N | Ev | Median (95% CI) | |
| Any | Yes | 206 | 177 | 7.2 (5.6, 7.4) | 243 | 204 | 3.6 (3.5, 4.6) | 0.58 (0.47, 0.72) |
| | No | 272 | 172 | 8.0 (7.2, 9.3) | 233 | 163 | 3.7 (3.6, 5.5) | 0.70 (0.56, 0.87) |
| Procedure | Yes | 99 | 80 | 7.4 (5.6, 9.2) | 112 | 93 | 3.6 (3.5, 5.3) | 0.41 (0.29, 0.57) |
| | No | 379 | 269 | 7.4 (6.2, 8.8) | 364 | 274 | 3.7 (3.6, 5.4) | 0.71 (0.59, 0.84) |
| Medication | Yes | 156 | 137 | 5.7 (5.5, 7.4) | 184 | 157 | 3.8 (3.5, 5.5) | 0.66 (0.51, 0.85) |
| | No | 322 | 212 | 8.6 (7.3, 9.2) | 292 | 210 | 3.7 (3.6, 4.7) | 0.66 (0.54, 0.80) |

Abbreviations: CI, confidence interval; Ev, events (death or progression); HR, hazard ratio; N, number of patients; PFS, progression-free survival.
Median estimated with the Kaplan–Meier product-limit method; 95% CI constructed with a generalised Brookmeyer and Crowley method; HR based on a Cox regression model including treatment group as a factor. Efron method used for correction of tied events; subsequent anti-cancer procedures excluded radiotherapy; subsequent anti-cancer medication excluded medications from procedures.

### 4.3.7 Adverse effects

- No report from the Committee for Medicinal Products for Human Use (CHMP) has been issued because lenvatinib has not yet received marketing authorisation from the EMA for use in HCC. As such, the ERG has relied on safety data available in the CS and CSR, including the draft

SmPC (CS, Appendix A), which combines evidence from REFLECT with a study in differentiated thyroid cancer (DTC). The draft SmPC summarises the safety profile of lenvatinib as follows:

- [redacted]
- [redacted]
- [redacted]

Table 19 summarises important safety data collected in REFLECT, and TEAEs of any grade that occurred in at least 10% of patients in either group. Diarrhoea, asthenia and fatigue are also listed, on the advice of the company's clinical experts. The company presented total TEAE data for lenvatinib and separately for each dose (8 mg and 12 mg), whereas the ERG has opted to present only the combined lenvatinib safety data for clarity. Dose of lenvatinib is based on body weight and the ERG did not observe notable differences between doses in CS Tables 20, 21 and 22.

The company highlights that the incidence of TEAEs, that is, the number of patients in each group who experienced one or more, does not account for the longer treatment exposure of lenvatinib compared with sorafenib (see Table 6), and present exposure-adjusted rates alongside overall incidence (see Table 19). The ERG considers that, since treatment benefits are related to treatment exposure, it is appropriate to capture the burden of TEAEs over the full time on treatment. Furthermore, adjustment for exposure does not account for TEAEs that emerge in the early stages of treatment. Where reported in the draft SmPC, [redacted].

Table 19 shows that almost all patients in both groups experienced at least one TEAE (98.7% and 99.4% for lenvatinib and sorafenib, respectively), and the proportion considered by the investigators to be related to study treatment was also similar (93.9% and 95.2%). Drug withdrawals (19.7%), and dose modifications (61.8%, including dose reductions and interruptions) due to TEAEs in the lenvatinib are in line with combined data DTC and HCC presented in the draft lenvatinib SmPC (20.2% and 62.3%, respectively) and were higher than the sorafenib group (14.5% and 55.6%, respectively). The draft lenvatinib SmPC lists common reasons for dose modification as decreased appetite, diarrhoea, proteinuria, hypertension, fatigue, palmar-plantar erythrodysaesthesia (PPE) and decreased platelet count, and common reasons for withdrawal as hepatic encephalopathy, fatigue, blood bilirubin

increased, proteinuria and hepatic failure. Without information about when withdrawals and modifications occurred in each group, it is unknown how much of the difference can be explained by imbalance in treatment exposure.

Table 19. Summary of TEAEs by patient incidence and treatment exposure (adapted from CS, Tables 20, 21, 22 and 23)

| Adverse event | Lenvatinib (N = 476) | | Sorafenib (N = 475) | |
|---|---|---|---|---|
| | N patients with ≥1 (%) | Episodes/PY | N patients with ≥1 (%) | Episodes/PY |
| TEAE | 470 (98.7) | 6,124 (18.89) | 472 (99.4) | 4,718 (19.73 |
| Treatment-related TEAE | 447 (93.9) | 3,546 (10.94) | 452 (95.2) | 2,865 (11.98) |
| SAE total | 205 (43.1) | 409 (1.26) | 144 (30.3) | 232 (0.97) |
| Fatal | 61 (12.8) | 47 (0.21) | 36 (7.6) | 36 (0.15) |
| Non-fatal | 189 (39.7) | 271 (1.18) | 128 (26.9) | 207 (0.87) |
| Drug withdrawal due to TEAE | 94 (19.7) | - | 69 ( 4.5) | - |
| Dose modification due to TEAE (all) | 294 (61.8) | - | 264 ( 5.6) | - |
| Reductions | 184 (38.7) | - | 1 5 (38.6) | - |
| Interruptions | 248 (52.1) | - | 193 (40.6) | - |
| **Specific TEAEs occurring in ≥10% of patients by preferred term (safety analysis set)** | | | | |
| Hypertension | 201 (42.2) | 253 (0.78) | 144 (30.3) | 166 (0.69) |
| Diarrhoea | 184 (38.7) | 327 (1.01) | 220 (46.3) | 351 (1.47) |
| Decreased appetite | 162 (34.0) | 196 (0.60) | 127 (26.7) | 139 (0.58) |
| Weight decreased | 147 (30.9) | 165 (0.51) | 106 (22.3) | 113 (0.47) |
| Fatigue | 141 (29.6) | 156 (0.48) | 119 (25.1) | 130 (0.54) |
| Palmar-plantar erythrodysaesthesia syndrome | 128 (26.9) | 142 (0.44) | 249 (52.4) | 289 (1.21) |
| Proteinuria | 117 (24.6) | 163 (0.50) | 54 (11.4) | 74 (0.31) |
| Dysphonia | 113 (23.7) | 131 (0.40) | 57 (12.0) | 66 (0.28) |
| Nausea | 93 (19.5) | 113 (0.35) | 68 (14.3) | 77 (0.32) |
| Platelet count decreased | 87 (18.3) | 117 (0.36) | 58 (12.2) | 67 (0.28) |
| Abdominal pain | 81 (17.0) | 108 (0.33) | 87 (18.3) | 106 (0.44) |
| Hypothyroidism | 78 (16.4) | 79 (0.24) | 8 (1.7) | 8 (0.03) |
| Vomiting | 77 (16.2) | 106 (0.33) | 36 (7.6) | 57 (0.24) |
| Constipation | 76 (16.0) | 89 (0.27) | 52 (10.9) | 60 (0.25) |
| Blood bilirubin increased | 71 (14.9) | 97 (0.30) | 63 (13.3) | 75 (0.31) |
| Pyrexia | 69 (14.5) | 80 (0.25) | 63 (13.3) | 72 (0.30) |
| Ascites | 68 (14.3) | 77 (0.24) | 44 (9.3) | 51 (0.21) |
| Oedema peripheral | 66 (13.9) | 85 (0.26) | 33 (6.9) | 40 (0.17) |
| Aspartate aminotransferase increased | 65 (13.7) | 82 (0.25) | 80 (16.8) | 100 (0.42) |
| Abdominal pain upper | 58 (12.2) | 73 (0.23) | 40 (8.4) | 48 (0.20) |
| Asthenia | 54 (11.3) | 76 (0.23) | 48 (10.1) | 54 (0.23) |
| Alanine aminotransferase increased | 53 (11.1) | 67 (0.21) | 52 (10.9) | 66 (0.28) |
| Back pain | 50 (10.5) | 55 (0.17) | 31 (6.5) | 32 (0.13) |
| Rash | 46 (9.7) | 51 (0.16) | 76 (16.0) | 87 (0.36) |
| Stomatitis | 45 (9.5) | 55 (0.17) | 56 (11.8) | 67 (0.28) |
| Alopecia | 14 (2.9) | 15 (0.05) | 119 (25.1) | 123 (0.51) |

Abbreviations: CS, company submission; PY, patient year; SAE, serious adverse event; TEAE, treatment-emergent adverse event.

Table 19 also indicates more patients in the lenvatinib group (43.1%) than the sorafenib group (30.3%) had serious adverse events (SAEs), and the difference between groups persists when adjusted for treatment exposure. The following TEAEs of any grade were more notably more common in the lenvatinib group than the sorafenib group, respectively [redacted]: hypertension (42.2% vs 30.3%), decreased appetite (34.0% vs 26.7%), weight decreased (30.9% vs 22.3%), proteinuria (24.6% vs 11.4%), dysphonia (23.7% vs 12.0%), nausea (19.5% vs 14.3%), platelet count decreased (18.3% vs 12.2%), hypothyroidism (16.4% vs 1.7%), vomiting (16.2% vs 7.6%), constipation (16.0% vs 10.9%), ascites (14.3% vs 9.3%), peripheral oedema (13.9% vs 6.9%), and back pain (10.5% vs 6.5%). Rates of some TEAEs were [redacted] when adjusted for treatment exposure; however, as highlighted above, the adjustment may not be an accurate reflection of when TEAEs occurred, or of tolerability in clinical practice if time on treatment is expected to differ between sorafenib and lenvatinib. TEAEs of any grade experienced by notably fewer patients in the lenvatinib group than the sorafenib group, respectively, were: diarrhoea (38.7% vs 46.3%), palmar-plantar erythrodysaesthesia syndrome (PPES; 26.9% vs 52.4%), rash (9.7% vs 16.0%), and alopecia (2.9% vs 25.1%; Table 19).

Table 20 lists TEAEs of grade 3 or above that were reflected in the company's economic model. TEAEs of grade 3 listed in the table are those that occurred in at least 5% of either group, plus diarrhoea, asthenia and fatigue on the advice of the company's clinical experts. More patients in the lenvatinib group than the sorafenib group experienced Grade 3, 4, or 5 TEAEs than the sorafenib group (75.0% vs 66.5%, respectively), and more in the lenvatinib group than the sorafenib group were judged by the investigators to be related to study treatment (56.7% vs 48.6%, respectively). Grade 3+ TEAEs adjusted by treatment exposure were similar (3.16 vs 3.33 episodes per patient year for the lenvatinib and sorafenib groups, respectively) and rates judged by to be related showed the opposite direction of effect to the incidence data (1.59 vs 1.80 for lenvatinib and sorafenib, respectively [CS, Table 21]). Grade 3 or 4 TEAEs that were more common in the lenvatinib group than the sorafenib group were: hypertension (23.3% vs 14.3%), weight decrease (7.6% vs 2.9%) and proteinuria (5.7% vs 1.7%). PPES was the only grade 3 or 4 TEAE that was less common in the lenvatinib group than the sorafenib group (2.9% vs 11.4%; [redacted]see CS, Appendix L.5, Table 23).

The ERG requested subgroup data for treatment exposure (see Table 6) and TEAEs at the clarification stage to assess any differences between the full population and Western subgroup. The ERG considered mean and median dose of lenvatinib similar in the full population and Western subgroup. Comparing Grade 3+ TEAEs with the full population, the Western subgroup experienced [redacted]fatigue and asthenia[redacted]platelet count than the full population, regardless of treatment received (Table 20). Differences between the incidence of decreased weight, proteinuria and PPES that favoured sorafenib in the full population were [redacted]in the Western subgroup

Table 20. Grade 3 and 4 TEAEs in REFLECT (adapted from CS, Tables 20 and 23, and company response to clarification, Table 10)

| | Lenvatinib (N = 476) | | | Sorafenib (N = 475) | | |
|---|---|---|---|---|---|---|
| | Full population | | Western | Full population | | Western |
| | Grade 3 | Grade 4 | 3 or 4 | Grade 3 | Grade 4 | 3 or 4 |
| Any Grade 3, 4 or 5 TEAE | 357 (75.0) | | - | 316 (66.5) | | - |
| Any Grade 3 or 4 TEAE | 260 (54.6) | 36 (7.6) | - | 248 (52.2) | 32 (6.7) | - |
| Treatment-related Grade 3+ TEAE | 270 (56.7) | | - | 231 (48.6) | | - |
| **Grade 3 or 4 TEAEs occurring in ≥5%* of patients by preferred term (safety analysis set)** | | | | | | |
| Hypertension | 111 (23.3) | 0 (0.0) | 36 (23.2) | 68 (14.3) | 0 (0.0) | 18 (11.5) |
| Weight decreased | 36 (7.6) | 0 (0.0) | 18 (11.6) | 14 (2.9) | 0 (0.0) | 11 (7.1) |
| Proteinuria | 27 (5.7) | 0 (0.0) | 5 (3.2) | 8 (1.7) | 0 (0.0) | 6 (3.8) |
| Platelet count decreased | 26 (5.5) | 0 (0.0) | 2 (1.3) | 14 (2.9) | 2 (0.4) | 1 (0.6) |
| Blood bilirubin increased | 24 (5.0) | 7 (1.5) | 11 (7.1) | 21 (4.4) | 2 (0.4) | 5 (3.2) |
| Aspartate aminotransferase increased | 21 (4.4) | 3 (0.6) | 7 (4.5) | 32 (6.7) | 6 (1.3) | 6 (3.8) |
| Gamma-glutamyl transferase increased | 20 (4.2) | 6 (1.3) | 10 (6.5) | 16 (3.4) | 3 (0.6) | 9 (5.8) |
| Palmar-plantar erythrodysaesthesia syndrome | 14 (2.9) | 0 (0.0) | 6 (3.9) | 54 (11.4) | 0 (0.0) | 14 (9.0) |
| Diarrhoea | 20 (4.2) | 0 (0.0) | 8 (5.2) | 20 (4.2) | 0 (0.0) | 11 (7.1) |
| Fatigue | 18 (3.8) | 0 (0.0) | 11 (7.1) | 17 (3.6) | 0 (0.0) | 11 (7.1) |
| Asthenia | 14 (2.9) | 0 (0.0) | 9 (5.8) | 11 (2.3) | 0 (0.0) | 9 (5.8) |

Abbreviations: TEAE, treatment-emergent adverse event.
*Grade ¾ diarrhoea, fatigue and asthenia occurred in less than 5% of either group but were considered clinically relevant.

In addition to the safety issues highlighted above, TEAEs with an incidence below the reporting thresholds used by the company that are highlighted in the SmPC for lenvatinib (based on data from REFLECT) were: [redacted].

The SmPC also highlights that lenvatinib [redacted]

## 4.4 Critique of the indirect comparison and/or multiple treatment comparison

No indirect comparison was conducted by the company because direct evidence of lenvatinib versus sorafenib, the comparator of interest, was available from the REFLECT study. As described in Section 3.3, while best supportive care was listed as a potential comparator in the final scope issued by NICE,

the ERG's clinical experts agreed with the company that sorafenib is the only relevant comparator. The ERG agrees with the company that the head-to-head REFLECT study provides the most reliable evidence to inform the decision problem of interest to this STA, and does not consider the incorporation of indirect evidence necessary or appropriate. Where there was uncertainty in the robustness or clinical plausibility of evidence from REFLECT, the ERG consulted its clinical experts and referred to related studies identified in the company's systematic literature review (SLR; see Section 4.1).

## 4.5 Summary and conclusions of clinical effectiveness sections

- The company proposes lenvatinib as an alternative to sorafenib as first-line systemic therapy for patients with advanced or unresectable HCC, Child–Pugh class A liver function, BCLC Stage C disease, or patients with BCLC Stage B disease who were not eligible for TACE.

- At the time of writing, the application to extend the European marketing authorisation for lenvatinib to include the treatment of HCC had not been approved; the draft indication is for the treatment of adult patients who have received no prior systemic therapy for HCC.

- While BSC was listed as a potential comparator in the scope, the ERG's clinical experts agreed that it would only be chosen if a patient refused or was considered unfit for systemic treatment.

- The company's SLR identified the international phase III, open-label RCT, REFLECT, as the only study providing direct evidence for the population, intervention, comparator (sorafenib) and outcomes outlined in the final scope issued by NICE:

    - Population: 954 adults with previously untreated, advanced or unresectable HCC. Patients were Child–Pugh class A and primarily BCLC stage B. Around two-thirds were from the Asia–Pacific region (N = 640), and a third from the Western region (N = 314), including 20 patients from the UK;

    - Intervention: lenvatinib once daily oral dose of 8 mg for patients weighing less than 60 kg, and 12 mg for patients weighing 60 kg or greater;

    - Comparator: sorafenib twice daily oral dose of 400 mg;

    - Outcomes: OS, PFS, TTP, response rates, adverse effects, and HRQoL. The primary objective of REFLECT was to show non-inferiority of lenvatinib compared with sorafenib in terms of OS, with an upper 95% CI of 1.08. The company's statistical approach was generally appropriate and well described. However, proportional hazards (PH) assessments conducted by the company suggest the assumption does not hold for OS or PFS, so HRs from Cox PH models should be interpreted with caution.

- Potential oversights in the company's SLR with regards to contextual evidence are unlikely to impact the submission given the robustness and maturity of evidence from RELFECT.

- The ERG's clinical experts consider REFLECT to provide evidence that is relevant to the decision problem, but highlight that the inclusion criteria mean results may not generalise to those with compromised liver function (Child–Pugh stage B or worse) or ECOG performance status of 2 or above.

- Lenvatinib met criteria for non-inferiority to sorafenib in the primary analysis of OS (full population, adjusted for stratification factors). At the final data cut-off, 73.4% of the lenvatinib group and 73.5% of the sorafenib group had died, and median OS in the full population was 13.6 months and 12.3 months, respectively (HR 0.92, 95% CI: 0.79 to 1.06).

- Median PFS and TTP were consistently longer in the lenvatinib group than the sorafenib group, regardless of the measure used (mRECIST or RECIST v1.1) or outcome assessor (investigator or IIR). The primary analyses were investigator-assessed with mRECIST (a validated modified set of criteria for HCC) and indicated median PFS of 7.4 months for lenvatinib and 3.7 months for sorafenib (HR 0.66, 95% CI: 0.57 to 0.77), and TTP of 8.9 months and 3.7 months, respectively (HR 0.63, 95% CI: 0.53 to 0.73).

- The ERG considers rules for censoring in the company's primary analyses of PFS and TTP (based on Food and Drug Administration [FDA] guidance) to favour lenvatinib. An alternative PFS analysis was provided at the clarification stage based guidance from the European Medicines Agency (EMA); median PFS in both groups was [redacted] impact on the relative treatment effect [redacted].

- While TTP is often preferred over PFS for HCC because cirrhosis-related deaths can mask drug benefit in PFS,[35] the similarity of group medians and HRs between TTP and PFS may reflect that patients have well-preserved liver function at baseline in the study.

- ORR was also higher with lenvatinib than with sorafenib in the primary investigator-assessed analysis based on mRECIST (24.1% vs 9.2%; OR 3.13, 95% CI: 2.15 to 4.56, p<0.00001). The magnitude of effect varied between the primary analysis and IIR assessments (mRECIST and RECIST v1.1), but the difference was statistically significant and clinically meaningful regardless of the measure chosen.

- Across HRQoL scales, there appears to be some symptom and TEAE benefits of lenvatinib compared with sorafenib ([redacted], role functioning,

pain, body image), and vice versa ([redacted]), but results [redacted].

- Most subgroup and supplementary analyses of OS and PFS were consistent with the primary analyses. Notable results that were inconsistent with the primary analyses included OS in the Western geographical subgroup, for which the HR lay in favour of sorafenib (HR 1.08, 95% CI: 0.82 to 1.42), [redacted]; possible reasons for the difference in effect are discussed with reference to the applicability of the full population to UK clinical practice in Section 1.5.2. While the PH assumption does not hold for PFS, the 95% CI for the HR indicate that the benefit of lenvatinib is not statistically significant in the Western subgroup (HR 0.81, 95% CI: 0.61 to 1.08), those with hepatitis C aetiology (HR 0.78, 95% CI: 0.56, 1.09), and other strata that included a small number of patients (age ≥75 years and female).

- Almost all patients in both groups experienced at least one TEAE (98.7% and 99.4% for lenvatinib and sorafenib, respectively), and the proportion considered by the investigators to be related to study treatment was also similar (93.9% and 95.2%). The ERG considered it more appropriate to compare incidence over the full course of treatment, rather than exposure-adjusted rates, because the former is consistent with the effectiveness evidence.

- Drug withdrawals due to TEAEs (19.7% vs 14.5% for lenvatinib and sorafenib, respectively), dose modifications due to TEAEs (61.8% vs 55.6%), SAEs (43.1% vs 30.3%) and Grade 3+ TEAEs (75.0% vs 66.5%; 56.7% vs 48.6% judged by the investigators to be related to study treatment) were all more common in the lenvatinib group than the sorafenib group.

- AEs of any grade that were more notably more frequent in the lenvatinib group than the sorafenib group ([redacted]) were hypertension (42.2% vs 30.3%), decreased appetite (34.0% vs 26.7%), weight decreased (30.9% vs 22.3%), proteinuria (24.6% vs 11.4%), dysphonia (23.7% vs 12.0%), nausea (19.5% vs 14.3%), platelet count decreased (18.3% vs 12.2%), hypothyroidism (16.4% vs 1.7%), vomiting (16.2% vs 7.6%), constipation (16.0% vs 10.9%), ascites (14.3% vs 9.3%), peripheral oedema (13.9% vs 6.9%), and back pain (10.5% vs 6.5%). TEAEs of any grade experienced by notably fewer patients in the lenvatinib group than the sorafenib group, respectively ([redacted]), were diarrhoea (38.7% vs 46.3%), PPES (26.9% vs 52.4%), rash (9.7% vs 16.0%), and alopecia (2.9% vs 25.1%).

- Grade 3–4 TEAEs that were more common ([redacted]) in the lenvatinib group than the sorafenib group were hypertension (23.3% vs 14.3%), weight

decrease (7.6% vs 2.9%) and proteinuria (5.7% vs 1.7%). PPES was the only grade 3–4 TEAE that was notably less common in the lenvatinib group than the sorafenib group (2.9% vs 11.4%).

### 4.5.1 Clinical issues

- Baseline imbalances that may favour the sorafenib group were noted in the proportion of patients with AFP levels ≥200 ng/ml and in aetiology of HCC (hepatitis B, hepatitis C, or alcohol). Adding the variables as covariates in supplementary analysis of OS had some impact on the relative treatment effect.

- The primary analyses of PFS and TTP were based on investigator assessments and may be prone to bias because REFLECT was an open-label study. Alternative analyses were available from retrospective IIR assessments, and results were similar.

- A version of RECIST modified to evaluate change more accurately in HCC (mRECIST) was used for the primary tumour assessments for PFS, TTP and ORR in REFLECT. mRECIST is endorsed by the European Association for the Study of the Liver (EASL), and the ERG's clinical experts advised that UK centres choose either mRECIST and RECIST v1.1. IIR results were available for comparison, and meaningful differences were only noted for ORR, which did not inform the economic model.

- The ERG was concerned that the majority Asia–Pacific population may reduce generalisability of the full REFLECT population to UK patients. The following differences were noted, but the ERG did not consider there to be sufficient evidence that results from the Western subgroup were more applicable to UK clinical practice than the mixed full population to justify the inevitable loss of precision by focussing on a subgroup:

    - *Baseline characteristics*: compared with the full population, the Western subgroup were heavier (~13% < 60 kg vs ~30%), had more heart disease (~23% NYHA class I or II vs ~10%), less underlying cirrhosis (40% vs 50%), less hepatitis B (20–25% vs ~50%), and more hepatitis C, alcohol, other, or unknown aetiology. Hepatitis C aetiology has been linked to increased benefit with sorafenib, so the full population may therefore overestimate the relative effect of lenvatinib for UK patients. However, evidence from separate Western and Asia–Pacific trials for sorafenib in HCC (SHARP[31] and the Asia–Pacific study[32]) suggest that the pattern of baseline differences does not necessarily impact relative treatment effects;

    - *Treatment exposure*: Treatment duration was longer in the lenvatinib group than the sorafenib group in the full population and the difference between groups was smaller

in the Western subgroup. Time on treatment and dose intensity was somewhat higher in both groups in the Western subgroup than the full population (Table 6), which is likely due to weight differences between the Western and Asia–Pacific subgroups;

- *Adverse events*: comparing Grade 3+ TEAEs with the full population, the Western subgroup experienced more fatigue and asthenia, and less decrease platelet count than the full population, regardless of treatment received. Differences between the incidence of decreased weight, proteinuria and PPES that favoured sorafenib in the full population were smaller in the Western subgroup;

- *Subsequent treatments*: neither the full population nor the Western subgroup were considered reflective of UK clinical practice with regards to the extent and type of subsequent treatments received. There was more imbalance in the Western subgroup, mostly due to patients in the sorafenib group being eligible for clinical trials after discontinuation of the study drug;

- *Clinical outcomes*: Results for OS and PFS in the Western subgroup were both less favourable than results for the full population, and less precise due to the restricted population. As may be expected from pattern of imbalance in subsequent interventions described above, results for the full population and Western subgroup [redacted] when the analyses included a binary variable adjustment for patients who received subsequent anti-cancer interventions.

- PH assessments conducted by the company suggest the assumption does not hold for OS or PFS in REFLECT, so hazard ratios calculated from Cox PH models should be interpreted with caution; independent statistical models were considered for both PFS and OS in the review of cost-effectiveness.

- Censoring in the primary analyses of PFS and TTP was based on guidance from the FDA and included censoring at treatment discontinuation if there was no disease progression. The primary analyses are likely to favour lenvatinib because treatment discontinuation for reasons other than progression (that is, due to TEAEs or patient choice) was more common in the lenvatinib group than the sorafenib group. A sensitivity analysis of PFS provided by the company at the clarification stage, based on EMA guidance, included all progressions and deaths as events; median PFS in both groups [redacted]. Comparing HRs may be unreliable because the PH assumption does not hold and K–M data were not available so the consequences of

using this analysis could not be evaluated in the economic model. The ERG considered the consistency in direction of effect to provide robust evidence for a PFS benefit with lenvatinib compared with sorafenib, although the rules for censoring mean the extent of benefit may be [redacted] by the primary PFS analysis in the economic model.

- In the lenvatinib group, median time on treatment (5.7 months) was shorter than the primary analyses of median PFS and TTP (7.4 to 8.9 months, respectively), which may reflect patients discontinuing for reasons other than disease progression (TEAEs and patient choice); median PFS, TTP and time on treatment were equivalent in the sorafenib group (3.7 months).

- The ERG advises caution in the interpretation of exposure-adjusted TEAE rates, because they are less likely than incidence to capture burden of TEAEs across the full course of treatment. Adjusting for exposure causes inconsistency with the effectiveness analyses, and does not account for the tendency of TEAEs to emerge in the early stages of treatment.

# 5 COST EFFECTIVENESS

## 5.1 Introduction

This section provides a structured description and critique of the systematic literature review and *de novo* economic evaluation submitted by the company. The company provided a written submission of the economic evidence along with an electronic version of the Microsoft® Excel based economic model. Table 21 summarises the location of the key economic information within the company's submission (CS).

Table 21. Summary of key information within the company's submission

| Information | Section (CS) |
|---|---|
| Details of the systematic review of the economic literature | B.3.1 |
| Model structure | B.3.2.2 |
| Technology | B.3.2.3 |
| Clinical parameters and variables | B.3.3 |
| Measurement and valuation of health effects and adverse events | B.3.4 |
| Resource identification, valuation and measurement | B.3.5 |
| Summary of base-case analysis inputs and assumptions | B.3.6 |
| Results | B.3.7 |
| Sensitivity analysis | B.3.8 |
| Subgroup analysis | B.3.9 |
| Validation | B.3.10 |
| Interpretation and conclusions | B.3.11 |
| Abbreviations used in table: CS, company submission. | |



## 5.2 Summary of the company's key results

The results of the company's deterministic base case analysis are presented in Table 22, and the results of the company's probabilistic sensitivity analysis are given in Table 23.

Table 22. Base-case results (taken from the company's response to clarification)

| Technologies | Total costs | Total LYG | Total QALYs | Incremental costs | Incremental LYG | Incremental QALYs | ICER |
|---|---|---|---|---|---|---|---|
| Sorafenib | £65,592 | 1.46 | 1.03 | - | - | - | - |
| Lenvatinib | [redacted] | 1.69 | [redacted] | [redacted] | 0.23 | [redacted] | Dominated |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. | | | | | | | |

Table 23. PSA results using 10,000 iterations (reproduced from the economic model provided at clarification)

| Technologies | Total costs | Total QALYs | Incremental costs | Incremental QALYs | ICER |
|---|---|---|---|---|---|
| Sorafenib | £65,688 | 1.03 | - | - | - |
| Lenvatinib | [redacted] | [redacted] | [redacted] | [redacted] | Dominated |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years | | | | | |

## 5.3 ERG comment on company's review of cost-effectiveness evidence

The company carried out a systematic literature review (SLR) to identify cost-effectiveness and resource use evidence in patients with treatment-naïve advanced, unresectable or metastatic hepatocellular carcinoma (HCC). Another SLR was carried out to identify studies reporting on the health-related quality of life (HRQoL) of patients with advanced, unresectable or metastatic HCC. All database searches were carried out in November 2017.

When conducting the SLRs, the company searched the MEDLINE and Embase databases, as well as the Cochrane Library, using the OVID platform. In addition, conference proceedings and grey literature, including relevant heath technology assessment (HTA) bodies, were also hand searched for evidence. Search strategies are provided in Appendix G for cost-effectiveness and resource use evidence and Appendix H for HRQOL evidence. In summary, search terms combined the population (patients with advanced, unresectable or metastatic HCC) with economic and quality of life terms, which the ERG considers to be inclusive.

Overall, a total of 46 cost-effectiveness and resource use studies met the pre-specified inclusion and exclusion criteria reported in Table 7 of Appendix G. Of the 46 included studies, 27 were cost-effectiveness analyses, while 19 were budget impact or cost analyses. A summary of the eight cost-effectiveness analyses conducted from a UK perspective is provided in Table 25 of the CS. However, none of the 19 included resource use studies were conducted in the UK and, therefore, data from those studies were not extracted by the company. However, it is important to clarify that the company considered the included cost-effectiveness studies (including TA189[34] and TA474[24]) to inform resource use inputs in their submission, although they were not reported again under resource use studies. Sources of resource and cost use data are described in greater detail in Section 5.4.9.

As for the HRQoL search, a total of 14 HRQoL studies met the pre-specified inclusion and exclusion criteria reported in Table 10 of Appendix H. A summary of the 14 HRQoL studies is provided in Table 11 of Appendix H. However, as described in Section 5.4.8, it was not necessary to incorporate data from any of the identified studies because EQ-5D-3L data directly from the REFLECT trial were used to populate the economic model.

The company did not provide the lists of excluded studies for either of the searches. For completeness, those lists were provided by the company at the clarification stage.

In summary, the ERG considers the inclusion criteria to be appropriate to capture recent and relevant published evidence. Due to time constraints, the ERG was unable to replicate the company's search and appraisal of identified abstracts for all databases.

## *5.4 Overview and critique of company's economic evaluation*

The company developed a *de novo* economic model to assess the cost effectiveness of lenvatinib compared with sorafenib in patients with untreated, advanced or unresectable, HCC and Child-Pugh Class A liver function.

The following subsections describe the company's methods in more detail including the model structure, the data sources used and applicability of the analysis in comparison to the NICE reference case.[57]

### 5.4.1 NICE reference case checklist

Table 24 summarises the ERG's assessment of the company's economic evaluation against the requirements set out in the NICE reference case checklist for the base case analysis, with reference to the NICE final scope outlined in Section 3.[1, 57]

Table 24. NICE reference case checklist

| Attribute | Reference case | Does the *de novo* economic evaluation match the reference case? |
|---|---|---|
| Decision problem | The scope developed by NICE | Yes |
| Comparator(s) | Alternative therapies routinely used in the NHS | Yes |
| Perspective costs | NHS and Personal Social Services | Yes |
| Perspective benefits | All health effects on individuals | Yes |
| Form of economic evaluation | Cost-utility analysis | Yes |
| Time horizon | Sufficient to capture differences in costs and outcomes | Yes. The time horizon was set at 20 years, which was deemed sufficient to capture the lifetime of patients. |
| Synthesis of evidence on outcomes | Systematic review | Yes. A systematic review was conducted to identify data sources for outcome measures including disease progression, mortality and quality of life. |
| Outcome measure | Quality adjusted life years | Yes |
| Health states for QALY | Described using a standardised and validated instrument | EQ-5D-3L |
| Benefit valuation | Time-trade off or standard gamble | Yes. Time-trade of valuation of the EQ-5D-3L. |

| | | |
|---|---|---|
| Source of preference data for valuation of changes in HRQoL | Representative sample of the public | Yes. EQ-5D-3L UK tariff. |
| Discount rate | An annual rate of 3.5% on both costs and health effects | Yes |
| Equity | An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit | Yes |
| Sensitivity analysis | Probabilistic sensitivity analysis | Yes |
| Abbreviations: EQ-5D, EuroQol-five dimensions questionnaire; HRQoL, health-related quality of life; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life year. | | |

## 5.4.2 Population

The company's economic analysis evaluates lenvatinib compared with sorafenib in an intended population of patients with untreated advanced or unresectable HCC and Child–Pugh Class A liver function. The key source of data for clinical outcomes was the intention-to-treat (ITT) population of the REFLECT trial; a phase III trial comparing lenvatinib with sorafenib with a population for which 99% of patients had Child-Pugh class A liver function.[58] The remainder had class B liver function.

The population in the trial were located in two main geographical regions; an Asia–Pacific region and a Western subgroup. The company's base case modelling includes all patients from both regions and no subgroup analyses were performed in the company's original economic model, despite subgroup analyses by region for the clinical outcomes being performed.

### *5.4.2.1 ERG critique*

The intended population in the company's analysis is in line with the NICE final scope of adults with unresectable HHC who have not previously received systemic treatment. The data used to inform the economic analysis is also largely in line with this intended population with only 1% of patients not having the Child-Pugh class A liver function.

The ERG considers the NICE final scope to have been fulfilled by the company's economic analysis using data from an appropriate source. The ERG considers the Western subgroup population to potentially be more reflective of the UK population and that a subgroup analysis may be more appropriate than the full population of the REFLECT trial.[58] However, following further analyses submitted in response to the ERG's clarification questions, the ERG considers it reasonable to assume that the efficacy is similar in the full population. This is discussed in more detail in Section 4 and 5.4.5.

## 5.4.3 Interventions and comparators

The company expect to receive marketing authorisation for lenvatinib in adult patients who have received no prior systemic therapy for HCC in [redacted]. However, this is currently only the draft

indication. The comparator treatment in the company's analysis is sorafenib, a NICE recommended treatment only for patients with advanced HCC and Child–Pugh class A liver function.

The company's model assumes that patients receiving lenvatinib or sorafenib will receive treatment until progression, death, or other withdrawal based on the protocol of the REFLECT trial,[58] the draft Summary of Product Characteristics (SmPC), and expected UK clinical practice.

Lenvatinib and sorafenib were modelled with dosing based on the REFLECT trial,[58] in which starting doses of 12 mg for patients weighing >60 kg, and 8 mg for patients weighing <60 kg, were used for lenvatinib. Sorafenib was given at a daily dose of 800 mg (400mg twice per day) regardless of body weight.

#### *5.4.3.1 ERG critique*

The ERG considers the intervention and comparator to appropriately reflect those of the decision problem set out in the NICE final scope. However, the mean dose used in the company's base case modelling is based on the mean dose of the full population of the REFLECT trial and may not fully REFLECT the dose used in the UK.[58] The ERG considers the Western subgroup dose may be more reflective of the UK population.

### 5.4.4 Modelling approach and model structure

The company submitted a *de novo* economic model in Microsoft® Excel using a partitioned survival structure with health-states for progression-free (PF) disease, progressed disease (PD), and death. This type of model uses survival models (e.g. parametric survival curves) for progression-free survival (PFS) and overall survival (OS) to form partitions in the population that define the proportion of patients in each health state. The PFS survival model determines the proportion of patients in the PF health state at a given time, while the OS survival model determines the proportion of patients who are alive at a given time, and therefore, also the proportion who are in the death health-state. The difference between these PFS and OS values determines the proportion of patients in the PD health state at a given time. The proportions are calculated in cycles every 28 days up to the time horizon of 20 years. Relevant unit costs and health-state utility values (HSUVs) are then applied at each cycle and summed over the time horizon to estimate the total costs and quality-adjusted life-years (QALYs) for each treatment to calculate the incremental cost effectiveness ratio (ICER) and perform the economic evaluation.

The model structure is shown in Figure 8 and the partitions used in the model are depicted in Figure 9. The ERG's critique of the modelling approach and structure is given in Section 5.4.4.1, while the subsequent sections describe and critique the data sources and statistical models used to inform the economic analysis in more detail.

Figure 8. Model structure (CS, page 74, Figure 8)

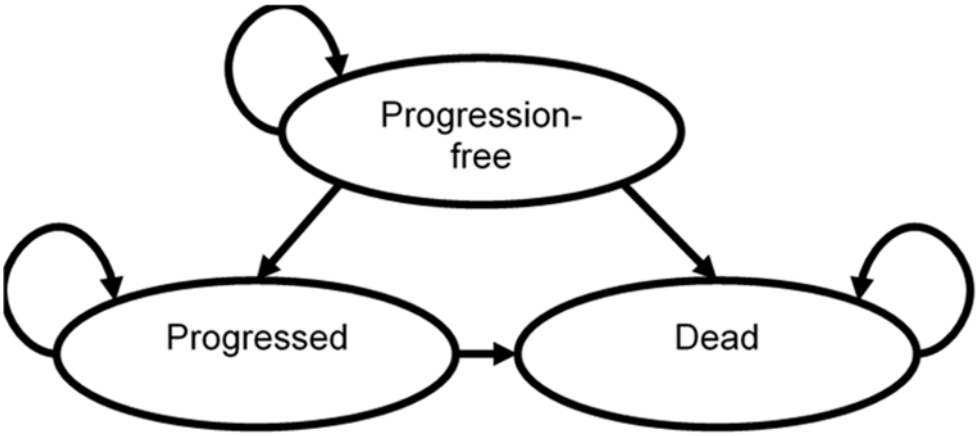
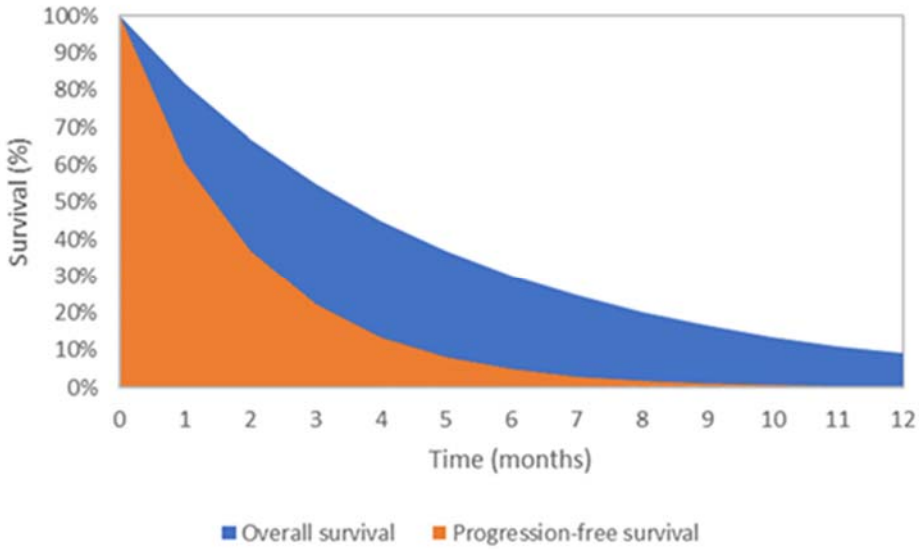



Figure 9. Example of survival partitions used in the model (CS, page 74, Figure 9)




#### *5.4.4.1 ERG critique*

The ERG considers the company's general approach and model structure to be appropriate for the evaluation conducted, and the company described their approach thoroughly and clearly. The electronic model is easy to navigate and is fully functional with an acceptable run time.

However, the ERG identified an error in the company's structure, which meant that there was an inconsistency in the application of the half cycle correction. The company used survival curve data to estimate the half-cycle corrected health state proportions that could be used to apply costs and utilities appropriately. However, the company were inconsistent when applying costs and instead used the survival curves directly, without a half-cycle correction, when applying costs such as post-progression drug costs and mortality costs.

The ERG requested this to be corrected at the clarification stage but the company did not consider it to be an error. However, the company's justification did not address the issue of the half-cycle correction not being fully applied to all costs and QALYs. The ERG, therefore, corrected this error and the results of the corrected company's base case are given in Section 6.1.

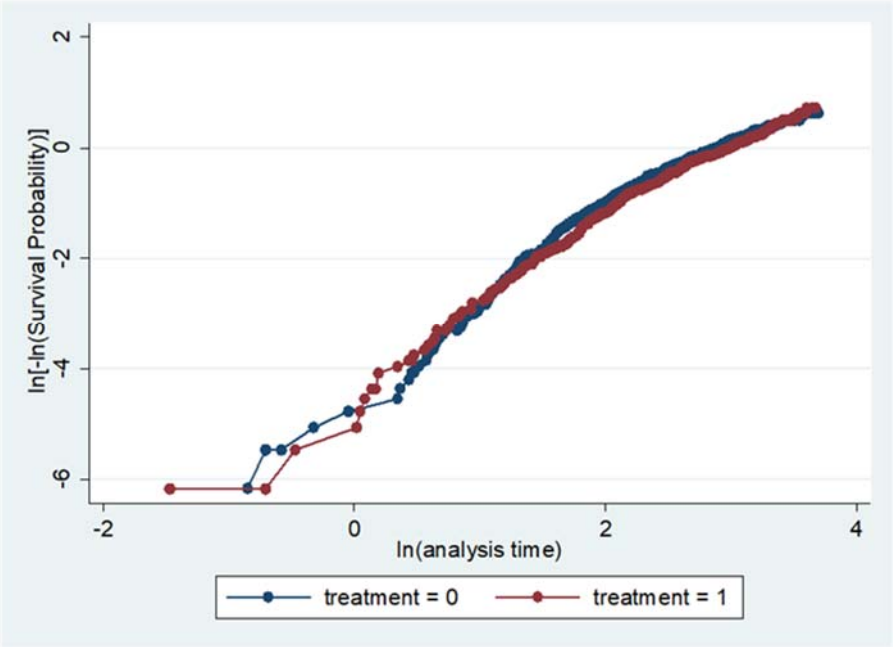
### 5.4.5 Treatment effectiveness

A key aspect of a partitioned survival model is the survival functions required to estimate the proportion of patients in each of the health-states at any given time. PFS and OS data were taken from the REFLECT trial,[58] which were used to fit a range of parametric survival models, so that extrapolated values beyond the trial period could be estimated.

#### *5.4.5.1 Assessment of proportional hazards*

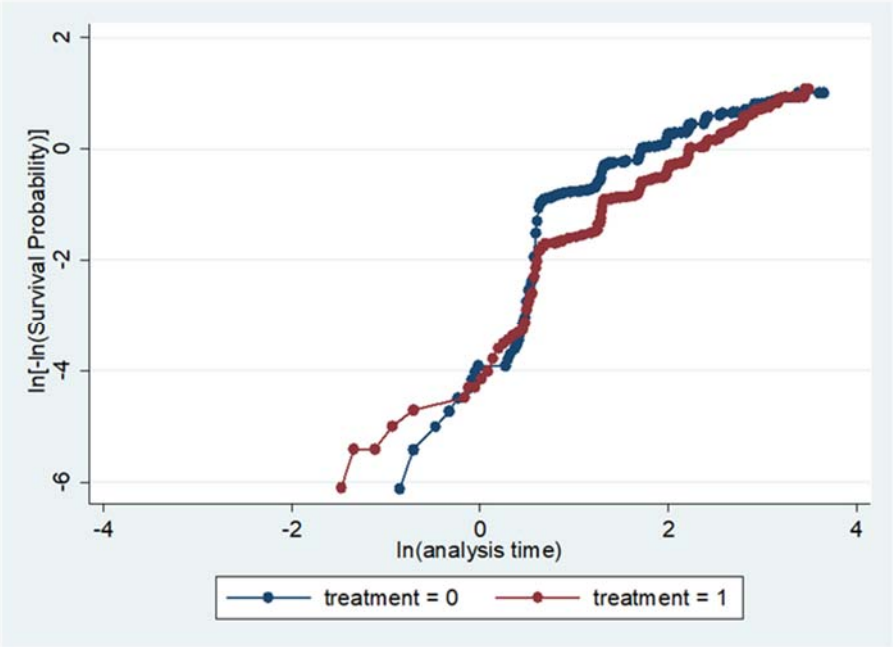
The company firstly assessed whether the data demonstrated proportionality in the hazards for both PFS and OS. This was done by assessing the log-cumulative hazard plots, shown in Figure 10 and Figure 11 for OS and PFS, respectively, as well as an assessment of the global Schoenfeld residual tests, which resulted in a p-values of 0.2902 and <0.0001, respectively.

Figure 10. Log-cumulative hazard plot for OS (CS, page 80, Figure 10)




Treatment = 0 (sorafenib); treatment = 1 (lenvatinib).
Abbreviations: OS, overall survival.

Figure 11. Log-cumulative hazard plot for PFS (CS, page 81, Figure 11)




Treatment = 0 (sorafenib); treatment = 1 (lenvatinib).
Abbreviations: PFS, progression-free survival.

For OS, the company states that the p-value of 0.2902 indicates no statistical deviation from the proportional hazards (PH) assumption, however, they also state that Figure 10 demonstrates log-cumulative hazard functions that are not linear with respect to log-time nor are they parallel, clearly indicated by the crossing of the curves at the beginning. The company conclude that a PH assumption would, therefore, be inappropriate.

The company state that the p-value of <0.0001 for PFS indicates a statistically significant deviation from the PH assumption. The company also state that the log-cumulative hazard plots show a crossing of the curves followed by a convergence, both indicating a lack of proportionality in the hazard functions.

The company, therefore, opted to consider independent statistical models for both PFS and OS, but firstly identified prognostic factors that were imbalanced in the REFLECT trial, which required adjustment in the parametric survival models. The methods adopted for this are described in Section 5.4.5.2.

#### *5.4.5.2 Survival modelling and identification of prognostic factors*

The company fitted multivariable parametric survival models to the PFS and OS data from the intention-to-treat population of the REFLECT trial.[58] For PFS modelling, the investigator assessed PFS outcomes were used.

The company used a range of standard parametric distributions to fit to the data; namely, the exponential, Weibull, Gompertz, log-logistic, lognormal and gamma functions. The most appropriate distribution was selected based on an assessment of the goodness-of-fit statistics, using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), as well as an assessment of the consistency in the extrapolations compared to previous findings in advanced HCC. The company did not assess the curves against the KM plots visually as the parametric survival curves were adjusted for imbalances in the baseline covariates and, therefore, would not be expected to provide a good fit to the 'unadjusted' data.

[redacted]

[redacted]

The multivariable parametric survival models were then fitted, including the application of a covariate adjustment based on the aforementioned set of variables. The outputted survival curves were based on the mean of covariates (MoC) approach, which essentially predicts the survival for the average patient based on the prediction equations derived from the multivariable regression. An alternative to this is the corrected group prognosis approach, which estimate the survival of all patients and calculates the average across the population; however, the company considered the MoC method to be more transparent and easier to incorporate the uncertainty within a probabilistic sensitivity analysis.

The resulting AIC and BIC statistics for multivariable adjusted OS models are given in Table 25 and Table 26 for lenvatinib and sorafenib, respectively. The equivalent values for PFS are given in Table 27 and Table 28, respectively.

Table 25. AIC and BIC: Lenvatinib, OS (CS, page 85, Table 27)

| Model | Obs | df | AIC | BIC | Mean OS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; Obs, observations; OS, overall survival.

Table 26. AIC and BIC: Sorafenib, OS (CS, page 86, Table 28)

| Model | Obs | df | AIC | BIC | Mean OS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; Obs, observations; OS, overall survival.

Table 27. AIC and BIC: Lenvatinib, PFS (CS, page 86, Table 29)

| Model | Obs | df | AIC | BIC | Mean PFS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; Obs, observations; PFS, progression-free survival.

Table 28. AIC and BIC: Sorafenib, PFS (CS, page 87, Table 30)

| Model | Obs | df | AIC | BIC | Mean PFS (months) |
|---|---|---|---|---|---|
| Weibull | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-normal | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Log-logistic | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Exponential | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gamma | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Gompertz | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; Obs, observations; PFS, progression-free survival.

The company determined that the best fitting model for OS was the log-logistic, with the lognormal and the gamma models also performing well. The company also noted that the log-logistic was considered the most suitable model in the original sorafenib appraisal (TA189) and the cancer drugs fund reconsideration (TA474).

For PFS, the company identified the lognormal model as the best fitting for lenvatinib with the gamma and log-logistic models performing similarly. For the sorafenib group, the gamma distribution had the best statistical fit, however, the company considered the extrapolation to be implausible. The next best fitting model, the lognormal, was therefore used in preference for the company's base case analysis.

All fitted survival models for both PFS and OS are given in Figure 12 and Figure 13 for lenvatinib and sorafenib, respectively.

Figure 12. PFS and OS survival models for lenvatinib (CS, page 88, Figure 12)

Abbreviations: OS, overall survival; PFS, progression-free survival.

Figure 13. PFS and OS survival models for sorafenib (CS, page 89, Figure 13)

Abbreviations: OS, overall survival; PFS, progression-free survival.

The company also provided scenario analyses for the alternative distributions, as well as providing unadjusted survival models within the economic model, and survival models adjusted only by AFP and stratification factors.

Following clarification questions, the company also provided additional analyses including the use of the corrected group prognosis method to estimate the survival outputs for the adjusted survival models, as well as fitting survival models specifically for the Western subgroup population, and survival models adjusted for subsequent anti-cancer interventions. These are discussed further within the ERG's critique in Section 5.4.5.3.

#### *5.4.5.3 ERG critique*

The ERG considers the company's general approach used in the modelling of PFS and OS to be reasonably sound with a few limitations that add uncertainty to the results. These key issues are described in the following subsections.

#### 5.4.5.3.1 Covariate adjustment applied to survival models

One key uncertainty with the analysis used in the company's base case relates to the covariate adjustment applied to each of the parametric survival models. The ERG considers the company's initial step, to assess the suitability of the inclusion of variables from a prognostic point of view by eliciting clinical expert opinion, to be appropriate. However, the ERG considers the second stage to assess the statistical suitability for the inclusion of each variable in the regression models to be potentially unreliable.

The company chose to determine the statistical significance of the predictive effect of each of the variables, and therefore the inclusion or exclusion of the variable for adjustment, by using a Cox PH model on the OS data, with an automated backwards stepwise variable selection applied. The variables identified from this model were then all included in the adjustment of the parametric models for both PFS and OS. The ERG considers it preferable to perform variable selection on each of the parametric models when they are fitted, as the underlying model will impact both the coefficients and the p-values, and therefore, could result in a different specification of covariates for adjustment.

Further to this, the adjusted parametric models for PFS are based on covariates not only selected based on a different underlying model, but also with different underlying data given that the Cox PH model was only applied to the OS data. This means the adjustments for the PFS models are potentially more unreliable than the OS models. This is a key uncertainty within the company's economic model, as the post-progression management costs are greater than the pre-progression management costs, so differences in PFS potentially have an important influence on the ICER. The ERG performed a range of scenario analyses to demonstrate the impact of potential differences in PFS on the company's base case ICER and these are reported in Section 6.2.

Another issue with using the Cox PH model is that each variable is assumed to have a relative effect on the hazard ratio. Although the proportionality of hazards between treatment group was tested, the remaining variables considered for inclusion in the multivariable regression model were not tested. The coefficients from this regression may not, therefore, be reliable even if the underlying Cox PH model and the PH assumption for the treatment effect is appropriate. The ERG could not explore this further without the patient level data informing the analysis; however, the ERG considers the method of adjustment used by the company to be preferable to the unadjusted analysis but is uncertain as to whether an alternative adjustment approach would make an important difference to the company's base case ICER. The company's approach does add uncertainty to the conclusions, but given that the data is from an RCT with only some imbalances in covariates, that may give some reassurance that the adjustments made are sufficient. Key imbalances that are likely to make an important difference are the

subsequent anti-cancer therapies received in each treatment group, which impacts the estimates of OS. This is discussed further in Section 5.4.5.3.2.

An additional minor point regarding the covariate adjustment is that the company also provided an analysis using the corrected group prognosis (CGP) approach to estimate adjusted model outputs, as opposed to the MoC method used in the company's base case analysis. The results provided by the company were similar using the two approaches, so the ERG considers the methodological differences in the two approaches to have negligible impact on the final ICER.

#### 5.4.5.3.2 Subsequent treatment adjustment

In response to clarification questions, the company provided an additional adjusted analysis that included a binary variable as an indicator of whether a patient had received any subsequent anti-cancer interventions. The ERG requested an analysis adjusting for these subsequent treatments because of the imbalances across the treatment groups in the REFLECT trial,[58] which potentially influence OS. The REFLECT trial data showed that more subsequent treatments were received in the sorafenib group than the lenvatinib group.[58] This is likely to have resulted in an increased post-sorafenib OS compared to the lenvatinib group. The imbalances were also more extreme in the Western subgroup population meaning the difference in treatment benefit may have been underestimated further within the Western subgroup.

Although the ERG considers the adjustment using a single binary variable to be somewhat crude, the ERG considers the company's adjustment to be preferable to the company's base case analysis. While the company's base case analysis does adjust for imbalances in some of the covariates, it does not adjust for subsequent anti-cancer interventions. Adjusting the OS models for subsequent anti-cancer interventions avoids the need to offset the potential benefits incurred by these treatments by applying the respective costs. This is particularly important given that no subsequent treatments are recommended beyond second-line in the UK. Removing subsequent treatment costs and adjusting the OS model appropriately also provides an economic evaluation that is more in line with the treatment pathway in the UK, as described in Section 2.2. This approach was, therefore, incorporated into the ERG's preferred base case analysis, as outlined in Section 6.3.

#### 5.4.5.3.3 Western subgroup population survival models

At the clarification stage, the ERG requested additional analyses based on the Western subgroup population as it considered that the full population of the REFLECT trial may not be fully reflective of the UK population. The company responded to the ERG's request with adjusted parametric survival models based on the Western subgroup. The analyses on the Western subgroup are based on a reduced population size (from 954 to 314 people) and, therefore, make the results less robust. The stratified Cox PH model results demonstrate that the survival in the lenvatinib group was worse than in the sorafenib

group (hazard ratio [HR] 1.08, 95% CI 0.82 to 1.42); however, the result was not significant and was potentially influenced by imbalances in subsequent anti-cancer therapies that were more apparent in the Western subgroup in favour of the sorafenib group than in the full population. When adjusting for these subsequent treatments, the results were similar in the Western subgroup and the full population. The ERG, therefore, considers the analysis based on the full population to be appropriate and that the efficacy in the full population of the REFLECT trial is likely to be similar to that expected in a UK population. This is discussed in more detail in Sections 3 and 4.

#### 5.4.5.3.4 Censoring in PFS data

For the company's analysis of PFS that was used to inform the company's base case in the economic model, data were censored for patients who had not experienced disease progression at the point of treatment discontinuation, even though these patients were still followed up further, beyond the point where their disease had progressed.

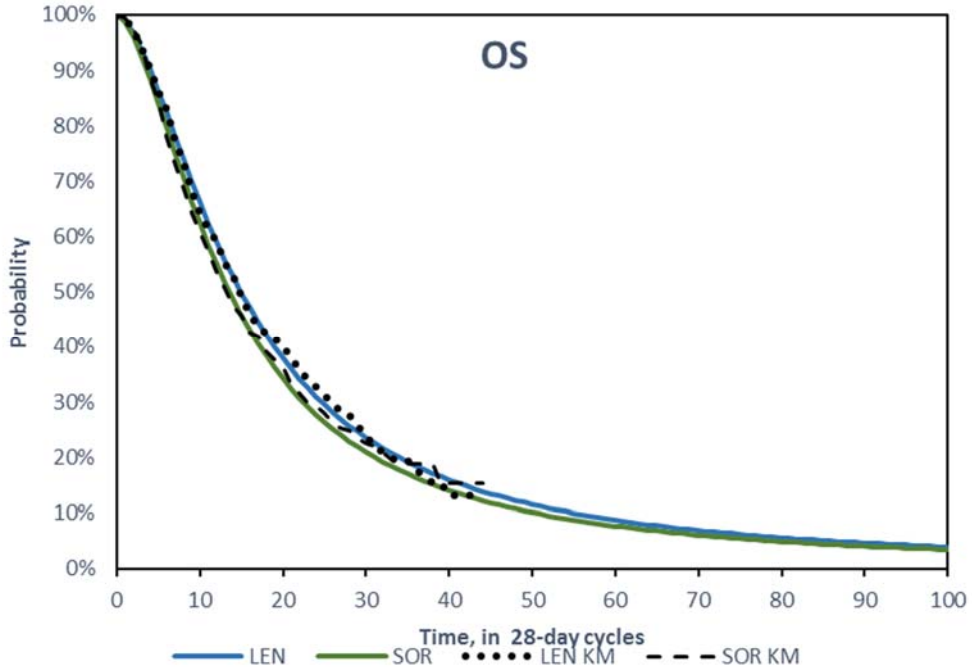
The consequence of this is that the PFS modelling used in the company's base case analysis was potentially overestimated, as indicated by the increased HR from 0.66 to [redacted] in the company's PFS scenario analysis, which included all progressed disease events. However, the ERG considers both of these analyses to be unreliable given that the PH assumption was shown by the company to be implausible. This issue is discussed in more detail in Section 4.2.3.

In the absence of any modelling using these data, the ERG explored a range of scenario analyses to diminish the treatment effect for PFS so as to assess the potential impact of this censoring on the company's base case ICER. To do this, the ERG varied the parameters in the company's best fitting PFS survival model. This is described in more detail along with the results of the scenario analyses presented in Section 6.2.

#### 5.4.5.3.5 Parametric model goodness-of-fit

For OS, the company chose the log-logistic model, which had the least AIC and BIC statistics for both the lenvatinib and sorafenib groups, and the ERG considers the extrapolations to be reasonable. The ERG sought clinical expert opinion to elicit a plausible maximum for the survival of a patient with untreated, advanced or unresectable HCC. A value of 10 years was considered to be plausible and the company's extrapolated curves are roughly in line with this, with around 1-2% of patients alive in each group at 10 years. These models were, therefore, kept as part of the ERG's preferred base case and are shown in Figure 14 along with the KM plot. A scenario analysis was undertaken using a similarly well fitted curve but with a slightly more conservative extrapolation; the generalised gamma curve. The results of this analysis are given in Section 6.2 and 6.4 for the company's base case and the ERG's base case, respectively.

Figure 14. Sorafenib OS (lognormal)




However, for PFS, the chosen lognormal models had the least AIC and BIC statistics for lenvatinib, with the log-logistic and gamma models performing similarly, but for sorafenib, the lognormal was not a good fit. The ERG's agrees with the use of the same functional form across treatment groups as advised in NICE TSD 14;[59] however, a reasonable fit should be considered in both groups. The gamma model was by far the best fit in terms of AIC and BIC as demonstrated in Table 28, and given that the gamma model was a reasonably good fit for lenvatinib, the ERG considers the gamma model to be the most suitable to estimate and extrapolate PFS for both treatment groups. This has been incorporated into the ERG's preferred base case, with the addition of the company's option within the model to prevent the fitted curves from crossing over, which happens with the PFS gamma models. The company's unadjusted lognormal models, which the company's base case analysis in the economic model was based on, are shown in Figure 15, and the ERG's preferred unadjusted gamma models are shown in Figure 16.

The ERG notes that the visual fit of the adjusted models could not be performed relative to the 'unadjusted' KM data. However, the ERG considers that the company should have assessed the unadjusted models against the KM data, as this would demonstrate that the underlying functional form may be considered appropriate. The adjustments to the model can of course affect the goodness-of-fit; however, if the unadjusted model does not fit the KM data then the functional form is unlikely to be suitable for the adjusted model either.

Figure 15. Lognormal PFS models used in company's base case (unadjusted)

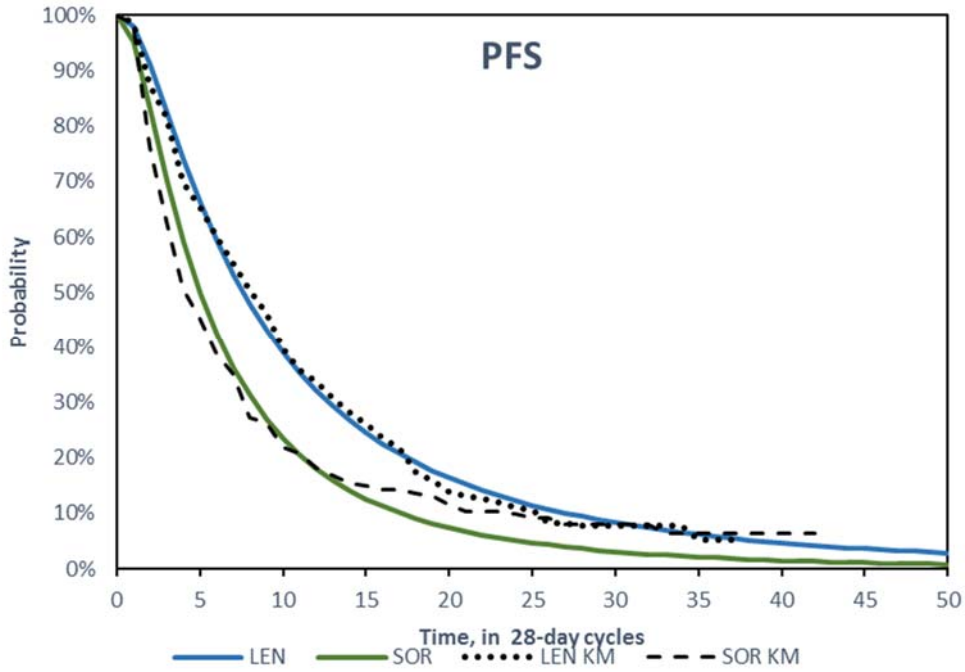
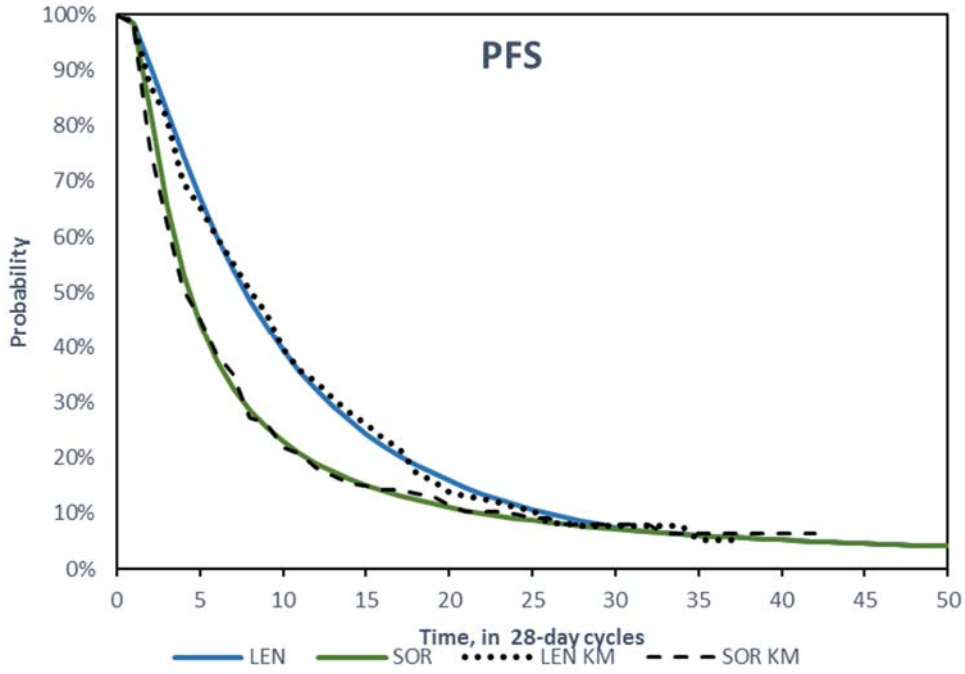



Figure 16. Gamma PFS models used in ERG's preferred base case (unadjusted)




#### 5.4.5.3.6 Overall summary of changes in ERG preferred base case

Overall, the ERG considered the analysis based on the full trial population to be appropriate and so no change was made regarding subgroup analyses. Although the covariate adjustment was considered to be potentially unreliable, the ERG considers it unlikely to make an important impact in comparison to the adjustments made directly to the parametric survival models when fitted. The ERG's preference is

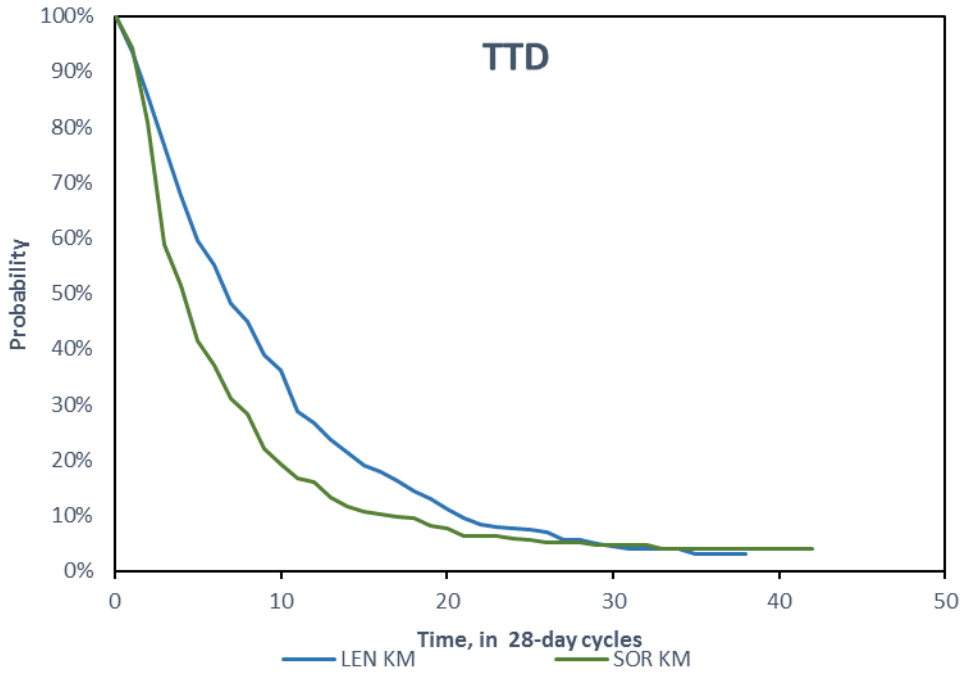
to include the adjustment to OS for subsequent anti-cancer interventions and so the adjusted analyses with this covariate included was included in the ERG's preferred base case.

The ERG considers there to be a potential overestimation of the treatment effect in the company's use of the PFS analysis with censoring for treatment discontinuation in patients without disease progression; however, the company have not provided survival models using the alternative data without this censoring. The ERG, therefore, could not perform further analyses to correct this issue, but have provided scenario analyses to assess the potential impact. One key change the ERG has made to the PFS modelling, is to change the functional form of the model to the gamma distribution, which the ERG considers to provide a better fit to the data.

### 5.4.6 Treatment discontinuation

The company did not conduct any modelling of time to treatment discontinuation (TTD) as the data in the REFLECT trial were almost complete except for 4% of patients remaining on treatment in the sorafenib group.[58] The company used the KM data directly in the model to estimate the proportion of patients on treatment at any particular time, and assumed that the 4% remaining on treatment in the sorafenib group all discontinued at the end of the trial follow-up period. The TTD KM plots from the REFLECT trial are given in Figure 17.

Figure 17. TTD KM plots from REFLECT



#### 5.4.6.1 ERG critique

The ERG considers the company's approach to estimating primary treatment costs using the data directly from the trial to be reasonable given the completeness of the data available.

### 5.4.7 Adverse events

In the economic model, the company modelled the impact of all grade 3–4 treatment-emergent adverse events (AEs) with an incidence of ≥5% in either arm of the REFLECT trial.[58] Diarrhoea, asthenia and fatigue occurred in less than 5% of patients, but were also included as the company's clinical experts expected those AEs to have significant clinical or economic impacts. The AE rates observed in the trial for the full population and applied in the model are presented in Table 29.

Table 29. Adverse events included in the economic model (adapted from Table 31 of the CS)

| Adverse event | Number of patients, n (%) | | Number of events per patient with ≥1 event | |
|---|---|---|---|---|
| | Lenvatinib | Sorafenib | Lenvatinib | Sorafenib |
| Hypertension | 111 (23.3%) | 68 (14.3%) | 1.12 | 1.09 |
| Weight decreased | 36 (7.6%) | 14 (2.9%) | 1.03 | 1.00 |
| Blood bilirubin increased | 31 (6.5%) | 23 (4.8%) | 1.10 | 1.04 |
| Proteinuria | 27 (5.7%) | 8 (1.7%) | 1.04 | 1.00 |
| Gamma-glutamyltransferase increased | 26 (5.5%) | 19 (4.0%) | 1.04 | 1.05 |
| Platelet count decreased | 26 (5.5%) | 16 (3.4%) | 1.12 | 1.00 |
| Aspartate aminotransferase increased | 24 (5.0%) | 38 (8.0%) | 1.04 | 1.11 |
| Diarrhoea | 20 (4.2%) | 20 (4.2%) | 1.20 | 1.20 |
| Fatigue† | 18 (3.8%) | 17 (3.6%) | 1.07 | 1.00 |
| Palmar-plantar erthrodysaesthesia syndrome | 14 (2.9%) | 54 (11.4%) | 1.07 | 1.17 |
| Asthenia | 14 (2.9%) | 11 (2.3%) | 1.07 | 1.00 |
| Abbreviations: n, number of patients. | | | | |

The impact of AEs on patients' quality of life is described in Section 5.4.8. As for the costs of managing AEs, the total cost is applied upfront in the first model cycle, based on the mean number of episodes per patient, as explained in Section 5.4.9.4.

#### 5.4.7.1 ERG critique

The ERG considers the company's estimation of AEs to be reasonable and the ERG's clinical experts confirmed that all the relevant AEs associated with lenvatinib and sorafenib were included in the economic analysis.

However, the ERG was unclear as to why the potential consequences of grade 5 AEs were not considered in the company's economic analysis. In response to a clarification question, the company

explained that grade 5 AEs were not included in their initial analysis to avoid double counting the costs of mortality. However, to address the ERG's concern that the potentially high cost of treating grade 5 AEs may have an impact on the company's base case ICER, the company explored a scenario which applied the cost of one hospitalisation to 12.8% and 7.6% of lenvatinib and sorafenib patients in the full population experiencing each grade 5 AE, respectively. However, the impact of including the costs to manage grade 5 AEs in the model was negligible with the absolute difference in costs reducing from [redacted] to [redacted].

During the clarification stage, the company also provided equivalent AE data for the Western subgroup that were considered in the economic model for the full population (Section 4.3.7). Despite these additional data, the company performed scenario analysis on the Western subgroup using AE data on the full population, which the ERG considers to be inconsistent. For completeness, the ERG explored a scenario where results for the Western subgroup were informed by AE data from the Western subgroup (Section 6.2). Nonetheless, the impact on the results was minimal with the absolute difference in costs reducing from [redacted] to [redacted].

Finally, the company did not include AEs associated with subsequent anti-cancer interventions in their analysis. However, as AEs for the primary therapies are not a key driver in the model, shown by the ERG's analysis which explores the impact of removing AE (Section 6.2), the ERG considers this to be a reasonable simplification. The number of people receiving subsequent anti-cancer interventions is only a subset of those receiving the primary therapies and, therefore, the impact of subsequent intervention AEs is unlikely to have a meaningful impact on the ICER.

## 5.4.8 Health-related quality of life

### *5.4.8.1 Health-related quality of life data obtained from the clinical trial*

During the REFLECT trial, patients completed the EQ-5D-3L questionnaire at the baseline visit, on day 1 of each 28-day treatment cycle, and at the off-treatment visit, which occurred within 30 days after the final dose of the study treatment. The EQ-5D data collected in REFLECT at each time point is provided in Appendix L of the CS.[58]

Using those data, the company generated mean utility values at baseline, in the progression-free health state and in the progressed health state (Table 30). The company also generated adjusted mean utility values, using a linear mixed model, controlling for prior treatment, age, sex, geographical region, baseline EQ-5D score and baseline ECOG-PS.

Table 30. Summary of EQ-5D data (reproduced from Table 32 of the CS)

| | | Full REFLECT population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| Baseline | N | 921 | 463 | 458 |
| | Mean (SE) | 0.829 (0.0067) | 0.823 (0.0101) | 0.836 (0.0088) |
| | Adjusted mean (SE) | N/A | 0.784 (0.0113) | 0.798 (0.0113) |
| Progression-free | N | 852 | 421 | 431 |
| | Mean (SE) | 0.745 (0.0079) | 0.750 (0.0105) | 0.740 (0.0118) |
| | Adjusted mean (SE) | N/A | 0.745 (0.0116) | 0.737 (0.0115) |
| Progressed | N | 755 | 373 | 382 |
| | Mean (SE) | 0.678 (0.0118) | 0.678 (0.0163) | 0.679 (0.0170) |
| | Adjusted mean (SE) | N/A | 0.665 (0.0188) | 0.656 (0.0185) |
| Abbreviations: CS, company submission; EQ-5D, EuroQol 5 dimensions; N, number of patients; SE, standard error | | | | |

Given that adjusted mean utility values for the progression-free and progressed health states were similar between the lenvatinib and sorafenib arms, with a small difference in favour of lenvatinib, the company conservatively assumed that utility values in the lenvatinib and sorafenib arms are equal to the mean values in the full REFLECT population.[58]

The company also concluded that as the treatment-specific utility values were similar, the impact of AEs on quality of life would also be similar. For this reason, the company did not explicitly model disutilities associated with AEs.

Although the EQ-5D-3L from REFLECT was considered to be the most appropriate source of HRQoL data, alternative sources including the recent NICE TAs for regorafenib (ID991) and sorafenib (TA474) were explored in scenario analysis.[24, 33] The results of those analysis are presented in Section 5.5.

#### *5.4.8.2 ERG critique*

The company measured changes in HRQoL for patients receiving lenvatinib and sorafenib directly from patients in the REFLECT trial using a generic preference-based measure, the EQ-5D, following the key components of the NICE reference case.[60] Nonetheless, the ERG would like to highlight that the size of the full population reduced from 463 and 458 for lenvatinib and sorafenib, respectively, at baseline, to [redacted] in both treatment arms from cycle 26 to 43. This notable reduction in patient numbers may increase the risk of attrition bias if sicker patients were lost to follow-up earlier and not included in the analysis. However, the ERG does not have patient level data to assess this further.

Furthermore, no formal comparison was made between the HRQoL values observed in the trial and data identified in the SLR, although the ERG welcomed the alternative sources explored by the company in scenario analysis.

As described in Section 5.4.5, the company provided an additional scenario analysis based on the Western subgroup during the clarification stage to address the ERG's concerns that there may be a difference in efficacy in the Western subgroup in addition to differences in treatment protocols and patient characteristics. Within that analysis the company applied pre- and post- progression HSUVs obtained from the Western subgroup to reflect the (lower) quality of life in the Western subgroup (Table 31). During validation of the revised model, the ERG found that full population utility values were implemented in the submitted model in error, although the correct results were reported by the company. This was subsequently corrected in the model by the ERG to ensure any subsequent analyses on the Western subgroup were accurate.

Table 31. Utility analysis based on the Western subgroup (reproduced from table 19 of the company's clarification responses)

| | | Full Western subgroup | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| Baseline | N | 285 | 144 | 141 |
| | Mean (SE) | 0.777 (0.0122) | 0.773 (0.0173) | 0.781 (0.0173) |
| | Adjusted mean (SE) | N/A | 0.756 (0.0193) | 0.766 (0.0198) |
| Progression-free | N | 268 | 129 | 139 |
| | Mean (SE) | 0.693 (0.0134) | 0.697 (0.0176) | 0.689 (0.0199) |
| | Adjusted mean (SE) | N/A | 0.690 (0.0176) | 0.691 (0.0174) |
| Progressed | N | 229 | 116 | 113 |
| | Mean (SE) | 0.633 (0.0189) | 0.652 (0.0251) | 0.614 (0.0285) |
| | Adjusted mean (SE) | N/A | 0.656 (0.0272) | 0.613 (0.0275) |

Abbreviations: N/A, not applicable; SE, standard error.

Overall, the ERG has two main issues regarding the company's approach to model HRQoL including: the linear mixed model and the quality of life of progression-free patients according to treatment status. Each of these is discussed in turn in the following subsections.

#### *5.4.8.3 Linear mixed model*

While the ERG considers it appropriate to perform a regression analysis to identify covariates that influence quality of life, the methods used by the company were not substantiated in the CS. During the clarification stage, the ERG asked the company to clarify how covariates for the linear mixed model were chosen and to provide further details on the steps taken to adjust mean utility values. Following this, the company explained that covariates for the linear mixed model were non-systematically pre-specified prior to commencement of the analysis of the REFLECT trial.

In light of the company's response, the ERG notes that the adjustment could have been improved if potentially influential factors identified following the commencement of the trial such as AFP levels and HCC aetiology were included. Furthermore, based on the company's response, it appears that an automated variable selection method (such as stepwise selection) was not considered, as all covariates (including all non-significant covariates) were retained in the linear mixed model (Table 20 of the

company's clarification responses). Despite this, non-significant covariates had relatively small coefficients which reduces the impact of overfitting due to random variation. Even so, the ERG considers the company's adjusted analysis to be potentially unreliable if influential factors are missing.

#### *5.4.8.4 Utility analysis by treatment status*

The company assumed AE disutilities were captured in the trial-derived EQ-5D utilities which the ERG considers to be a reasonable assumption, as incorporating additional disutilities could lead to double counting if the HRQoL impact of AEs has already been captured by the HSUVs. However, given the difference between TTD and PFS, the ERG requested the company to provide adjusted and unadjusted treatment-specific utilities for the on- and off-treatment periods for those who are progression-free. Following this request, the company provided results for both the full population (Table 32) and Western subgroup (Table 33) in their response.

Table 32. Utility analysis based on treatment status (overall population) (reproduced from Table 17 of the company's clarification responses)

| | Statistic | Total study population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| Baseline | n | 885 | 443 | 442 |
| | Mean (SE) | 0.833 (0.0067) | 0.830 (0.0100) | 0.837 (0.0090) |
| | Adj Mean (SE)[3] | N/A | 0.789 (0.0114) | 0.796 (0.0115) |
| Progression-Free Survival (On Treatment)[1] | n | 846 | 417 | 429 |
| | Mean (SE) | 0.747 (0.0079) | 0.752 (0.0105) | 0.742 (0.0117) |
| | Adj Mean (SE)[3] | N/A | 0.748 (0.0116) | 0.739 (0.0114) |
| Progression-Free Survival (Off Treatment)[2] | n | 53 | 27 | 26 |
| | Mean (SE) | 0.689 (0.0430) | 0.713 (0.0544) | 0.664 (0.0677) |
| | Adj Mean (SE)[3] | N/A | 0.700 (0.0888) | 0.694 (0.1100) |

[1] The average of all post-baseline pre-progression on-treatment EQ-5D HUI scores among all patients progression free.
[2] The average of all post-baseline pre-progression off-treatment (after last dose of study medication) EQ-5D HUI scores among all patients progression free.
[3] Least-squared means adjusted for prior treatment, age, sex, geographical region, baseline EQ-5D HUI score, and baseline ECOG-PS.
Program: (t_ea_hui_MM_NICE.sas) (22MAR18:12:34:20)
Analysis datasets: adqs, adttdef
Abbreviations: Adj, Adjusted; ECOG PS, Eastern Cooperative Oncology Group Performance Status; EQ-5D, EuroQol Five Dimension Health Survey; SE, Standard Error; N/A, Not Applicable; HUI, Health Utility Index.

Table 33. Utility analysis based on treatment status (Western subgroup) (reproduced from Table 18 of the company's clarification responses)

| | Statistic | Western population | Lenvatinib | Sorafenib |
|---|---|---|---|---|
| Baseline | n | 273 | 138 | 135 |
| | Mean (SE) | 0.780 (0.0125) | 0.778 (0.0176) | 0.781 (0.0179) |
| | Adj Mean (SE)[3] | N/A | 0.757 (0.0199) | 0.763 (0.0203) |
| Progression-Free Survival (On Treatment)[1] | n | 266 | 127 | 139 |
| | Mean (SE) | 0.695 (0.0134) | 0.701 (0.0177) | 0.689 (0.0200) |
| | Adj Mean (SE)[3] | N/A | 0.697 (0.0177) | 0.693 (0.0174) |
| Progression-Free Survival (Off Treatment)[2] | n | 19 | 11 | 8 |
| | Mean (SE) | 0.714 (0.0471) | 0.654 (0.0667) | 0.797 (0.0561) |
| | Adj Mean (SE)[3] | N/A | 0.606 (0.0606) | 0.757 (0.0938) |

[1] The average of all post-baseline pre-progression on-treatment EQ-5D HUI scores among all patients progression free.
[2] The average of all post-baseline pre-progression off-treatment (after last dose of study medication) EQ-5D HUI scores among all patients progression free.
[3] Least-squared means adjusted for prior treatment, age, sex, geographical region, baseline EQ-5D HUI score, and baseline ECOG-PS.
Program: (t_ea_hui_MM_NICE.sas) (22MAR18:12:34:20)
Analysis datasets: adqs, adttdef
Abbreviations: Adj, Adjusted; ECOG PS, Eastern Cooperative Oncology Group Performance Status; EQ-5D, EuroQol Five Dimension Health Survey; SE, Standard Error; N/A, Not Applicable.

For patients who are progression-free in the full population, on-treatment utility values were higher than off-treatment utility values in both treatment arms. The same was also true for the lenvatinib arm in the Western subgroup. Conversely, a higher utility off-treatment than on-treatment was demonstrated in the sorafenib arm in the Western subgroup.

One reason for the latter observation in the Western subgroup may be the sparsity of data contributing to the result, but other reasons could include differences in sorafenib tolerability or access to subsequent treatments following discontinuation of sorafenib, both of which have been highlighted by the company when comparing regions in other contexts. Nonetheless, as the difference between lenvatinib and sorafenib was not found to be statistically significant in each analysis, the company applied the average utility values across the two treatment arms for each of the on-treatment and off-treatment states, and included this analysis as an option in the revised model provided at clarification.

Applying utilities for the on- and off-treatment periods for those who are progression-free in the full population reduced the total QALYs in both treatment arms, but more so for lenvatinib, leading to a reduction in incremental QALYs from [redacted] to [redacted]. As for the Western subgroup, a result was not provided by the company in their response. For completeness, this scenario was explored by the ERG but the effect of treatment status on the results was found to be minimal with an increase in incremental QALYs from [redacted] to [redacted].

## 5.4.9 Resources and costs

The cost components included in the economic model are listed below and discussed in detail in the following subsections:

- Costs associated with the intervention and comparator (Section 5.4.9.1);
- Costs associated with subsequent anti-cancer medication (Section 5.4.9.2);
- Disease management costs (Section 5.4.9.3);
- Costs of managing AEs (Section 5.4.9.4);
- End-of-life costs (Section 5.4.9.5).

#### *5.4.9.1 Costs associated with the intervention and comparator*

Drug acquisition costs used in the model for lenvatinib and sorafenib are presented in Table 34. For lenvatinib, a simple patient access scheme (PAS) discount of [redacted] has been agreed with the Department of Health, and this was applied to the list price in the company's analysis. As for sorafenib, only the list price is presented, although details of the approved PAS, along with results of the economic analysis incorporating this PAS, can be found in the confidential appendix. For both treatments, the daily dose was modified in REFLECT, as needed, according to the toxicity management plan. Moreover, once a patient's dose was reduced for either treatment, it could not be increased. Consequently, the company modelled the mean doses received in REFLECT rather than the recommended daily doses stated previously in Section 5.4.3. Daily doses based on the full population used to inform the company's base case analysis are given in Table 34.

Table 34. Drug aquisition costs applied in the economic model

| Drug | Units per pack | mg per unit | List cost per pack[a] | PAS discount | PAS cost per pack | Cost per mg | mg per day b | Cost per day | Drug cost per 28-day cycle |
|---|---|---|---|---|---|---|---|---|---|
| Lenvatinib | 30 | 4 | £1,437 | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] |
| Sorafenib | 112 | 200 | £3,577 | 0% | £3,577 | £0.16 | 663.8 | £106 | £2,968 |

Abbreviations: PAS, patient access scheme
[a] Drug cost information obtained from BNF[61]
[b] mean dose intensity based on the full population safety analysis set (Table 19 of the CS)

The company assumed that both lenvatinib and sorafenib would not be associated with administration costs as patients would orally self-administer both treatments. The company also assumed that monitoring and testing requirements for both treatments would be the same based on clinical advice received and the independent assessment presented in the Multiple Technologies Appraisal (MTA) of lenvatinib and sorafenib in the treatment of thyroid cancer (GID-TA10101).[62] Finally, in the base case analysis, it was assumed that there were no costs of wastage incurred, but a scenario analysis was performed in which discontinuation of treatment was associated with wastage of 7 days' worth of drug costs.

Based on the assumptions outlined above, the final cost per 28-day cycle was [redacted] for lenvatinib, with the application of the agreed PAS, and for sorafenib, the cost was £2,968. In order to apply drug acquisition costs for lenvatinib and sorafenib in the model, the company used TTD data from the REFLECT trial to determine the proportion of patients on treatment in each cycle of the model. The company's approach for applying TTD in the model is outlined and critiqued in Section 5.4.6.

#### *5.4.9.2 Costs associated with subsequent anti-cancer medication*

After discontinuing therapy in REFLECT, patients in the sorafenib arm were permitted to continue sorafenib as a subsequent anti-cancer medication, and lenvatinib patients were permitted to switch to sorafenib. Patients in the sorafenib arm were also eligible for second-line trials (including regorafenib). Some patients received other subsequent anti-cancer medications and procedures in REFLECT; however, only sorafenib and regorafenib are considered in the model as they are the only medications licensed in the UK for this indication.

Drug acquisition costs for the subsequent medications applied in the model for the lenvatinib and sorafenib groups are presented in Table 35, which are based on list prices. However, both medications are associated with confidential PASs. Using data on the full population in REFLECT, the company calculated the one-off acquisition cost of subsequent anti-cancer medication. Other costs associated with subsequent anti-cancer medication (i.e. administration, monitoring and testing) were not applied in the model or discussed in the CS.

Table 35. Costs of subsequent anti-cancer medication applied in the economic model (adapted from Table 41 and 42 of the CS)

| Subsequent anti-cancer medication | Units per pack | mg per unit | Cost per pack [a] | Cost per mg | mg per day | Cost per day | % of progressed patients [d] | | Mean duration (days) | | Weighted duration (days) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | LEN | SOR | LEN | SOR | LEN | SOR |
| Sorafenib | 112 | 200 | £3,577 | £0.16 | 664[b] | £106 | [redacted] | [redacted] | [redacted] | [redacted] | 33.28 | 11.45 |
| Regorafenib | 84 | 40 | £3,744 | £1.11 | 144[c] | £160 | [redacted] | [redacted] | [redacted] | [redacted] | 0.56 | 4.28 |
| Total cost [e] | - | - | - | - | - | - | - | - | - | - | **£3,617** | **£1,900** |

Abbreviations: LEN, lenvatinib; SOR, sorafenib

[a] drug cost information obtained from BNF[61]; [b] assumed to be the same as first line; [c] mean dose taken from the RESORCE trial; [d] calculated as the number of patients using each medication divided by the number of patients experiencing either progression or death from the progression-free state; [e] weighted average cost is applied to all individuals leaving the PFS state in the cycle in which this occurs.

#### *5.4.9.3 Disease management costs*

Estimates of resource use in the progression-free and progressed health states were based on two surveys, commissioned by the marketing authorisation holder for sorafenib. The initial survey, presented in TA189, reported the proportion of patients expected to incur physician visits, laboratory tests, radiological tests, hospitalisations, hospital follow-ups and social care, and the frequency those resources are received each month.[34] The updated survey, presented in TA474, only reported the costs

associated with physician visits, laboratory tests, radiological tests and hospitalisations.[24] The results of those two surveys are summarised in Table 36 and Table 37.

To reflect Appraisal Committee preference during the 2016 reconsideration of sorafenib, the company pooled the two surveys. Due to differences in data reported in TA189 and TA474, new unit costs for the initial survey (provided in Table 37 of the CS) were taken from NHS reference costs 2016–2017 and the Personal Social Services Research Unit (PSSRU) 2017.[63, 64] For the updated survey, the 2015–2016 costs presented were inflated to 2016–2017 costs. Finally, a weighted average cost based on the number of clinicians responding to each survey was used to inform the economic model (Table 38).

Table 36. Disease management costs (taken from the economic model)

| Cost Item | | Unit cost | Progression-free | | | | Progressed | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Initial survey (TA189)[34] | | | Updated survey (TA474)[24] | Initial survey (TA189)[34] | | | Updated survey (TA474)[24] |
| | | | % of patients | Frequency per month | Frequency per cycle | | % of patients | Frequency per month | Frequency per cycle | |
| Physician visits | Oncologist | £173 | 100% | 0.75 | 0.69 | - | 100% | 0.38 | 0.35 | - |
| | Hepatologist | £217 | 100% | 0.17 | 0.16 | - | 100% | 0.50 | 0.46 | - |
| | Macmillan nurse | £42 | 100% | 0.50 | 0.46 | - | 100% | 1.00 | 0.92 | - |
| | Gastroenterologist | £141 | 100% | 0.08 | 0.07 | - | 100% | 0.00 | 0.00 | - |
| | Radiologist | £74 | 100% | 0.08 | 0.07 | - | 100% | 0.00 | 0.00 | - |
| | Clinical nurse specialist | £42 | 100% | 0.50 | 0.46 | - | 100% | 0.25 | 0.23 | - |
| | Palliative care physician/ nurse | £42 | 100% | 0.13 | 0.12 | - | 100% | 0.75 | 0.69 | - |
| | **Total costs** | **-** | **-** | **-** | **£212** | **£89** | - | - | **£237** | **£581** |
| Laboratory tests | AFP test | £25 | 75% | 0.83 | 0.76 | - | 38% | 1.00 | 0.92 | - |
| | Liver function test | £25 | 50% | 0.67 | 0.62 | - | 25% | 1.00 | 0.92 | - |
| | INR | £3 | 50% | 0.67 | 0.62 | - | 0% | 0.00 | 0.00 | - |
| | Complete blood count | £3 | 75% | 1.00 | 0.92 | - | 50% | 1.00 | 0.92 | - |
| | Biochemistry | £1 | 50% | 1.00 | 0.92 | - | 25% | 1.00 | 0.92 | - |
| | Endoscopy | £500 | 25% | 0.33 | 0.30 | - | 0% | 0.00 | 0.00 | - |
| | **Total costs** | **-** | **-** | **-** | **£63** | **£293** | **-** | **-** | **£16** | **£295** |
| Radiological tests | CT scan (abdominal) | £102 | 73% | 0.33 | 0.30 | - | 73% | 0.39 | 0.36 | - |
| | MRI scan (abdominal) | £145 | 28% | 0.33 | 0.30 | - | 28% | 0.50 | 0.46 | - |
| | **Total costs** | **-** | **-** | **-** | **£35** | **£24** | **-** | **-** | **£45** | **£3** |
| Hospitalisation | Hospitalisation | £1,924 | 46% | 0.16 | 0.15 | - | 48% | 0.40 | 0.37 | - |
| | **Total costs** | **-** | **-** | **-** | **£130** | **£40** | **-** | **-** | **£340** | **£6** |
| Hospital follow-up | Specialist | £217 | 100% | 0.25 | 0.23 | - | 100% | 3.00 | 2.76 | - |
| | GP | £37 | 100% | 1.50 | 1.38 | - | 100% | 1.50 | 1.38 | - |
| | Nurse | £42 | 100% | 1.75 | 1.61 | - | 100% | 2.00 | 1.84 | - |
| | **Total costs** | **-** | **-** | **-** | **£169** | **NR** | **-** | **-** | **£726** | **NR** |

Abbreviations: AFP, Alpha-fetoprotein; CT, computerised tomography; GP, general practitioner; INR, international normalized ratio; MRI, magnetic resonance imaging; NR, not reported

Table 37. Social care costs presented in the initial survey (TA189) (taken from the economic model)

| Resource | Cost per day | Cost source | Progression-free | | | | Progressed | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | % of patients | Days per month | Days per cycle | % funded by the NHS | % of patients | Days per month | Days per cycle | % funded by the NHS |
| Residential care | £111 | PSSRU 2017. Average of two sources: private sector residential care for older people (£632 per week = £90.29 per day) and private sector residential care homes for adults requiring physical support (£131 per day).[64] | 2% | 0.00 | 0.00 | 0% | 3% | 6.43 | 5.92 | 100% |
| Day care | £91 | PSSRU 2017. Day care for adults requiring physical support.[64] | 2% | 0.00 | 0.00 | 0% | 23% | 5.36 | 4.93 | 0% |
| Home care | £165 | PSSRU 2017. Home care for adults requiring physical support. Average of two hourly rates: £25.62 (services provided in-house) and £15.52 (provision by external providers). Assumes 8 hours per day.[64] | 7% | 4.00 | 3.68 | 50% | 28% | 12.86 | 11.83 | 100% |
| Hospice | £503 | Marie Curie Cancer Care. Understanding the cost of end of life care in different settings.[65] Inflated from 2003/04 to 2016/17 using HCHS Pay and Prices Index. | 0% | 0.00 | 0.00 | 0% | 18% | 14.00 | 12.88 | 43% |
| **Total costs** | - | | **-** | **-** | **-** | **£21** | **-** | **-** | **-** | **£1,066** |

Abbreviations: HCHS, hospital and community health service; PSSRU, personal social services research unit

Table 38. Per-cycle health state costs (reproduced from Table 38 of the CS)

| Resource | Progression-free | Progressed |
|---|---|---|
| Physician visits | £159.63 | £384.40 |
| Laboratory tests | £161.78 | £135.56 |
| Radiological tests | £30.04 | £27.25 |
| Hospitalisation | £91.52 | £196.78 |
| Hospital follow-up* | £168.50 | £726.26 |
| Social care* | £21.19 | £1,066.07 |
| **Total** | **£632.67** | **£2,536.32** |
| *Not reported in the 2016 reconsideration of sorafenib by NICE; these values are therefore based only on the survey results presented in the original sorafenib submission to NICE. | | |

#### *5.4.9.4 Cost of managing adverse events*

The company included the costs of managing grade 3–4 AEs in the model. During the clarification stage, the company also explored the cost of managing grade 5 AEs in the model as a scenario analysis. The proportions of patients experiencing each AE in the model have been previously reported in Section 5.4.6. The resource use and total cost to manage each grade 3–4 AE is given in Table 39. As for grade 5 AEs, the total cost was assumed to equal £617.11 according to the average cost of a non-elective short-stay hospitalisation in NHS reference costs 2016/17.[63]

In order to apply the costs of managing AEs in the model, the total cost was applied upfront in the first model cycle, based on the mean number of episodes per patient. The total cost per patient was calculated by weighting the cost to treat AEs (Table 39) by the rates observed in the REFLECT trial (Table 29). The company also accounted for patients experiencing more than one episode of each AE in their calculation (Table 29). Following this, the one-off cost to manage grade 3–4 AEs was £581 for lenvatinib and £439 for sorafenib. As for the company's scenario which included grade 5 AEs, the one-off cost increased to £660 and £486, respectively.

Table 39. Cost of adverse events (reproduced from Table 39 of the CS)

| Adverse event | Items | Unit cost | Reference |
|---|---|---|---|
| Aspartate amino transferase increased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| Asthenia | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| | Nurse visit | £42.00 | PSSRU 2017 – Nurse (GP practice) – cost per hour, including qualifications[64] |
| | Total | £659.11 | - |
| Blood bilirubin increased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| | Outpatient contact | £172.67 | NHS reference costs 2016/17 – WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology)[63] |

| | | | |
|---|---|---|---|
| | CT scan | £101.57 | NHS reference costs 2016/17. Average of all computerised tomography currency codes (adult only), weighted by activity (RD20A, RD21A, RD22Z, RD23Z, RD24Z, RD25Z, RD26Z, RD27Z, RD28Z)[63] |
| | Total | £891.35 | - |
| Diarrhoea | Hospitalisation | £588.54 | NHS reference costs 2016/17 – FZ91K Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 6-10 – non-elective short-stay[63] |
| Fatigue | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| | Nurse visit | £42.00 | PSSRU 2017 – Nurse (GP practice) – cost per hour, including qualifications[64] |
| | Total | £659.11 | - |
| Gamma-glutamyl transferase increased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| Hypertension | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| | Outpatient contact | £172.67 | NHS reference costs 2016/17 – WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology)[63] |
| | GP contacts (x2) | £37.00 (x2) | PSSRU 2017 – General practitioner – cost per surgery consultation lasting 9.22 minutes – including direct care staff costs, with qualification costs[64] |
| | Total | £863.78 | - |
| Palmar-plantar erthrodysaesthesia syndrome | Hospitalisation | £431.64 | NHS reference costs 2016/17 – JD07J Skin Disorders without Interventions, with CC score 2-5 – non-elective short stay[63] |
| Platelet count decreased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| Proteinuria | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| | Outpatient contact | £172.67 | NHS reference costs 2016/17 – WF01A Consultant-led, Non-Admitted Face-to-Face Attendance, Follow-up (medical oncology)[63] |
| | Total | £789.78 | - |
| Weight decreased | Hospitalisation | £617.11 | NHS reference costs 2016/17 – average cost of non-elective short-stay[63] |
| | Dietician | £30.00 | PSSRU 2017 – dietitians/speech and language therapists - cost per working hour, Band 4[64] |
| | Total | £647.11 | - |
| Abbreviations: CC, complications and comorbidities; CT, computerised tomography; GP, general practitioner; PSSRU, personal social services research unit | | | |

#### *5.4.9.5 End-of-life costs*

A one-off end-of-life cost was applied to patients who die in the model. According to Georghiou and Bardsley 2014, end-of-life care comprises of hospital contacts, local authority-funded social care, nurse visits and GP visits.[66] In response to clarification questions, the company inflated the costs of those resources to 2016/17 prices and this change was incorporated into the company's revised base case analysis. The corrected end-of-life cost per patient applied in the revised model was £7,658.92.

#### *5.4.9.6 ERG critique*

Resource use is based on estimates reported in previous NICE appraisals, the draft SmPC for lenvatinib and the company's clinical experts' input.[24, 34, 62] The estimates used are based on the 2016/17 prices, with unit costs obtained from published sources such as the NHS national schedule of reference costs, the PSSRU and the British National Formulary (BNF), which is in line with the NICE reference case.[60, 61, 63, 64] The ERG validated all the costs from the sources cited, and checked that prices are correctly inflated when necessary, and that the formulae are generally correct and sound in the economic model. Following this, the ERG found a few minor discrepancies between the sources and the submission, which were corrected by the company at clarification. However, the ERG identified a few additional issues in the company's revised model provided at clarification regarding the implementation of subsequent anti-cancer interventions.

The ERG considers that the cost estimations in the model are generally correct and sound, but disagrees with the company's decision to exclude drug wastage costs in the base case analysis. This approach was also highlighted by the Appraisal Committee for TA474, who agreed that the most plausible ICER should account for drug wastage for up to 7 days.[24] However, the ERG considers a value of 7 days to be chosen quite arbitrarily as no justification was provided in either submission. For completeness, the ERG explored a scenario where drug costs were based on the planned number of capsules and tablets required each day (Section 6.2), effectively assuming that dose reductions account for wastage. Nonetheless, the impact of the ERG's scenario on the full population (including only sorafenib and regorafenib as subsequent interventions) was modest with the absolute difference in costs reducing from [redacted] to [redacted] (excluding the ERG's corrections to the model).

The ERG considered that some aspects of post-progression management costs, particularly social care, were potentially double counted by the company. During the clarification stage, the ERG asked the company to explain the difference between the social care costs applied to patients with progressed disease and end-of-life care costs. The company in their response agreed there may be some double counting for all components of end of life care costs, given that each of these aspects of resource use is costed in the progressed health state. To address this issue, the company showed that excluding end of life care costs in the full population had a negligible impact on the absolute difference in costs with a reduction of [redacted] (from [redacted] to [redacted]).

The ERG's clinical experts confirmed that the resource use was generally in line with what would be expected in UK clinical practice. However, clinical experts disagreed with the subsequent anti-cancer medications modelled by the company.

Firstly, as noted in Section 5.4.9.2, the company did not justify their assumption that subsequent anti-cancer medications (sorafenib or regorafenib) incur no administration, monitoring or testing costs. In

response to the ERG's clarification question, the company explained that no monitoring or testing costs were applied to avoid double counting, given that monitoring and testing costs are included in the post-progression state. Even so, the ERG considers that the surveys posed to the clinicians in TA189[34] and TA474[24] on resource use may not have explicitly stated the subsequent anti-cancer medications or procedures under consideration. Nonetheless, the ERG considers this to be a minor issue given the relatively low number of patients receiving subsequent anti-cancer medications in the base case analysis.

More importantly, the ERG does not consider the estimates of subsequent anti-cancer medications used in the model to be reflective of clinical practice in the UK, or the REFLECT trial. With regards to the former, clinical experts advised the ERG that patients in UK clinical practice would not receive subsequent anti-cancer medications, firstly as regorafenib is not recommended by NICE (ID991)[33] and secondly because it would be counterintuitive to offer regorafenib or sorafenib after failure on the same class of therapy. As for the latter, not all the subsequent anti-cancer medications received in the REFLECT trial are licensed for use in the UK, but as patients received them in the trial, their potential impact on survival is incorporated in effectiveness data. As noted in Section 4.3.6.2, survival was longer for patients who had received post-treatment anti-cancer interventions than those who did not, hence, offsetting this benefit with appropriate costs is justifiable.

To explore the uncertainty surrounding subsequent anti-cancer interventions (medications and procedures), the ERG asked the company to provide two scenarios: one which includes the full list of subsequent anti-cancer interventions patients in REFLECT received (including medications without a UK license) and one which excludes subsequent anti-cancer interventions.

During the clarification stage, the company provided a scenario for the ERG where all subsequent anti-cancer interventions observed in ≥1% of either arm in REFLECT were applied to the full population, except for tegafur-uracil, tivantinib and investigation drugs, whose costs were not available. The cost of subsequent anti-cancer medications and procedures applied in the model for the full population are reported Table 40 and Table 41, respectively.

To calculate doses dependent on body weight, or surface area, the company used the average weight (68.3kg) of patients included in REFLECT and body surface area (BSA) of 1.79$m^2$ reported by Sacco *et al.* 2010.[67] During validation, the ERG found Sacco *et al.* 2010 was conducted in the UK and reported BSAs separately for men (1.91$m^2$) and women (1.71$m^2$). Given that HCC is more common in men than women, the ERG notes that the company could have implemented a weighted BSA in the model.

Nonetheless, the ERG's calculation of BSA using the height and weight of patients in REFLECT was similar (1.77$m^2$) and the estimated cost of subsequent anti-cancer interventions (medications and

procedures) amounts to a total of £5,162 (£3,950 +£1,212) for lenvatinib and £4,203 (£2,887 + £1,316) for sorafenib in the company's analysis. However, upon inspection of the company's revised model, the ERG found that the scenario analysis was informed by medication costs alone, this was subsequently corrected by the ERG to also include the cost of procedures (Section 6).

Table 40. Acquisition cost of subsequent anti-cancer medications applied in scenario analysis (taken from the economic model provided at clarirfication)

| Drug | Units /pack | mg /unit | Cost/ pack* | Cost/ mg | mg/ day | mg/m2/ day | Cost/ day | % of progressed patients | | Mean duration (days) | | Weighted duration (days) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | LEN | SOR | LEN | SOR | LEN | SOR |
| **Oral medication** | | | | | | | | | | | | | |
| Sorafenib | 112 | 200 | £3,576.56 | £0.16 | 663.8 | - | £105.99 | [redacted] | [redacted] | [redacted] | [redacted] | 33.3 | 11.4 |
| Regorafenib | 84 | 40 | £3,744.00 | £1.11 | 144 | - | £160.46 | [redacted] | [redacted] | [redacted] | [redacted] | 0.6 | 4.3 |
| Capecitabine | 60 | 300 | £76.04 | £0.00 | - | 1250 | £6.30 | [redacted] | [redacted] | [redacted] | [redacted] | 0.8 | 1.1 |
| Gimeracil with oteracil potassium/tegafur | 84 | 20 | £248.40 | £0.15 | - | 25 | £4.96 | [redacted] | [redacted] | [redacted] | [redacted] | 2.4 | 0.5 |
| Thalidomide | 28 | 50 | £298.48 | £0.21 | 200 | - | £42.64 | [redacted] | [redacted] | [redacted] | [redacted] | 0.3 | 0.9 |
| Cabozantinib | 30 | 60 | £5,143.00 | £2.86 | 60 | - | £171.43 | [redacted] | [redacted] | [redacted] | [redacted] | 0.0 | 3.4 |
| Protein kinase inhibitors | 30 | 150 | £1,631.53 | £0.36 | 150 | - | £54.38 | [redacted] | [redacted] | [redacted] | [redacted] | 3.0 | 0.7 |
| **IV medication** | | | | | | | | | | | | | |
| Fluorouracil | 1 | 500 | £6.08 | £0.01 | - | 200 | £4.35 | [redacted] | [redacted] | [redacted] | [redacted] | 8.9 | 3.5 |
| Gemcitabine | 1 | 2000 | £26.86 | £0.01 | - | 1000 | £2.58 | [redacted] | [redacted] | [redacted] | [redacted] | 2.1 | 2.8 |
| Doxorubicin | 1 | 200 | £391.40 | £1.96 | - | 60 | £10.01 | [redacted] | [redacted] | [redacted] | [redacted] | 0.5 | 4.0 |
| Cisplatin | 1 | 50 | £25.11 | £0.50 | - | 15 | £2.41 | [redacted] | [redacted] | [redacted] | [redacted] | 9.7 | 3.6 |
| Nivolumab | 1 | 40 | £439.00 | £10.98 | - | 3 | £58.94 | [redacted] | [redacted] | [redacted] | [redacted] | 0.0 | 2.9 |
| Oxaliplatin | 1 | 200 | £591.26 | £2.96 | - | 85 | £32.13 | [redacted] | [redacted] | [redacted] | [redacted] | 2.0 | 2.6 |
| Folinic acid | 1 | 50 | £24.70 | £0.49 | | 20 | £0.00 (£1.76) | [redacted] | [redacted] | [redacted] | [redacted] | 1.5 | 0.5 |
| **Total cost** | **-** | **-** | **-** | **-** | **-** | **-** | **-** | **-** | **-** | **-** | **-** | **£3,950 (£3,952)** | **£2,887 (£2,888)** |

* Drug costs were taken from the BNF[61]; the cheapest cost per mg was assumed where alternative costs were available.
Abbreviations: IV, intravenous; LEN, lenvatinib; SOR, sorafenib
Folinic acid calculation corrected by the ERG in parentheses

Table 41. Cost of subsequent anti-cancer interventions applied in scenario analysis (taken from the economic model provided at clarirfication)

| Procedure | Cost per procedure | Cost source | % of progressed patients | |
|---|---|---|---|---|
| | | | Lenvatinib | Sorafenib |
| High frequency ablation | £998.00 | NHS reference costs 2016/17[63]. Weighted average of all radiofrequency ablation currency codes. | [redacted] | [redacted] |
| Radiotherapy | £198.48 | NHS reference costs 2016/17[63]. Weighted average of all radiotherapy currency codes. | [redacted] | [redacted] |
| Radiotherapy to bone | £198.48 | NHS reference costs 2016/17[63]. Weighted average of all radiotherapy currency codes. | [redacted] | [redacted] |
| Radiotherapy to brain | £198.48 | NHS reference costs 2016/17[63]. Weighted average of all radiotherapy currency codes. | [redacted] | [redacted] |
| Radiotherapy to liver | £198.48 | NHS reference costs 2016/17[63]. Weighted average of all radiotherapy currency codes. | [redacted] | [redacted] |
| Radiotherapy to lymph nodes | £198.48 | NHS reference costs 2016/17[63]. Weighted average of all radiotherapy currency codes. | [redacted] | [redacted] |
| Regional chemotherapy | £4,401.68 | NHS reference costs 2016/17[63]. Percutaneous, Chemoembolisation or Radioembolisation, of Lesion of Liver. | [redacted] | [redacted] |
| Transcatheter arterial chemoembolisation | £4,401.68 | NHS reference costs 2016/17[63]. Percutaneous, Chemoembolisation or Radioembolisation, of Lesion of Liver. | [redacted] | [redacted] |
| **Total cost** | **-** | **-** | **£1,212** | **£1,316** |

Abbreviations: NHS, National Health Service.

The ERG has three minor concerns regarding the company's scenario analysis which includes all subsequent anti-cancer interventions observed in ≥1% of either arm in REFLECT. Firstly, the company did not consider the difference in resource and cost use between oral and IV medication administrations. Secondly, the company applied the weighted cost of all radiotherapy currency codes to each type of radiotherapy procedure which may underestimate the cost of radiotherapy for the indications in REFLECT. And thirdly, it was unclear in the company's response how missing data on the duration of treatment was handled by the company; i.e. whether the last observation was used as a proxy for the end date or whether those observations were excluded from the calculation of the mean. Nonetheless, given that the route of medications, types of radiotherapy and missing data were reasonably balanced across the treatment arms, the ERG did not explore these issues further.

One further limitation of the company's scenario analysis, the ERG would like to reiterate, is the imbalance in investigational drugs between lenvatinib and sorafenib and the uncertainty regarding their composition (Table 42). Consequently, if data were available on the composition of investigational drugs, the impact on cost-effectiveness could be large, although the direction of their inclusion in relation to costs is uncertain in the absence of data.

Table 42. Investigational drugs by treatment arm and region

| Subsequent interventions received during survival follow-up, n (%) | Lenvatinib | | Sorafenib | |
|---|---|---|---|---|
| | Full | Western | Full | Western |
| Investigational drugs | [redacted] | [redacted] | [redacted] | [redacted] |

As for the Western subgroup, subsequent anti-cancer interventions (medications or procedures) were not costed in a scenario analysis because the Western subgroup analysis controls for subsequent anti-cancer interventions and, therefore, the company concluded it would be inconsistent to apply different subsequent intervention costs by treatment arm. However, the ERG notes that the adjustment was based only on the categorisation of patients into those who received any subsequent anti-cancer procedure or medication and those who received neither. As such, the adjustment does not control for any imbalances in the types of intervention received in each treatment arm. The ERG was unable to explore an analysis, which included different subsequent intervention costs by treatment arm, as the company did not provide the percentage of progressed patients receiving subsequent interventions or the mean duration of subsequent interventions for the Western subgroup. As such, the ERG recommends that the company's analysis of the Western subgroup is interpreted with caution.

Within the company's Western subgroup analysis, the ERG would also like to note that the mean dose intensity was taken from the company's base case analysis informed by the full population (lenvatinib, 9.4 mg/day; sorafenib 669.1mg/day), rather than the mean dose intensity in the Western subgroup provided at clarification (lenvatinib [redacted] mg/day; sorafenib, [redacted] mg/day). However, given that the ERG prefers to include the costs of drug wastage in its base case analysis, this issue becomes somewhat redundant.

Finally, as for the company's scenario which excludes subsequent anti-cancer interventions, the impact on the full population was notable and favoured lenvatinib (as lenvatinib was associated with higher subsequent anti-cancer medication costs than sorafenib in the base case analysis) with the difference in costs increasing from [redacted] to [redacted]. A scenario was not presented by the company for the Western subgroup as the multivariable models for PFS and OS estimated in the full population included region as a covariate and adjusted for the use of post-progression therapies in the company's Western subgroup analysis.

Overall, while there was variation in the type and extent of subsequent interventions, neither the full population nor the Western subgroup are considered reflective of UK clinical practice. Moreover, the company's results from the full population and Western subgroup are robust to all additional analyses which change the approach to cost subsequent anti-cancer interventions.

## 5.5 Results included in company's submission

The company presented deterministic and probabilistic results. The base case results were calculated deterministically (using mean parameter values) as well as probabilistically (assessing the simultaneous effect of parameter uncertainty).

The company also carried out a series of univariate sensitivity analyses and scenario analyses to test the robustness of model results to changes in model parameters and structural assumptions. Base case results are presented in Section 5.5.1, while the results of deterministic and probabilistic sensitivity analyses are presented in Sections 5.5.2 and 5.5.3, respectively. No subgroup analyses were performed in the company's original model, although a Western subgroup was provided at clarification as requested by the ERG, as discussed in Section 5.4.5.

## 5.5.1 Base case results

The original multivariable analyses contained Child–Pugh score as a continuous variable. However, during the clarification stage, Child–Pugh class as a categorical variable (A vs B) was considered more appropriate by the company. Consequently, all analyses used to generate results in the revised economic model were based on re-estimated statistical models. In response to clarification questions, the company also corrected end-of-life care costs in their revised analyses. All other aspects of the analyses remain the same.

The results of the company's revised base case analysis using PAS prices for lenvatinib are presented in Table 43. As described previously in Section 5.4.5, the base-case analysis was based on multivariable adjustments to the PFS and OS curves for imbalances in baseline characteristics (including AFP and HCC aetiology) but not for imbalances in subsequent treatments.

According to the company's base case analysis, lenvatinib is expected to extend patients' lives by around 2.7 months compared to sorafenib. This translates to an incremental average QALY gain for lenvatinib of [redacted] QALYs. Overall, sorafenib is dominated by lenvatinib, in that lenvatinib is less costly and more effective.

Table 43. Base-case results (taken from the company's response to clarification)

| Technologies | Total costs | Total LYG | Total QALYs | Incremental costs | Incremental LYG | Incremental QALYs | ICER |
|---|---|---|---|---|---|---|---|
| Sorafenib | £65,592 | 1.46 | 1.03 | - | - | - | - |
| Lenvatinib | [redacted] | 1.69 | [redacted] | [redacted] | 0.23 | [redacted] | Dominated |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. | | | | | | | |

## 5.5.2 Deterministic sensitivity analysis

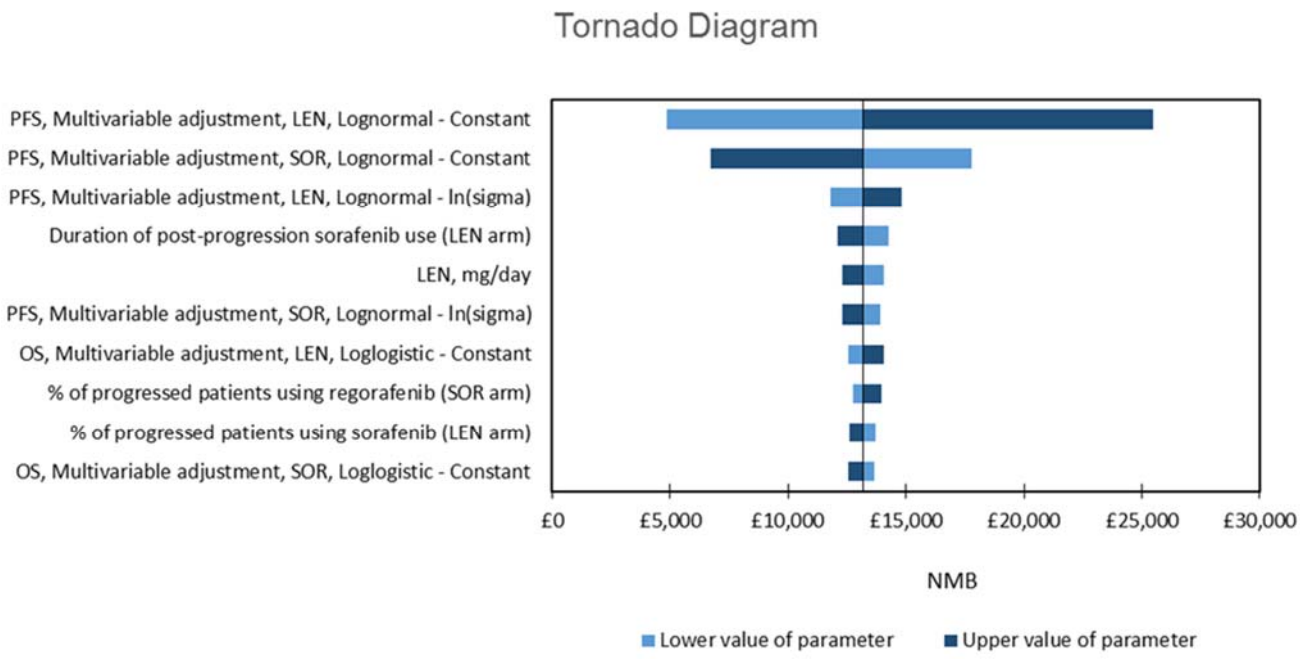
The company carried out univariate sensitivity analyses to assess the impact of varying the values of parameters from their means over a plausible range determined by either the 95% confidence interval

(CI), or +/- 15% where no estimates of precision were available. Upper and lower ranges of included parameters are presented in Appendix N of the CS. The company also carried out scenario analyses changing assumptions surrounding:

- Drug wastage;
- Covariate adjustments (including unadjusted parametric models and adjustments for AFP only);
- OS distribution;
- PFS distribution;
- Resource use and costs;
- Discount rates;
- Time horizon;
- Utility data.

Assuming a cost-effectiveness threshold of £50,000 per QALY, the company presented the results of deterministic sensitivity analysis using the net monetary benefit (NMB) in the CS. During the clarification stage, results of sensitivity analysis from the revised model were only provided in the economic model. For completeness, the results of the univariate analyses and scenario presented in the revised model using the lenvatinib PAS price are presented in Figure 18 and Table 44, respectively.

Figure 18. Results of univariate sensitivity analysis (tornado diagram) (reproduced from the economic model provided at clarification)




Abbreviations: LEN, lenvatinib; NMB, net monetary benefit; PFS, progression free survival; OS, overall survival; SOR, sorafenib

Table 44. Results of scenario analysis for lenvatinib versus sorafenib (reproduced from the economic model provided at clarification)

| # | Scenario | Incremental Costs | Incremental QALYs | NMB | % change from base-case NMB | ICER |
|---|---|---|---|---|---|---|
| - | Base-case | [redacted] | [redacted] | [redacted] | 0% | Dominant |
| 1 | Drug wastage included | [redacted] | [redacted] | [redacted] | 2% | Dominant |
| 2 | Mortality costs excluded | [redacted] | [redacted] | [redacted] | -1% | Dominant |
| 3 | No covariate adjustment | [redacted] | [redacted] | [redacted] | -5% | Dominant |
| 4 | Adjustment for AFP and stratification factors only | [redacted] | [redacted] | [redacted] | -2% | Dominant |
| 5 | OS distribution: log-normal | [redacted] | [redacted] | [redacted] | -3% | Dominant |
| 6 | OS distribution: gamma | [redacted] | [redacted] | [redacted] | -1% | Dominant |
| 7 | OS distribution: Weibull | [redacted] | [redacted] | [redacted] | 0% | Dominant |
| 8 | OS distribution: Gompertz | [redacted] | [redacted] | [redacted] | 0% | Dominant |
| 9 | OS distribution: exponential | [redacted] | [redacted] | [redacted] | -5% | Dominant |
| 10 | PFS distribution: log-logistic | [redacted] | [redacted] | [redacted] | -3% | Dominant |
| 11 | PFS distribution: gamma† | [redacted] | [redacted] | [redacted] | -1% | Dominant |
| 12 | PFS distribution: Weibull | [redacted] | [redacted] | [redacted] | 0% | Dominant |
| 13 | PFS distribution: Gompertz† | [redacted] | [redacted] | [redacted] | 0% | Dominant |
| 14 | PFS distribution: exponential | [redacted] | [redacted] | [redacted] | -5% | Dominant |
| 15 | Resource use costs halved (all states) | [redacted] | [redacted] | [redacted] | -4% | Dominant |
| 16 | Resource use costs doubled (all states) | [redacted] | [redacted] | [redacted] | 7% | Dominant |
| 17 | Target dose assumed | [redacted] | [redacted] | [redacted] | 13% | Dominant |
| 18 | Discount rates of 1.5% | [redacted] | [redacted] | [redacted] | 3% | Dominant |
| 19 | Time horizon: 1 year | [redacted] | [redacted] | [redacted] | -44% | Dominant |
| 20 | Time horizon: 2 years | [redacted] | [redacted] | [redacted] | -25% | Dominant |
| 21 | Time horizon: 5 years | [redacted] | [redacted] | [redacted] | -3% | Dominant |
| 22 | Sorafenib utility data used (committee preference) [a] | [redacted] | [redacted] | [redacted] | 2% | Dominant |
| 23 | Regorafenib utility data used [b] | [redacted] | [redacted] | [redacted] | 4% | Dominant |
| 24 | Post-progression utility of 0.5 | [redacted] | [redacted] | [redacted] | 8% | Dominant |
| 25 | 60% discount applied to sorafenib | [redacted] | [redacted] | [redacted] | -86% | [redacted] |

Abbreviations: AFP, alpha fetoprotein; ICER, incremental cost-effectiveness ratio; NMB, net monetary benefit; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life-year.

[a] PFS utility, 0.76; progressed utility, 0.68; [b] PFS utility, 0.797; progressed utility 0.749

† In these scenarios the PFS curves for lenvatinib and sorafenib are assumed equivalent at the point they would otherwise cross each other.

According to the univariate sensitivity analysis, the main drivers of the model are the constant terms (baseline hazard of events) for the base-case PFS and OS models (lognormal and loglogistic) for each of the lenvatinib and sorafenib arms. As for the scenario analysis, results were most sensitive to a 60% discount to the sorafenib list price and shorter time horizons.

### 5.5.3 Probabilistic sensitivity analysis

The company performed probabilistic sensitivity analysis (PSA) to assess the joint parameter uncertainty around the base case results. The results are based on 10,000 PSA iterations.

The ERG considers the parameters and respective distributions chosen for PSA, outlined in Appendix N of the CS, to be generally sound. In summary, utilities, AE rates and the proportion of patients receiving disease management were varied using a beta distribution, while unit costs and the frequency of resource use were varied using a gamma distribution. Drug acquisition cost were kept constant. Where the covariance structure between parameters was known (survival curves), correlated random draws were sampled from a multivariate normal distribution.

The results of the company's PSA in the revised model using PAS prices for lenvatinib are presented in Table 45. According to the company's analysis, lenvatinib was associated with mean cost-savings of [redacted] and mean incremental QALYs of [redacted] which the ERG considers to be comparable to the deterministic base case results. Furthermore, the ERG could produce very similar PSA results when they replicated the analysis using 10,000 iterations.

Table 45. PSA results using 10,000 iterations (reproduced from the economic model provided at clarification)

| Technologies | Total costs | Total QALYs | Incremental costs | Incremental QALYs | ICER |
|---|---|---|---|---|---|
| Sorafenib | £65,688 | 1.03 | - | - | - |
| Lenvatinib | [redacted] | [redacted] | [redacted] | [redacted] | Dominated |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years | | | | | |

The scatterplots and cost-effectiveness acceptability curves (CEACs) are presented in Figure 19 and Figure 20 using PAS prices for lenvatinib. The majority of PSA iterations lie in the south-east quadrant of the cost-effectiveness plane. Moreover, the probability that lenvatinib was cost-effective at a threshold of £50,000 per QALY was 100%.

Figure 19. Cost-effectiveness plane (reproduced from the economic model provided at clarification)

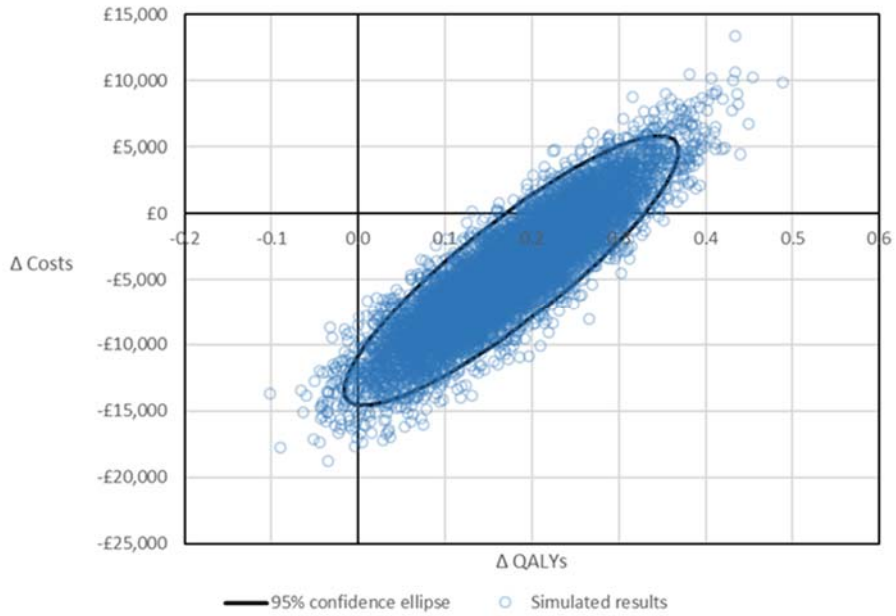
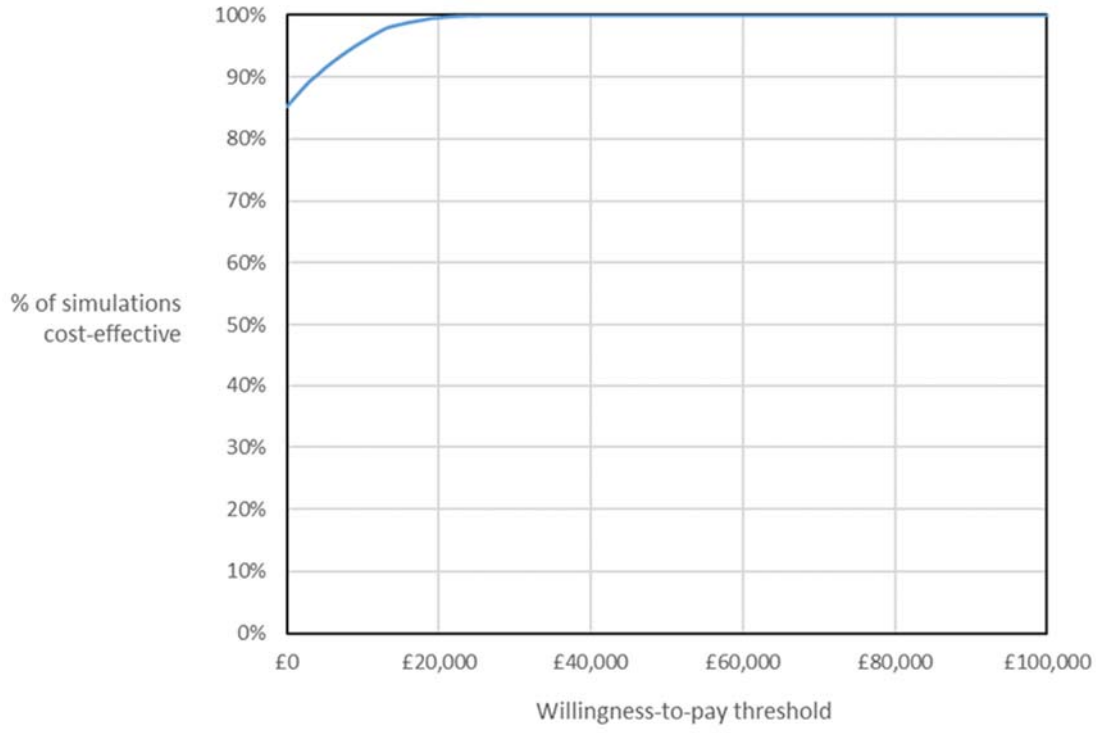



Figure 20. Cost-effectiveness acceptability curve (CEAC; reproduced from the economic model provided at clarification)




## 5.5.4 Model validation

The ERG conducted a thorough validation of the company's economic model to check for inconsistencies with the CS and to test the functionality of the model. The model was found to be generally sound with no serious issues.

# 6 ADDITIONAL WORK UNDERTAKEN BY THE ERG

## 6.1 Model corrections

The ERG identified a minor error at the clarification stage whereby the company had not incorporated some of the cost components within the half-cycle correction; namely the post-progression treatment costs and mortality costs. The company provided a justification for this in response to clarification questions; however, the ERG did not consider it to address the issue. The results with the ERG's correction applied are very similar to the company's submitted base case (re-submitted after clarification questions) and the results are given in Table 46.

Table 46. Company's corrected base-case results

| Technologies | Total costs | Total LYG | Total QALYs | Incremental costs | Incremental LYG | Incremental QALYs | ICER |
|---|---|---|---|---|---|---|---|
| Sorafenib | £65,574 | 1.46 | 1.03 | - | - | - | - |
| Lenvatinib | [redacted] | 1.69 | [redacted] | [redacted] | 0.23 | [redacted] | Dominant |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. | | | | | | | |

## 6.2 ERG scenario analysis

The ERG conducted a number of scenario analyses around the company's corrected base case ICER. Each scenario is described in the following paragraphs.

*Scenario 1: Subsequent anti-cancer intervention adjustment to OS*

The company's subsequent anti-cancer intervention adjustment for OS was applied to account for the imbalance in subsequent treatments received in each group of the trial. The associated costs were removed. The ERG considered the adjusted analysis to provide a more reliable method to account for benefits received from subsequent treatments, rather than applying relevant costs, which may cause greater bias if the costs of the treatments are not equally valued in terms of the benefits received.

*Scenario 2: Gamma distribution for PFS*

The ERG considered the gamma distribution to provide a better fit to the data for both groups of the REFLECT trial, as the company's lognormal curves did not provide a sufficiently good fit for the sorafenib group. An adjustment to prevent the curves crossing was also applied, as the ERG considered the curves to be otherwise implausible.

*Scenario 3: Gamma distribution for OS*

Although the company's chosen log-logistic OS curves did provide the best fit to the data, the gamma distribution also provided a reasonably good fit to the data but with a more conservative extrapolation,

so this was applied to assess the impact on the results. An adjustment to prevent the curves crossing was also applied, as the ERG considered the curves to be otherwise implausible.

*Scenario 4: Mortality adjustment for newly progressed patients*

The company's post-progression drug costs were applied to 'newly progressed' patients; however, the proportion of patients assumed to be newly progressed was taken as the difference in PFS between cycles. This is an overestimate of the proportion of newly progressed patients as some of those patients who leave the PFS state are people who have died. The ERG's preferred approach is to apply the mortality rate between cycles to PFS and take this off the company's estimate. This is not an accurate estimate as it assumes mortality is the same for progressed and non-progressed patients; however, the ERG considers it preferable to the clearly overestimated approach. An adjustment was also applied to ensure that the estimated proportion was non-negative.

*Scenario 5: All post-progression interventions (including procedures)*

This scenario analysis applies the costs of all post-progression therapies occurring in at least 1% of the trial population, and includes post-progression procedures as well as post-progression drugs.

*Scenario 6: Scenario 4 and 5*

Combines changes from scenario 4 and scenario 5.

*Scenario 7: Scenario 1 with Western subgroup with Western subgroup AEs*

This scenario applies the changes made from Scenario 1 as well as applying the Western subgroup survival analyses and the Western subgroup AE data.

*Scenario 8: Assumes full costs of drugs (without dose reductions)*

The company assumed a value of 7 days of treatment for the scenario analyses including wastage of drugs. The ERG was concerned that this arbitrary value may underestimate the potential for wastage, so the ERG conducted a scenario analysis that rounded up the mean dose received to the planned dose. This does not necessarily reflect wastage accurately but gives an indication of a potential upper bound if we assume that the value of dose reductions accounts for wastage.

*Scenario 9: Removes AE costs*

Simply removes the costs applied for the management of AEs to assess how great an impact they have.

Table 47. Scenario analyses around company's corrected base case ICER

| | Results per patient | Sorafenib (1) | Lenvatinib (2) | Incremental value (2-1) |
|---|---|---|---|---|
| **0** | **Company's corrected base case** | | | |
| | Total costs (£) | £65,574 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.03 | [redacted] | [redacted] |
| | ICER | | | Dominant |
| **1** | **Subsequent anti-cancer intervention adjustment to OS and removal of post-progression therapy costs** | | | |
| | Total costs (£) | £60,243 | [redacted] | [redacted] |
| | LYs | 1.35 | 1.65 | 0.30 |
| | QALYs | 0.95 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **2** | **Gamma distribution for PFS (with prevention of curves crossing)** | | | |
| | Total costs (£) | £61,397 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.04 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **3** | **Gamma distribution for OS (with prevention of curves crossing)** | | | |
| | Total costs (£) | £62,338 | [redacted] | [redacted] |
| | LYs | 1.36 | 1.54 | 0.18 |
| | QALYs | 0.96 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **4** | **Mortality adjustment for newly progressed patients** | | | |
| | Total costs (£) | £64,979 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.03 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **5** | **All post-progression interventions (including procedures)** | | | |
| | Total costs (£) | £67,824 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.03 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **6** | **Scenario 4 and 5** | | | |
| | Total costs (£) | £66,508 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.03 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **7** | **Scenario 1 with Western subgroup with Western subgroup AEs** | | | |
| | Total costs (£) | £66,019 | [redacted] | [redacted] |
| | LYs | 1.64 | 1.75 | 0.11 |
| | QALYs | 1.08 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **8** | **Assumes full costs of drugs (without dose reductions)** | | | |

| | | | | |
|---|---|---|---|---|
| | Total costs (£) | £70,234 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.03 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| **9** | **Removal of AEs** | | | |
| | Total costs (£) | £65,135 | [redacted] | [redacted] |
| | LYs | 1.46 | 1.69 | 0.23 |
| | QALYs | 1.03 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| Abbreviation used in the table: AE, adverse event; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALYs, quality-adjusted life years. | | | | |

## 6.3 ERG base case ICER

The ERG's preferred base case analysis is based on two key changes to the company's corrected base case analysis presented in Section 6.1. The changes are outlined in the following paragraphs.

The company's subsequent anti-cancer intervention adjustment for OS was applied to account for the imbalance in subsequent treatments received in each group of the trial. The company's applied subsequent anti-cancer intervention costs were removed. The ERG considered the adjusted analysis to provide a more reliable method to account for benefits received from subsequent treatments, rather than applying relevant costs, which may cause greater bias if the costs of the treatments are not equally valued in terms of the benefits received.

The second change the ERG made was to apply the gamma distribution for PFS. This also required the company's correction to prevent the two treatment curves from crossing over, which the ERG considered to be implausible and, therefore, a necessary correction. The ERG chose this distribution as it provided a good fit to both groups, whereas the company's chosen lognormal curve did not provide a good fit for the sorafenib group.

The results of the ERG's changes are presented, incorporating each change cumulatively, in Table 48. An ICER for each individual change compared to the base case is also given.

Table 48. ERG base case ICER

| Results per patient | Sorafenib (1) | Lenvatinib (2) | Incremental value (2-1) |
|---|---|---|---|
| **Company's corrected base case** | | | |
| Total costs (£) | £65,574 | [redacted] | [redacted] |
| QALYs | 1.03 | [redacted] | [redacted] |
| ICER | | | Dominant |
| **Post-progression adjustment to OS and removal of post-progression therapy costs** | | | |
| Total costs (£) | £60,243 | [redacted] | [redacted] |
| QALYs | 0.95 | [redacted] | [redacted] |
| **ICER (compared with base case)** | | | Dominant |

| | | | |
|---|---|---|---|
| **ICER with all changes incorporated** | | | Dominant |
| **Gamma distribution for PFS (with prevention of curves crossing)** | | | |
| Total costs (£) | £56,237 | [redacted] | [redacted] |
| QALYs | 0.96 | [redacted] | [redacted] |
| **ICER (compared with base case)** | | | Dominant |
| **ICER with all changes incorporated** | | | Dominant |
| **ERG's preferred base case ICER** | | | **Dominant** |
| Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALYs, quality-adjusted life years. | | | |

In the company's economic model, the covariance matrix for the parameters of the company's preferred survival distributions was provided. This was used by the company to generate correlated random draws from the sample space of the survival model parameters using a multivariate normal distribution.

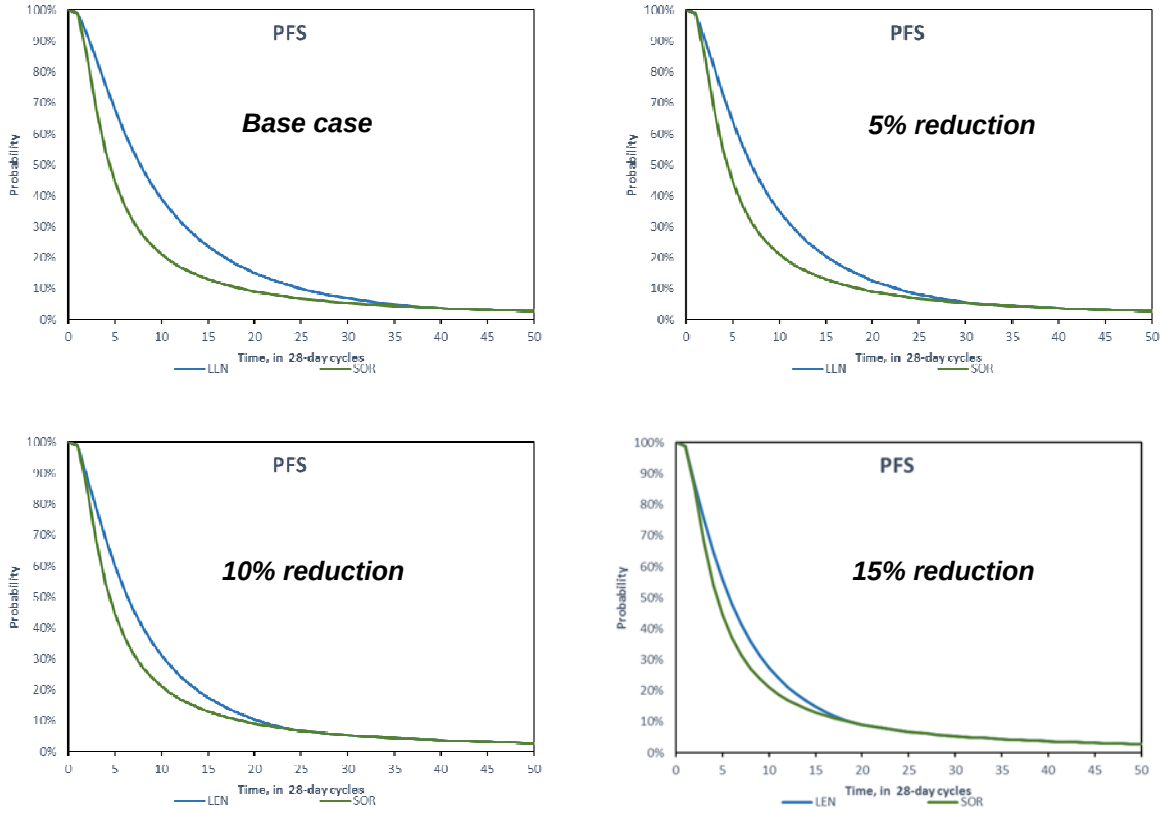
For the alternative survival distributions, the covariance matrices were not provided and, hence, the ERG was unable to incorporate the uncertainty around the ERG's preferred survival models into a PSA. The ERG considers the results of a PSA without this uncertainty incorporated to be unreliable and potentially misleading as an assessment of the impact of parameter uncertainty on the ICER. Therefore, the ERG did not conduct a PSA for the ERG's preferred base case analysis.

## *6.4 Scenario analyses around ERG base case ICER*

The following outlines the scenario analyses undertaken by the ERG around the ERG's preferred base case analysis.

The ERG was concerned that the PFS difference between the two treatment groups may have been overestimated by including treatment discontinuation as a censor in the analysis of PFS. Therefore, the ERG conducted a scenario analysis to reduce the scale of the lenvatinib PFS gamma curve by 5%, 10% and 15%. The results of these analyses are given in Table 49, labelled as Scenario 1, 2, and 3, respectively. The resulting PFS curves after the reductions in the scale parameters, along with the ERG's preferred base case curves, are displayed in Figure 21.

Figure 21. PFS gamma curves with reductions in the scale parameters




Abbreviations: LEN, lenvatinib; PFS, progression-free survival; SOR, sorafenib.

The ERG also performed a scenario analysis to change the distribution of the OS curve to the gamma function. This function provided a similar goodness-of-fit to the data but with a more conservative extrapolation. The results of this scenario are given in Table 49, labelled as Scenario 4.

Table 49. Scenario analyses around ERG's preferred base case ICER

| | Results per patient | Sorafenib (1) | Lenvatinib (2) | Incremental value (2-1) |
|---|---|---|---|---|
| **0** | **ERG's preferred base case** | | | |
| | Total costs (£) | £56,237 | [redacted] | [redacted] |
| | LYs | 1.35 | 1.65 | 0.30 |
| | QALYs | 0.96 | [redacted] | [redacted] |
| | ICER | | | Dominant |
| **1** | **Reduce scale of gamma PFS function by 5%** | | | |
| | Total costs (£) | £56,237 | [redacted] | [redacted] |
| | LYs | 1.35 | 1.65 | 0.30 |
| | QALYs | 0.96 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | £2,085 |
| **2** | **Reduce scale of gamma PFS function by 10%** | | | |
| | Total costs (£) | £56,237 | [redacted] | [redacted] |
| | LYs | 1.35 | 1.65 | 0.30 |
| | QALYs | 0.96 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | £8,490 |

| | | | | |
|---|---|---|---|---|
| **3** | **Reduce scale of gamma PFS function by 15%** | | | |
| | Total costs (£) | £56,237 | [redacted] | [redacted] |
| | LYs | 1.35 | 1.65 | 0.30 |
| | QALYs | 0.96 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | £14,024 |
| **4** | **Gamma distribution for OS (with prevention of curves crossing)** | | | |
| | Total costs (£) | £55,765 | [redacted] | [redacted] |
| | LYs | 1.31 | 1.55 | 0.24 |
| | QALYs | 0.93 | [redacted] | [redacted] |
| | **ICER (compared with base case)** | | | Dominant |
| | Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALYs, quality-adjusted life years. | | | |

# 7 END OF LIFE

The company propose that lenvatinib meets both criteria outlined by the National Institute for Health and Care Excellence (NICE) for an end of life treatment. The company's assessment and ERG's comments are provided in Table 50.

The company refers to the 2012 joint European Society for Medical Oncology and European Society of Digestive Oncology (ESMO–ESDO) clinical practice guidelines as evidence for the first criterion, which break down prognosis by Barcelona Clinic Liver Cancer (BCLC) stages. The guidelines state that median OS for those with advanced HCC (BCLC Stage C), the relevant population for this Single Technology Appraisal, is 4 to 8 months based on natural history, and 6 to 11 months for those treated with sorafenib (current standard of care).[21] Median OS in the sorafenib group of REFLECT (12.3 months), and mean survival for sorafenib from the company's updated base case (17.5 months) and the ERG's base case (16.2 months), all fall below the threshold of 24 months.

The second criterion was met based on the company's original base case extrapolation of OS (incremental mean OS benefit of 3.1 months), but not in the updated base case provided at the clarification stage (2.7 months). Incremental mean OS based on the ERG's preferred assumptions was 3.6 months, including adjustment for subsequent treatments, and therefore meets the second criteria.

Overall, there is little uncertainty that lenvatinib meets the first end-of-life criterion concerning life expectancy, but whether the second criterion it met depends on the preferred assumptions on which the extrapolation of OS is based in the review of cost effectiveness.



Table 50. End of life considerations

| NICE criterion | Company assessment | ERG comment |
|---|---|---|
| The treatment is indicated for patients with a short life expectancy, normally less than 24 months | CS section B.1.3, page 12. The current life expectancy for patients with advanced HCC is less than 1 year.[21] | ESMO-ESDO guidelines give median OS based on natural history as 4 to 8 months, and 6 to 11 months with sorafenib, for people with BCLC Stage C HCC.<br>From the company submission:<br>• Sorafenib median OS in REFLECT (primary analysis) = 12.3 months (upper quartile 25.4 months).<br>• Sorafenib mean OS = 17.5 months in company base case, 16.2 months in ERG's base case. |
| 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 | CS section B.3.11, page 116. Results of the cost-effectiveness analysis show an incremental mean overall survival benefit for lenvatinib of 3.1 months compared with sorafenib. | Incremental mean OS benefit from the company's economic model:<br>• Company base case = 2.7 months;<br>• ERG base case = 3.6 months. |

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; CS, company submission; ERG, evidence review group; ESMO-ESDO, joint European Society for Medical Oncology and European Society of Digestive Oncology guidance; HCC, hepatocellular carcinoma; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; OS, overall survival.

# 8 OVERALL CONCLUSIONS

REFLECT demonstrates similar overall survival (OS) between lenvatinib and sorafenib, which is standard of care in the UK, in a mixed Asia–Pacific (67.1%) and Western population (32.9%). Evidence from REFLECT may not generalise to patients with untreated HCC who have compromised liver function (Child–Pugh stage B or worse) or Eastern Cooperative Oncology Group (ECOG) performance status of 2, but is otherwise considered appropriate to a UK population with untreated, advanced or unresectable hepatocellular carcinoma (HCC).

OS in both groups is likely to be longer than would be seen in UK practice due to the type and extent of subsequent therapies received, and the direction of imbalance between groups is likely to bias OS in favour of sorafenib. As such, analyses that adjust for imbalances in subsequent therapies, and for imbalances in baseline characteristics that are considered prognostic of outcome, are likely to be the least biased and most applicable to UK patients.

Independent tumour assessments for key secondary outcomes of progression and response confirmed results from the primary investigator measures, which are more prone to bias from the study's open-label design. Lenvatinib was consistently superior to sorafenib for progression-free survival (PFS), time to progression (TTP) and objective response rate (ORR). The choice of censoring in the primary analysis of PFS [redacted] effectiveness of lenvatinib, but Kaplan–Meier data for a supplementary analysis considered less biased were not available to assess the impact on the ICER.

After consideration of differences between the full population and Western subgroup (including baseline characteristics, treatment exposure, adverse events, and subsequent therapies), the full population was considered suitable for the primary analysis. Subgroup results suggested the effect of lenvatinib on PFS compared with sorafenib may be reduced in the Western population (HR 0.81, 95% CI: 0.61 to 1.08), but OS was similar between the Western subgroup and full population when both were adjusted for subsequent treatment.

Patients in the lenvatinib group spent longer on treatment, and were more likely to discontinue treatment for reasons other than progression (including adverse events). Nonetheless, incidence of adverse effects is likely to be the most appropriate way to characterise tolerability across a course of treatment. Drug withdrawals and dose modifications due to TEAEs, serious adverse events and Grade 3+ TEAEs were all more common in the lenvatinib group than the sorafenib group. Some tolerability issues were reflected in specific quality of life metrics, but patients taking lenvatinib are likely to have similar overall quality of life to those taking sorafenib.

The company's economic analyses were generally sound but with a few weaknesses and some remaining uncertainties. The key weakness is in the method applied for the covariate adjustment of the

parametric models. The ERG considers it inappropriate to apply a variable selection procedure on one model (i.e. the Cox proportional hazards model) and then apply that set of variables directly to a parametric model for adjustment. The set of variables selected may differ if the selection procedure was applied to the parametric models considered for the model.

The key source of uncertainty lies in the censoring approach used by the company for the PFS analysis, for which the impact on the economic analysis is difficult to quantify. This censoring approach may overestimate the benefits of lenvatinib, which potentially underestimates the costs and overestimates the QALYs gained. The results of the analyses should, therefore, be interpreted with caution.

## 8.1 Implications for research

Should lenvatinib be approved, observational evidence of its use in a UK population would help to resolve the remaining uncertainty about the applicability of REFLECT to UK patients. For example, a real-world study or retrospective chart review could assess whether differences between the trial and UK clinical practice - such as population differences, medication compliance and subsequent therapies - impact the effectiveness of lenvatinib for UK patients.

# 9 REFERENCES


1. National Institute for Health and Care Excellence (NICE). Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]: Final scope2017 6 Mar 2018. Available from: https://www.nice.org.uk/guidance/gid-ta10150/documents/final-scope.

2. National Cancer Registration and Analysis Service. Trends in incidence of primary liver cancer subtypes2007 12 Mar 2018. Available from: http://www.ncin.org.uk/publications/data_briefings/trends_in_incidence_of_primary_liver_cancer_subtypes.

3. Office for National Statistics. Registrations of newly diagnosed cases of cancer (third digit): site, sex and age, England, 20152015 12 Mar 2018. Available from: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland/2015/cancerregistrations2015final22.05.2017.xls.

4. Cancer Research UK. Liver cancer incidence statistics2018 12 Mar 2018. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer/incidence#heading-Zero.

5. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology. 2007;132(7):2557-76.

6. Pons F, Varela M, Llovet JM. Staging systems in hepatocellular carcinoma. HPB (Oxford). 2005;7(1):35-41.

7. Das D, Chattopadhyay D, Aslam T, Patanwala I, Walia D, Rose J, et al., editors. NAFLD and the changing face of hepatocellular cancer (HCC).2011: Gut.

8. King J, Palmer DH, Johnson P, Ross P, Hubner RA, Sumpter K, et al. Sorafenib for the Treatment of Advanced Hepatocellular Cancer - a UK Audit. Clin Oncol (R Coll Radiol). 2017;29(4):256-62.

9. Parkin DM. 11. Cancers attributable to infection in the UK in 2010. Br J Cancer. 2011;105 Suppl 2:S49-56.

10. Cancer Research UK. Liver cancer. Symptoms2015 12 Mar 2018. Available from: http://www.cancerresearchuk.org/about-cancer/liver-cancer/symptoms.

11. Cancer Research UK. Liver cancer. Diagnosis. 2015.

12. Cancer Research UK. BCLC staging system and the Child-Pugh system2015 12 Mar 2018. Available from: http://www.cancerresearchuk.org/about-cancer/liver-cancer/stages/bclc-staging-system-child-pugh-system.

13. Llovet JM, Pena CE, Lathia CD, Shan M, Meinhardt G, Bruix J, et al. Plasma biomarkers as predictors of outcome in patients with advanced hepatocellular carcinoma. Clin Cancer Res. 2012;18(8):2290-300.

14. Pascual S, Herrera I, Irurzun J. New advances in hepatocellular carcinoma. World J Hepatol. 2016;8(9):421-38.

15. Silva JP, Gorman RA, Berger NG, Tsai S, Christians KK, Clarke CN, et al. The prognostic utility of baseline alpha-fetoprotein for hepatocellular carcinoma patients. J Surg Oncol. 2017;116(7):831-40.

16. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5(6):649-55.

17. Office for National Statistics. Cancer survival in England: adults diagnosed between 2011 and 2015 and followed up to 2016.2017 12 Mar 2018. Available from: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed/20112015/adultcancersurvivalcorrectionfinal.xls.

18. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg. 1973;60(8):646-9.

19. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol. 2006;44(1):217-31.

20. Cancer EAfSotL-EOfRaTo. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. Eur J Cancer. 2012;48(5):599-641.

21. Verslype C, Rosmorduc O, Rougier P, Group EGW. Hepatocellular carcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012;23 Suppl 7:vii41-8.

22. Ryder SD, British Society of G. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (HCC) in adults. Gut. 2003;52 Suppl 3:iii1-8.

23. National Institute for Health and Care Excellence (NICE). Liver cancers pathway: interventional procedures for treating hepatocellular carcinoma2017 14 Mar 2018. Available from: https://pathways.nice.org.uk/pathways/liver-cancers.

24. National Institute for Health and Care Excellence (NICE). Sorafenib for treating advanced hepatocellular carcinoma: technology appraisal guidance [TA474]2017 01 Feb 2018. Available from: https://www.nice.org.uk/guidance/ta474.

25. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. 2018.

26. European Association for Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol [Internet]. 2018 30 Apr 2018.; In press. Available from: http://www.easl.eu/medias/cpg/2018/Hepatocellular%20carcinoma/English-report.pdf.

27. Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet. 2018;Published Online February 9, 2018.

28. Tamai T, Hayato S, Hojo S, Suzuki T, Okusaka T, Ikeda K, et al. Dose Finding of Lenvatinib in Subjects With Advanced Hepatocellular Carcinoma Based on Population Pharmacokinetic and Exposure-Response Analyses. J Clin Pharmacol. 2017;57(9):1138-47.

29. Bruix J, Raoul JL, Sherman M, Mazzaferro V, Bolondi L, Craxi A, et al. Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma: subanalyses of a phase III trial. J Hepatol. 2012;57(4):821-9.

30. Jackson R, Psarelli EE, Berhane S, Khan H, Johnson P. Impact of Viral Status on Survival in Patients Receiving Sorafenib for Advanced Hepatocellular Cancer: A Meta-Analysis of Randomized Phase III Trials. J Clin Oncol. 2017;35(6):622-8.

31. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008;359(4):378-90.

32. Cheng AL, Kang YK, Chen Z, Tsao CJ, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol. 2009;10(1):25-34.

33. National Institute for Health and Care Excellence (NICE). Regorafenib for previously treated unresectable hepatocellular carcinoma: technology appraisal in development [ID991]2018 09 Feb 2018. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10112.

34. National Institute for Health and Care Excellence (NICE). Sorafenib for the treatment of advanced hepatocellular carcinoma: technology appraisal guidance [TA189]2010 09 Feb 2018. Available from: https://www.nice.org.uk/guidance/ta189.

35. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008;100(10):698-711.

36. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30(1):52-60.

37. Lencioni R, Montal R, Torre F, Park JW, Decaens T, Raoul JL, et al. Objective response by mRECIST as a predictor and potential surrogate end-point of overall survival in advanced HCC. J Hepatol. 2017;66(6):1166-72.

38. Meyer T, Palmer DH, Cheng AL, Hocke J, Loembe AB, Yen CJ. mRECIST to predict survival in advanced hepatocellular carcinoma: Analysis of two randomised phase II trials comparing nintedanib vs sorafenib. Liver Int. 2017;37(7):1047-55.

39. Food and Drug Administration. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics2007 3 Apr 2018. Available from: https://www.fda.gov/downloads/Drugs/Guidances/ucm071590.pdf.

40. European Medicine Agency (EMA). Appendix 1 to the guidance on the evaluation of anticancer medicinal products in man2012 3 Apr 2018. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137126.pdf.

41. Lencioni R, Kudo M, Finn RS, Qin S, Han KH, Ikeda K, et al. Independent imaging review (IIR) results in a phase 3 trial of lenvatinib (LEN) vs sorafenib (SOR) in first-line treatment of patients (pts) with unresectable hepatocellular carcinoma (uHCC). Poster presented at the 2018 Gastrointestinal Cancers Symposium, January 18-20, San Francisco, California, USA. 2018.

42. Abou-Alfa GK, Niedzwieski D, Knox JJ, Kaubisch A, Posey J, Tan BR, et al. Phase III randomized study of sorafenib plus doxorubicin versus sorafenib in patients with advanced hepatocellular carcinoma (HCC): CALGB 80802 (Alliance). J Clin Oncol. 2016;34(15_suppl):4003-.

43. Cainap C, Qin S, Huang WT, Chung IJ, Pan H, Cheng Y, et al. Linifanib versus Sorafenib in patients with advanced hepatocellular carcinoma: results of a randomized phase III trial. J Clin Oncol. 2015;33(2):172-9.

44. Cheng AL, Kang YK, Lin DY, Park JW, Kudo M, Qin S, et al. Sunitinib versus sorafenib in advanced hepatocellular cancer: results of a randomized phase III trial. J Clin Oncol. 2013;31(32):4067-75.

45. Johnson PJ, Qin S, Park JW, Poon RT, Raoul JL, Philip PA, et al. Brivanib versus sorafenib as first-line therapy in patients with unresectable, advanced hepatocellular carcinoma: results from the randomized phase III BRISK-FL study. J Clin Oncol. 2013;31(28):3517-24.

46. Assenat E, Boige V, Thézenas S, Pageaux GP, Peron JM, Becouarn Y, et al. Sorafenib (S) alone versus S combined with gemcitabine and oxaliplatin (GEMOX) in first-line treatment of advanced hepatocellular carcinoma (HCC): Final analysis of the randomized phase II GONEXT trial (UNICANCER/FFCD PRODIGE 10 trial). J Clin Oncol. 2013;31(15_suppl):4028.

47. Cheng AL, Guan Z, Chen Z, Tsao CJ, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients with advanced hepatocellular carcinoma according to baseline status: subset analyses of the phase III Sorafenib Asia-Pacific trial. Eur J Cancer. 2012;48(10):1452-65.

48. Haruna Y, Inoue A, Kawamoto S. Efficacy and safety of vitamin K with sorafenib combination treatment against hepatocellular carcinoma: Open-label, randomized phase II study. J Clin Oncol. 2017;35(15_suppl):e15665.

49. Ikeda M, Shimizu S, Sato T, Morimoto M, Kojima Y, Inaba Y, et al. Sorafenib plus hepatic arterial infusion chemotherapy with cisplatin versus sorafenib for advanced hepatocellular carcinoma: randomized phase II trial. Ann Oncol. 2016;27(11):2090-6.

50. Ji YX, Zhang ZF, Lan KT, Nie KK, Geng CX, Liu SC, et al. Sorafenib in liver function impaired advanced hepatocellular carcinoma. Chin Med Sci J. 2014;29(1):7-14.

51. Lee FA, Zee BC, Cheung FY, Kwong P, Chiang CL, Leung KC, et al. Randomized Phase II Study of the X-linked Inhibitor of Apoptosis (XIAP) Antisense AEG35156 in Combination With Sorafenib in Patients With Advanced Hepatocellular Carcinoma (HCC). Am J Clin Oncol. 2016;39(6):609-13.

52. Lencioni R, Kudo M, Ye SL, Bronowicki JP, Chen XP, Dagher L, et al. GIDEON (Global Investigation of therapeutic DEcisions in hepatocellular carcinoma and Of its treatment with sorafeNib): second interim analysis. Int J Clin Pract. 2014;68(5):609-17.

53. Palmer DH, Hussain SA, Smith AJ, Hargreaves S, Ma YT, Hull D, et al. Sorafenib for advanced hepatocellular carcinoma (HCC): impact of rationing in the United Kingdom. Br J Cancer. 2013;109(4):888-90.

54. Vilgrain V, Pereira H, Assenat E, Guiu B, Ilonca AD, Pageaux GP, et al. Efficacy and safety of selective internal radiotherapy with yttrium-90 resin microspheres compared with sorafenib in locally advanced and inoperable hepatocellular carcinoma (SARAH): an open-label randomised controlled phase 3 trial. Lancet Oncol. 2017;18(12):1624-36.

55. Yoon SM. Sorafenib versus transarterial chemoembolization plus external beam radiotherapy in patients with hepatocellular carcinoma with major vascular invasion (START): a randomized phase 2 trial. AASLD LiverLearning®. 2017;66(Supplement 1):729A.

56. Eisai Inc. Statistical analysis plan: E7080-G000-304. Data on file. 2016.

57. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013 - 5. The Reference Case 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case.

58. Cheng AL, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F. Phase III trial of lenvatinib (LEN) vs sorafenib (SOR) in first-line treatment of patients (pts) with unresectable hepatocellular carcinoma (uHCC). J Clin Oncol. 2017;35(15_suppl):4001.

59. School of Health and Related Research (ScHARR). Nice DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Report by the Decision Support Unit. 2013.

60. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 20132013 01 Feb 2018. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword.

61. British National Formulary (BNF). BNF Online 2017 [Dec 2017]. Available from: https://www.bnf.org/.

62. National Institute for Health and Care Excellence (NICE). Lenvatinib and sorafenib for treating differentiated thyroid cancer after radioactive iodine [ID1059]. NICE Guidance in development (ID1059) [Internet]. 2018 01 Feb 2018. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10101.

63. NHS Improvement. Reference Cost Collection: National Schedule of Reference Costs, 2016-17 - NHS trusts and NHS foundation trusts2017 01 Feb 2018. Available from: https://improvement.nhs.uk/resources/reference-costs/.

64. Curtis L, Burns A. Unit Costs of Health and Social Care. 2017.

65. Marie Curie Cancer Care. Understanding the cost of end of life care in different settings. 2012.

66. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. Nuffield Trust, 2014.

67. Sacco JJ, Botten J, Macbeth F, Bagust A, Clark P. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. PLoS One. 2010;5(1):e8933.

# 10 APPENDICES

## *10.1 Results of the company's clinical effectiveness systematic literature review SLR*

Table 51. Summary of included studies from the company's clinical effectiveness SLR, and selected excluded studies

| Study ID | Registration | Location (centres) | N | Intervention | Comparator | ERG note |
|---|---|---|---|---|---|---|
| **Studies included in the submission** | | | | | | |
| REFLECT[27] | NCT01761266 | Asia, Europe, North America | 954 | Lenvatinib | Sorafenib | Sole clinical evidence for this STA. |
| **Other studies included by the company that provide contextual evidence only** | | | | | | |
| SHARP[31] | NCT00105443 | Europe, Americas, Australia | 602 | Sorafenib | Placebo | Evidence for TA474. Child–Pugh A. |
| ASIA–PACIFIC[47] | NCT00492752. | Asia | 226 | Sorafenib | Placebo | Child–Pugh A. |
| Ji 2014[50] | Unknown | China | 189 | Sorafenib | Placebo | Child–Pugh B/C. |
| Lee 2016[51] | NCT00882869 | Hong Kong | 51 | AEG35156 + sorafenib | Sorafenib | Child–Pugh A. |
| GONEXT[46] | NCT00941967 | France | 94 | Sorafenib, gemcitabine, oxaliplatin | Sorafenib | Abstract only. |
| Haruna 2017[48] | UMIN000007855 | Japan | 44 | Sorafenib + Vitamin K | Sorafenib | Abstract only |
| Ikeda 2016[49] | UMIN000005703 | Japan | 108 | Sorafenib + cisplatin | Sorafenib | Child–Pugh A/B. |
| START[55] | NCT01901692 | South Korea | 90 | TACE + RT | Sorafenib | Liver-confined HCC with MVI |
| SARAH[54] | NCT01482442 | France | 467 | Sorafenib | SIRT | Child–Pugh A/B. |
| **Additional studies that the ERG consider relevant contextual evidence** | | | | | | |
| Cheng 2013[44] | NCT00699374 | Americas, Australia, Europe, Asia | 1074 | Sunitinib | Sorafenib | All Child–Pugh A population studies terminated early or primary endpoint not met. All report outcomes for the sorafenib arm with which to compare the clinical plausibility of REFLECT data. |
| Abou-Alfa 2016[42] | NCT01015833 | North America | 346 | Sorafenib + doxorubicin | Sorafenib | |
| Cainap 2015[43] | NCT01009593 | Americas, Australia, Europe, Asia | 1035 | Linifanib | Sorafenib | |
| Johnson 2013[45] | NCT00858871 | Americas, Australia, Europe, Asia | 1155 | Brivanib | Sorafenib | |
| GIDEON[52] | N/A | International | 1571 | Sorafenib (global prospective study) | | Ongoing real-world study of sorafenib, including those excluded from RCTs. |
| Palmer 2013[53] | N/A | UK | 133 | Sorafenib (retrospective study) | | Compares those who were eligible for sorafenib with those who were not |

Abbreviations: MVI, macroscopic vascular invasion; RT, external beam radiotherapy; SIRT, selective internal radiotherapy with yttrium-90 resin microspheres; SLR, systematic literature review; STA, Single Technology Appraisal; TACE, transcatheter arterial chemoembolisation.

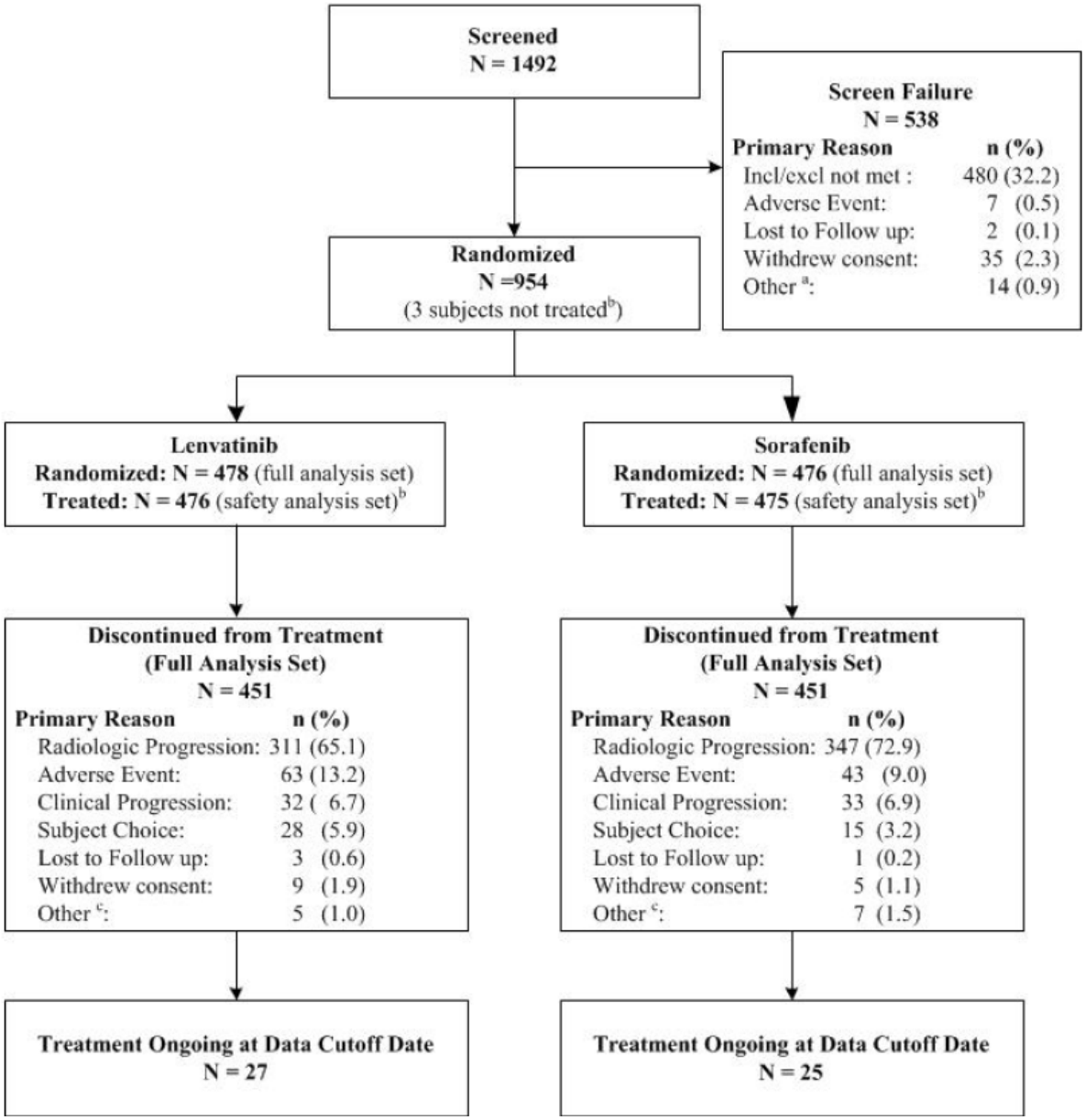
## *10.2 Quality assessment of REFLECT*

Table 52. Company's quality assessment of REFLECT with validation by the ERG (adapted from CS Appendix B.2.5, Table 10)

| Quality indicator | Company's assessment | ERG comment |
|---|---|---|
| Was randomisation carried out appropriately? | Yes. Patients were assigned to treatment based on a computer-generated randomisation scheme. Allocation of randomisation numbers was performed using an interactive voice/web response system based on stratification factors. | Randomisation was carried out appropriately; low risk of bias. |
| Was the concealment of treatment allocation adequate? | Yes. Patients were assigned to treatment based on a computer-generated randomisation scheme. | Treatment allocation was concealed appropriately; low risk of bias. |
| Were the groups similar at the outset of the study in terms of prognostic factors? | Yes. Stratified randomisation was performed which generates separate schedules for subsets of participants defined by potentially important prognostic factors (region; macroscopic portal vein invasion or extrahepatic spread of both; ECOG; and body weight). Demographic and other baseline characteristics were generally well balanced between treatment arms with the exception of some notable differences in disease characteristics as described in Section B.2.3.2. | The clinical relevance of imbalances highlighted by the company is uncertain, namely alpha-fetoprotein levels and hepatitis C aetiology. The ERG considers there to be an unclear risk of bias (see ERG critique, Section 4.2.2). |
| Were the care providers, participants and outcome assessors blind to treatment allocation? | No. This was an open label trial. This open-label design was chosen in the interests of patient safety (see Section B.2.13.2). OS was the primary outcome and assessment bias was therefore unlikely. For the secondary endpoints of PFS, TTP and ORR, tumour assessments were performed by the investigator and there was therefore a risk of bias. However, PFS results from a post-hoc, blinded, retrospective IIR using both mRECIST and RECIST 1.1 criteria were consistent with the investigator assessments of PFS (see Section B.2.13.1). | REFLECT was open-label but alternative IIR results were available for some outcomes to minimise bias, and the study included a Data Integrity Protection Plan, to mask data fields from the clinical and statistical team. Patient reported outcomes, investigator-assessed outcomes and adverse events (although these recorded regardless of whether they were regarded to be related to treatment) are considered at high risk of bias. |
| Were there any unexpected imbalances in drop-outs between groups? | No. The overall drop-outs were generally well-balanced between treatment arms and the primary reasons for treatment discontinuation were also well-balanced between treatment arms (as per the detailed CONSORT diagram presented in Appendix D, Section D.2. | A similar number of patients discontinued each treatment, but radiological progression was a more common reason in the sorafenib arm, and adverse events and patient choice were more common in the lenvatinib arm (CS Appendix D.2, Figure 2).<br>The ERG considered there to be a moderate risk of bias. |
| Is there any evidence to suggest that the authors measured more outcomes than they reported? | No there are no reported changes to the planned analysis in the CSR and the outcomes listed in the study protocol and CSR are consistent. | Results for all relevant outcomes and analyses described in the study protocol were available. The ERG considered REFLECT to be at low risk of reporting bias. |
| Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data? | Yes. The primary analysis was based on a full analysis set (intent to treat analysis set) including all patients who were randomised. At the time of the primary analysis if a patient was alive or did not have disease progression, they were censored. OS patients lost to follow-up were censored at the last date the patient was known to be alive, and patients who remained alive will be censored at the time of data cut-off. For PFS the censoring rules followed FDA guidance (45). | Efficacy outcomes included all randomised patients, and safety outcomes included all but 3 patients who did not receive study treatment.<br>OS censoring was appropriate but PFS may be biased because patients were censored if they discontinued treatment for a reason other than disease progression, which was more common in the lenvatinib group (CS Table 14). |
| Abbreviations: CONSORT, Consolidated Standards of Reporting Trials; CSR, clinical study report'; ECOG, Eastern Cooperative Oncology Group; FDA, United States Food and Drug Administration; mRECIST, modified Response Evaluation Criteria in Solid Tumours; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; TTP, time to progression. | | |

## *10.3 Participant flow*

Figure 22. REFLECT CONSORT diagram (reproduced from CS Appendix D.2, Figure 2)




[a] Other reasons for screening failure varied, with the most common reasons being expiration of the 21-day screening window (n=4) and worsening of the patient's condition (n=3); [b]Two patients randomised to lenvatinib were not treated as they were randomised in error and 1 patient randomised to sorafenib chose not to receive treatment; therefore, the Safety Analysis Set includes 476 patients in the lenvatinib arm and 475 patients in the sorafenib arm; [c]"Other" reasons for discontinuation in the lenvatinib arm included randomisation in error (n=2; not treated); patient required surgery (n=2) and investigator choice (n=1). In the sorafenib arm, "other" reasons included investigator choice (n=5); need for a prohibited medication (warfarin; 1 patient); and discontinuation to undergo liver transplantation (n=1).

## *10.4 Baseline characteristics*

Table 53. REFLECT demographic and baseline characteristics for the full population; reproduced from CS Table 6)

| | Lenvatinib | | | Sorafenib |
|---|---|---|---|---|
| | 8 mg† (N = 151) | 12 mg† (N = 327) | Total (N = 478) | (N = 476) |
| Age (years) | | | | |
| Mean (SD) | 63.1 (12.30) | 60.4 (11.32) | 61.3 (11.69) | 61.2 (12.01) |
| Median | 65.0 | 62.0 | 63.0 | 62.0 |
| Q1, Q3 | 56.0, 72.0 | 53.0, 68.0 | 54.0, 70.0 | 54.0, 70.0 |
| Min, max | 20, 86 | 24, 88 | 20, 88 | 22, 88 |
| Age group (years), n (%) | | | | |
| <65 | 69 (45.7) | 201 (61.5) | 270 (56.5) | 283 (59.5) |
| ≥65 to <75 | 56 (37.1) | 94 (28.7) | 150 (31.4) | 126 (26.5) |
| ≥75 | 26 (17.2) | 32 (9.8) | 58 (12.1) | 67 (14.1) |
| Sex, n (%) | | | | |
| Male | 106 (70.2) | 299 (91.4) | 405 (84.7) | 401 (84.2) |
| Female | 45 (29.8) | 28 (8.6) | 73 (15.3) | 75 (15.8) |
| Region, n (%) | | | | |
| Western‡ | 21 (13.9) | 136 (41.6) | 157 (32.8) | 157 (33.0) |
| Asia–Pacific‡ | 130 (86.1) | 191 (58.4) | 321 (67.2) | 319 (67.0) |
| Race, n (%) | | | | |
| White | 17 (11.3) | 118 (36.1) | 135 (28.2) | 141 (29.6) |
| Black/African American | 0 (0.0) | 7 (2.1) | 7 (1.5) | 6 (1.3) |
| Asian | 134 (88.7) | 200 (61.2) | 334 (69.9) | 326 (68.5) |
| American Indian/Alaskan native | 0 (0.0) | 1 (0.3) | 1 (0.2) | 0 (0.0) |
| Hawaiian/Pacific Islander | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (0.2) |
| Other | 0 (0.0) | 1 (0.3) | 1 (0.2) | 2 (0.4) |
| Weight (kg) | | | | |
| Mean (SD) | 52.7 (4.90) | 75.9 (14.40) | 68.6 (16.32) | 68.1 (13.90) |
| Median | 53.0 | 72.0 | 66.2 | 67.0 |
| Q1, Q3 | 50.0, 56.5 | 65.3, 82.0 | 57.0, 76.2 | 57.6, 77.0 |
| Min, max | 39, 60 | 60, 142 | 39, 142 | 39, 123 |
| Body weight group | | | | |
| <60 kg | 151 (100.0) | 2 (0.6) | 153 (32.0) | 146 (30.7) |
| ≥60 kg | 0 (0.0) | 325 (99.4) | 325 (68.0) | 330 (69.3) |
| ECOG PS, n (%) | | | | |
| 0 | 93 (61.6) | 211 (64.5) | 304 (63.6) | 301 (63.2) |
| 1 | 58 (38.4) | 116 (35.5) | 174 (36.4) | 175 (36.8) |
| NYHA classification, n (%) | | | | |
| I | 4 (2.6) | 33 (10.1) | 37 (7.7) | 44 (9.2) |
| II | 1 (0.7) | 7 (2.1) | 8 (1.7) | 6 (1.3) |
| Not applicable | 145 (96.0) | 287 (87.8) | 432 (90.4) | 426 (89.5) |
| Missing | 1 (0.7) | 0 (0.0) | 1 (0.2) | 0 (0.0) |
| Child–Pugh score, n (%) | | | | |
| 5 | 111 (73.5) | 257 (78.6) | 368 (77.0) | 357 (75.0) |
| 6 | 40 (26.5) | 67 (20.5) | 107 (22.4) | 114 (23.9) |

| | | | | |
|---|---|---|---|---|
| 7 | 0 (0.0) | 3 (0.9) | 3 (0.6) | 4 (0.8) |
| 8 | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (0.2) |
| Macroscopic portal vein invasion, n (%) | | | | |
| Yes | 38 (25.2) | 71 (21.7) | 109 (22.8) | 90 (18.9) |
| No | 113 (74.8) | 256 (78.3) | 369 (77.2) | 386 (81.1) |
| Extrahepatic spread, n (%) | | | | |
| Yes | 91 (60.3) | 200 (61.2) | 291 (60.9) | 295 (62.0) |
| No | 60 (39.7) | 127 (38.8) | 187 (39.1) | 181 (38.0) |
| Macroscopic portal vein invasion, extrahepatic spread, or both, n (%) | | | | |
| Yes | 105 (69.5) | 224 (68.5) | 329 (68.8) | 336 (70.6) |
| No | 46 (30.5) | 103 (31.5) | 149 (31.2) | 140 (29.4) |
| Underlying cirrhosis, n (%) | | | | |
| Yes§ | 75 (49.7) | 168 (51.4) | 243 (50.8) | 231 (48.5) |
| No | 76 (50.3) | 159 (48.6) | 235 (49.2) | 245 (51.5) |

Abbreviations: CRF, case report form; ECOG PS, Eastern Cooperative Oncology Group Performance Status; HCC, hepatocellular carcinoma; kg, kilograms; NYHA, New York Heart Association; Q, quartile; SD, standard deviation.
† 8mg and 12 mg were the lenvatinib starting doses based on patients' body weight (<60 kg, ≥60 kg) at Baseline; ‡ Western region consists of North America and Europe including Russia and Israel; Asia–Pacific region consists of China, Hong Kong, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand; § The proportion of patients with underlying cirrhosis at baseline (49.7%) was likely underestimated as this information was collected on the CRF under medical history, and the presence or absence of cirrhosis was verified only when needed to confirm the clinical diagnosis of HCC.

Table 54. REFLECT disease history and characteristics (reproduced from CS Table 7)

| | **Lenvatinib N=478** | **Sorafenib N=476** |
|---|---|---|
| Time since first diagnosis (months) | | |
| Mean (SD) | 21.1 (30.17) | 23.3 (34.66) |
| Median | 8.2 | 9.0 |
| Q1, Q3 | 1.6, 27.3 | 2.0, 27.2 |
| Min, max | 0, 180 | 0, 250 |
| Age at first diagnosis (years) | | |
| Mean (SD) | 59.6 (11.57) | 59.3 (11.54) |
| Median | 61.0 | 60.0 |
| Q1, Q3 | 52.0, 68.0 | 52.0, 67.0 |
| Min, max | 15, 87 | 20, 85 |
| BCLC stage, n (%) | | |
| B: Intermediate stage | 104 (21.8) | 92 (19.3) |
| C: Advanced stage | 374 (78.2) | 384 (80.7) |
| Involved disease sites†, n (%) | | |
| Liver | 441 (92.3) | 430 (90.3) |
| Lung | 163 (34.1) | 144 (30.3) |
| Lymph nodes | 127 (26.6) | 141 (29.6) |
| Bone | 51 (10.7) | 43 (9.0) |
| Other | 82 (17.2) | 97 (20.4) |
| Number of involved disease sites per patient, n (%) | | |
| 1 | 207 (43.3) | 207 (43.5) |
| 2 | 167 (34.9) | 183 (38.4) |
| ≥3 | 103 (21.5) | 86 (18.1) |

| | | |
|---|---|---|
| Factor of carcinogenesis‡, n (%) | | |
| Hepatitis B | 251 (52.5) | 228 (47.9) |
| Hepatitis C | 91 (19.0) | 126 (26.5) |
| Alcohol | 36 (7.5) | 21 (4.4) |
| Other | 38 (7.9) | 32 (6.7) |
| Unknown | 62 (13.0) | 69 (14.5) |
| Baseline alpha-fetoprotein level (ng/mL) | | |
| Mean (SD) | 17507.5 (105137.39) | 16678.5 (94789.46) |
| Median | 133.1 | 71.2 |
| Q1, Q3 | 8.0, 3730.6 | 5.2, 1081.8 |
| Min, max | 0, 1567470 | 0, 1446396 |
| Baseline alpha-fetoprotein group, n (%) | | |
| <200 ng/mL | 255 (53.3) | 286 (60.1) |
| ≥ 200 ng/mL | 222 (46.4) | 187 (39.3) |
| Missing | 1 (0.2) | 3 (0.6) |
| Ammonia level (μg/dL) | | |
| Mean (SD) | 38.2 (29.98) | 36.7 (32.90) |
| Median | 31.8 | 30.0 |
| Q1, Q3 | 22.0, 45.0 | 21.0, 42.3 |
| Min, max | 4, 246 | 4, 473 |
| Concomitant systemic antiviral therapy for Hepatitis B or Hepatitis C, n (%) | 163 (34.1) | 149 (31.3) |

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma; Q, quartile; SD, standard deviation.
† Patients may be counted in more than 1 disease site; ‡ Based on the combined data from HCC diagnosis and medical history. Patients may be counted in more than 1 factor.

Table 55. REFLECT demographcis, baseline characteristics and disease history by region (reproduced from CS Appendix, Table 4)

| | **Lenvatinib** | | **Sorafenib** | |
|---|---|---|---|---|
| | **Asia–Pacific N=321** | **Western N=157** | **Asia–Pacific N=319** | **Western N=157** |
| Age (years) | | | | |
| Mean (SD) | 60.0 (11.76) | 63.8 (11.15) | 60.2 (11.87) | 63.3 (12.06) |
| Median | 60.0 | 64.0 | 61.0 | 64.0 |
| Q1, Q3 | 51.0, 69.0 | 58.0, 72.0 | 52.0, 68.0 | 58.0, 72.0 |
| Min, max | 30, 86 | 20, 88 | 26, 88 | 22, 85 |
| Age group (years), n (%) | | | | |
| <65 | 191 (59.5) | 79 (50.3) | 202 (63.3) | 81 (51.6) |
| ≥65 to <75 | 95 (29.6) | 55 (35.0) | 78 (24.5) | 48 (30.6) |
| ≥75 | 35 (10.9) | 23 (14.6) | 39 (12.2) | 28 (17.8) |
| Sex, n (%) | | | | |
| Male | 277 (86.3) | 128 (81.5) | 269 (84.3) | 132 (84.1) |
| Female | 44 (13.7) | 29 (18.5) | 50 (15.7) | 25 (15.9) |
| Body weight group | | | | |
| <60 kg | 132 (41.1) | 21 (13.4) | 127 (39.8) | 19 (12.1) |
| ≥60 kg | 189 (58.9) | 136 (86.6) | 192 (60.2) | 138 (87.9) |
| Ammonia level (ng/dL) | | | | |

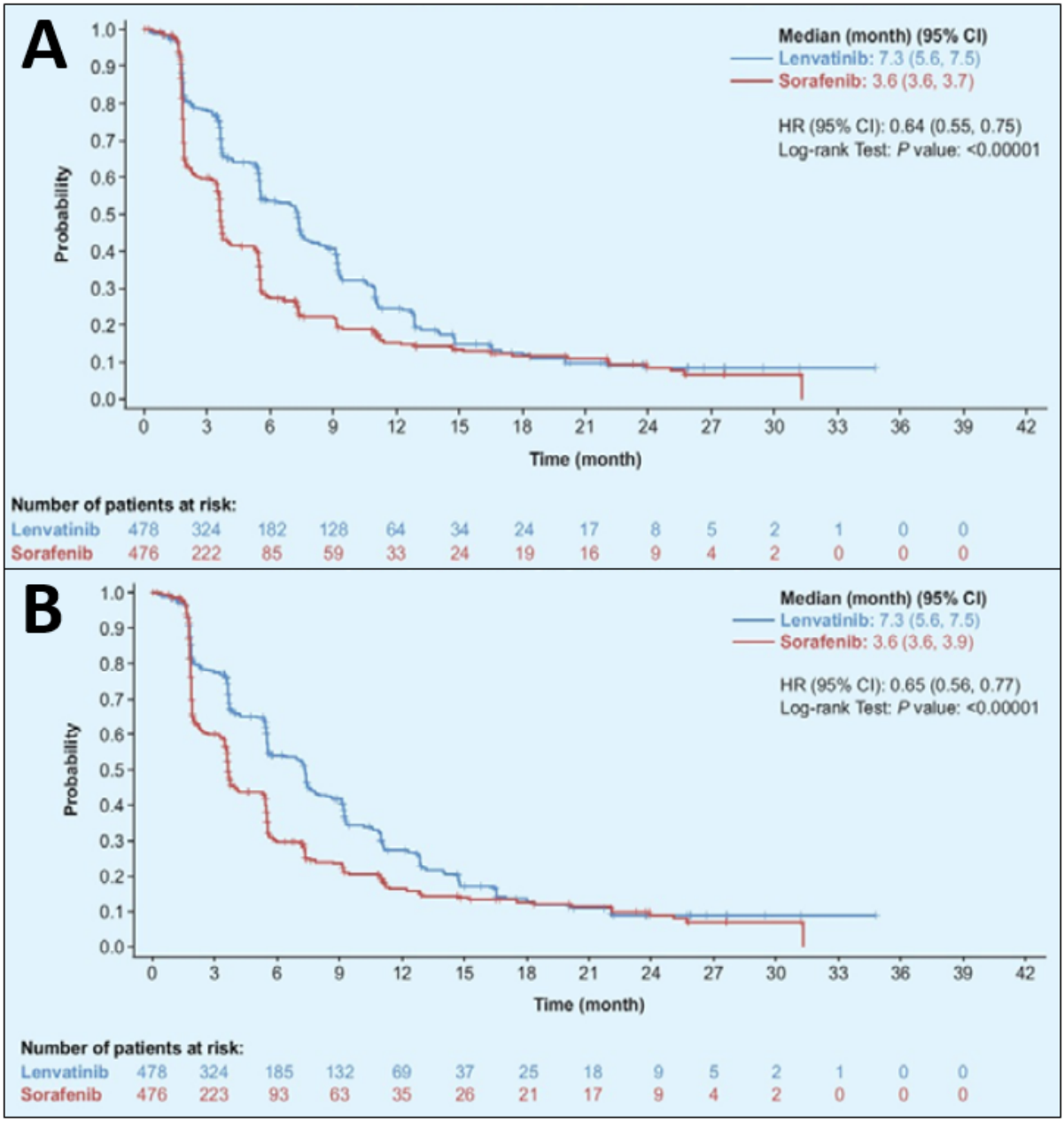
| | | | | |
|---|---|---|---|---|
| N | 315 | 140 | 309 | 144 |
| Mean (SD) | 36.3 (32.54) | 42.5 (22.74) | 32.9 (24.64) | 45.0 (44.86) |
| Median | 28.6 | 38.5 | 28.0 | 37.9 |
| Q1, Q3 | 20.0, 42.0 | 27.0, 52.4 | 19.0, 39.0 | 26.1, 48.6 |
| Min, max | 4, 246 | 7, 132 | 4, 199 | 6, 473 |
| ECOG PS, n (%) | | | | |
| 0 | 206 (64.2) | 98 (62.4) | 204 (63.9) | 97 (61.8) |
| 1 | 115 (35.8) | 59 (37.6) | 115 (36.1) | 60 (38.2) |
| NYHA classification | | | | |
| I | 8 (2.5) | 29 (18.5) | 12 (3.8) | 32 (20.4) |
| II | 0 (0.0) | 8 (5.1) | 1 (0.3) | 5 (3.2) |
| Not applicable | 312 (97.2) | 120 (76.4) | 306 (95.9) | 120 (76.4) |
| Missing | 1 (0.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Child–Pugh score, n (%) | | | | |
| 5 | 255 (79.4) | 113 (72.0) | 256 (80.3) | 101 (64.3) |
| 6 | 65 (20.2) | 42 (26.8) | 63 (19.7) | 51 (32.5) |
| Macroscopic portal vein invasion, n (%) | | | | |
| Yes | 74 (23.1) | 35 (22.3) | 64 (20.1) | 26 (16.6) |
| No | 247 (76.9) | 122 (77.7) | 255 (79.9) | 131 (83.4) |
| Extrahepatic spread, n (%) | | | | |
| Yes | 193 (60.1) | 98 (62.4) | 193 (60.5) | 102 (65.0) |
| No | 128 (39.9) | 59 (37.6) | 126 (39.5) | 55 (35.0) |
| Macroscopic portal vein invasion, extrahepatic spread, or both, n (%) | | | | |
| Yes | 220 (68.5) | 109 (69.4) | 221 (69.3) | 115 (73.2) |
| No | 101 (31.5) | 48 (30.6) | 98 (30.7) | 42 (26.8) |
| Underlying cirrhosis†, n (%) | | | | |
| Yes | 180 (56.1) | 63 (40.1) | 169 (53.0) | 62 (39.5) |
| No | 141 (43.9) | 94 (59.9) | 150 (47.0) | 95 (60.5) |
| Time since first diagnosis (months) | | | | |
| Mean (SD) | 23.9 (31.70) | 15.4 (25.93) | 26.8 (37.61) | 16.1 (26.40) |
| Median | 10.3 | 5.6 | 11.4 | 5.3 |
| Q1, Q3 | 1.6, 35.0 | 1.6, 17.0 | 3.1, 34.2 | 1.5, 14.6 |
| Min, max | 0, 180 | 0, 150 | 0, 250 | 0, 144 |
| Age at first diagnosis (years) | | | | |
| Mean (SD) | 58.1 (11.45) | 62.6 (11.26) | 58.0 (11.23) | 62.0 (11.72) |
| Median | 59.0 | 63.0 | 59.0 | 63.0 |
| Q1, Q3 | 50.0, 67.0 | 57.0, 70.0 | 51.0, 66.0 | 57.0, 70.0 |
| Min, max | 27, 86 | 15, 87 | 25, 82 | 20, 85 |
| BCLC stage, n (%) | | | | |
| B: Intermediate stage | 70 (21.8) | 34 (21.7) | 65 (20.4) | 27 (17.2) |
| C: Advanced stage | 251 (78.2) | 123 (78.3) | 254 (79.6) | 130 (82.8) |
| Involved disease sites‡, n (%) | | | | |
| Liver | 295 (91.9) | 146 (93.0) | 285 (89.3) | 145 (92.4) |
| Lung | 117 (36.4) | 46 (29.3) | 110 (34.5) | 34 (21.7) |
| Lymph nodes | 76 (23.7) | 51 (32.5) | 83 (26.0) | 58 (36.9) |
| Bone | 40 (12.5) | 11 (7.0) | 31 (9.7) | 12 (7.6) |
| Other | 46 (14.3) | 36 (22.9) | 64 (20.1) | 33 (21.0) |

| | | | | |
|---|---|---|---|---|
| Number of involved disease sites per patient, n (%) | | | | |
| 1 | 144 (44.9) | 63 (40.1) | 145 (45.5) | 62 (39.5) |
| 2 | 110 (34.3) | 57 (36.3) | 112 (35.1) | 71 (45.2) |
| ≥3 | 67 (20.9) | 36 (22.9) | 62 (19.4) | 24 (15.3) |
| Factor of carcinogenesis, n (%) | | | | |
| Hepatitis B | 212 (66.0) | 39 (24.8) | 197 (61.8) | 31 (19.7) |
| Hepatitis C | 50 (15.6) | 41 (26.1) | 70 (21.9) | 56 (35.7) |
| Alcohol | 17 (5.3) | 19 (12.1) | 8 (2.5) | 13 (8.3) |
| Other | 17 (5.3) | 21 (13.4) | 11 (3.4) | 21 (13.4) |
| Unknown | 25 (7.8) | 37 (23.6) | 33 (10.3) | 36 (22.9) |
| Baseline alpha-fetoprotein level (ng/mL) | | | | |
| N | 315 | 156 | 309 | 154 |
| Mean (SD) | 10078.5 (39198.91) | 32508.2 (173397.79) | 16460.6 (77052.08) | 17115.6 (123204.55) |
| Median | 168.0 | 78.9 | 100.6 | 27.0 |
| Q1, Q3 | 11.1, 4186.8 | 5.2, 3102.1 | 7.5, 1416.0 | 3.6, 622.0 |
| Min, max | 1, 518877 | 0, 1567470 | 1, 925460 | 0, 1446396 |
| Baseline serum alpha-fetoprotein group, n (%) | | | | |
| <200 ng/mL | 164 (51.1) | 91 (58.0) | 182 (57.1) | 104 (66.2) |
| ≥200 ng/mL | 157 (48.9) | 65 (41.4) | 137 (42.9) | 50 (31.8) |
| Missing | 0 (0.0) | 1 (0.6) | 0 (0.0) | 3 (1.9) |

Abbreviations: BCLC, Barcelona Clinic Liver Centre; CS, company submission; ECOG PS, Eastern Cooperative Oncology Group performance status; n, number of patients; NYHA, New York Heart Association; Q, quartile; SD, standard deviation.

## *10.5 Supportive analyses*

Figure 23. Kaplan–Meier curves for IIR PFS according to mRECIST (A) and RECIST v1.1 (B; adapted from CS Appendix, Figures 6 and 7)




Abbreviations: CI, confidence interval; CS, company submission; HR hazard ratio; IIR independent imaging review; mRECIST, modified response evaluation criteria in solid tumours; PFS, progression-free survival; RECIST v1.1, response evaluation criteria in solid tumours version 1.1.

## *10.6 Subgroup analyses*

Table 56. Overall survival subgroup analyses for stratification factors and baseline imbalances (adapted from CS Appendices, Table 5)

| | | Lenvatinib (N = 478) | | | Sorafenib (N = 476) | | | HR, lenvatinib vs sorafenib |
|---|---|---|---|---|---|---|---|---|
| | | N | Events | Median months (95% CI) | N | Events | Median months (95% CI) | |
| Overall | | 478 | 351 | 13.6 (12.1 to 14.9) | 476 | 350 | 12.3 (10.4, 13.9) | 0.92 (0.79, 1.06) |
| Age | <65 years | 270 | 203 | 12.4 (10.9, 14.1) | 283 | 204 | 11.4 (9.7, 13.9) | 0.94 (0.77, 1.15) |
| | ≥65 to <75 years | 150 | 107 | 15.3 (12.5, 20.7) | 126 | 94 | 12.3 (10.3, 14.3) | 0.84 (0.63, 1.12) |
| | ≥75 years | 58 | 41 | 13.4 (10.1, 20.9) | 67 | 52 | 17.8 (8.4, 19.7) | 0.84 (0.53, 1.33) |
| Sex | Male | 405 | 293 | 13.4 (12.0, 14.4) | 401 | 293 | 12.4 (10.5, 14.2) | 0.91 (0.77, 1.07) |
| | Female | 73 | 58 | 15.3 (10.4, 19.2) | 75 | 57 | 11.4 (8.0, 14.6) | 0.84 (0.56, 1.26) |
| Region | Asia–Pacific | 321 | 243 | 13.5 (11.7, 15.1) | 319 | 248 | 11.0 (9.6, 12.5) | 0.86 (0.72, 1.02) |
| | Western | 157 | 108 | 13.6 (11.5, 17.7) | 157 | 102 | 14.2 (11.9, 18.0) | 1.08 (0.82, 1.42) |
| Macroscopic portal vein invasion or extrahepatic spread, or both* | Yes | 329 | 250 | 11.5 (9.4, 13.4) | 336 | 259 | 9.8 (8.7, 11.4) | 0.87 (0.73, 1.04) |
| | No | 149 | 101 | 18.0 (14.6, 20.7) | 140 | 91 | 18.0 (15.7, 21.3) | 1.05 (0.79, 1.40) |
| ECOG PS | 0 | 304 | 221 | 14.6 (13.4, 18.0) | 301 | 223 | 12.8 (11.6, 14.6) | 0.88 (0.73, 1.06) |
| | 1 | 174 | 130 | 10.7 (8.6, 13.5) | 175 | 127 | 10.3 (8.8, 13.5) | 0.97 (0.76, 1.25) |
| Body weight | <60 kg | 153 | 110 | 13.4 (10.5, 15.7) | 146 | 113 | 10.3 (8.7, 15.9) | 0.85 (0.65, 1.11) |
| | ≥60 kg | 325 | 241 | 13.7 (12.0, 15.6) | 330 | 237 | 12.5 (11.1, 14.2) | 0.95 (0.79, 1.14) |
| AFP at baseline | <200 ng/mL | 255 | 167 | 19.5 (15.0, 22.4) | 286 | 193 | 16.3 (14.2, 18.8) | 0.91 (0.74, 1.12) |
| | ≥200 ng/mL | 222 | 183 | 10.4 (8.7, 12.0) | 187 | 154 | 8.2 (7.0, 9.7) | 0.78 (0.63, 0.98) |
| Aetiology | Hepatitis B | 259 | 196 | 13.4 (11.6, 14.6) | 244 | 186 | 10.2 (8.6, 12.4) | 0.83 (0.68, 1.02) |
| | Hepatitis C | 103 | 75 | 15.3 (11.2, 20.2) | 135 | 97 | 14.1 (11.4, 18.5) | 0.91 (0.66, 1.26) |
| | Alcohol | 33 | 22 | 14.1 (5.9, 22.5) | 23 | 15 | 11.9 (7.8, 24.0) | 1.03 (0.47, 2.28) |
| BCLC staging | Stage B | 104 | 71 | 18.5 (14.6, 21.2) | 92 | 65 | 17.3 (14.6, 18.8) | 0.91 (0.65, 1.28) |
| | Stage C | 374 | 280 | 11.8 (10.4, 13.7) | 384 | 285 | 10.3 (9.4, 12.3) | 0.92 (0.77, 1.08) |

Abbreviations: AFP, alpha fetoprotein; BCLC, Barcelona Clinic Liver Cancer; CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; HBV, hepatitis B virus; HCV, hepatitis C virus; HR, hazard ratio.
Subgroup analyses not presented here included region by China versus rest of world and macroscopic portal vein invasion and extrahepatic spread as separate variables. Subgroup analyses to explore use of subsequent anti-cancer procedures and medications are discussed in Section 4.3.6.

Table 57. PFS subgroup analyses for stratification factors and baseline imbalances (adapted from CS Appendices, Table 6)

| | | Lenvatinib (N = 478) | | | Sorafenib (N = 476) | | | HR, lenvatinib vs sorafenib |
|---|---|---|---|---|---|---|---|---|
| | | N | Events | Median months (95% CI) | N | Events | Median months (95% CI) | |
| Overall | | 478 | 349 | 7.4 (6.9, 8.8) | 476 | 367 | 3.7 (3.6, 4.6) | 0.66 (0.57, 0.77) |
| Age | <65 years | 270 | 201 | 7.3 (5.6, 9.0) | 283 | 223 | 3.6 (3.5, 3.7) | 0.67 (0.55, 0.82) |
| | ≥65 to <75 years | 150 | 113 | 7.4 (6.2, 9.2) | 126 | 100 | 5.3 (3.6, 5.6) | 0.61 (0.46, 0.82) |
| | ≥75 years | 58 | 35 | 7.8 (6.1, 10.6) | 67 | 44 | 5.5 (3.7, 5.6) | 0.59 (0.34, 1.02) |
| Sex | Male | 405 | 298 | 7.4 (7.0, 9.0) | 401 | 308 | 3.7 (3.6, 4.1) | 0.66 (0.56, 0.77) |
| | Female | 73 | 51 | 7.4 (5.4, 9.2) | 75 | 59 | 4.6 (3.5, 7.2) | 0.75 (0.49, 1.13) |
| Region | Asia–Pacific | 321 | 249 | 7.3 (5.6, 9.0) | 319 | 264 | 3.6 (3.4, 3.7) | 0.61 (0.51, 0.73) |
| | Western | 157 | 100 | 7.4 (6.2, 9.3) | 157 | 103 | 5.5 (3.7, 7.4) | 0.81 (0.61, 1.08) |
| Macroscopic portal vein invasion or extrahepatic spread, or both* | Yes | 329 | 246 | 7.3 (5.6, 7.5) | 336 | 265 | 3.6 (3.3, 3.7) | 0.64 (0.54, 0.77) |
| | No | 149 | 103 | 9.2 (7.3, 11.0) | 140 | 102 | 5.6 (4.1, 7.3) | 0.73 (0.55, 0.97) |
| ECOG PS | 0 | 304 | 220 | 7.4 (7.2, 9.1) | 301 | 233 | 3.7 (3.5, 4.6) | 0.63 (0.52, 0.76) |
| | 1 | 174 | 129 | 7.3 (5.5, 8.4) | 175 | 134 | 3.7 (3.6, 5.5) | 0.70 (0.55, 0.90) |
| Body weight | <60 kg | 153 | 111 | 7.4 (5.4, 9.2) | 146 | 121 | 3.6 (3.5, 4.1) | 0.61 (0.46, 0.79) |
| | ≥60 kg | 325 | 238 | 7.4 (6.9, 9.0) | 330 | 246 | 3.7 (3.6, 5.4) | 0.69 (0.58, 0.83) |
| AFP at baseline | <200 ng/mL | 255 | 186 | 9.0 (7.4, 9.2) | 286 | 209 | 5.4 (3.7, 5.7) | 0.68 (0.55, 0.83) |
| | ≥200 ng/mL | 222 | 163 | 5.5 (4.2, 7.2) | 187 | 157 | 2.4 (1.9, 3.6) | 0.59 (0.47, 0.75) |
| Aetiology | Hepatitis B | 259 | 205 | 7.3 (5.6, 9.1) | 244 | 199 | 3.6 (2.6, 3.6) | 0.62 (0.50, 0.75) |
| | Hepatitis C | 103 | 70 | 7.4 (5.5, 9.1) | 135 | 103 | 5.3 (3.6, 5.6) | 0.78 (0.56, 1.09) |
| | Alcohol | 33 | 19 | 8.8 (5.5, 14.7) | 23 | 18 | 3.9 (2.0, 7.2) | 0.27 (0.11, 0.66) |
| BCLC staging | Stage B | 104 | 72 | 9.1 (7.2, 9.4) | 92 | 66 | 5.5 (3.6, 7.2) | 0.70 (0.50, 0.99) |
| | Stage C | 374 | 277 | 7.3 (5.6, 7.6) | 384 | 301 | 3.7 (3.6, 3.7) | 0.63 (0.53, 0.75) |

Abbreviations: AFP, alpha fetoprotein; BCLC, Barcelona Clinic Liver Cancer; CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group Performance Status; HBV, hepatitis B virus; HCV, hepatitis C virus; HR, hazard ratio.
Subgroup analyses not presented here included region by China versus rest of world and macroscopic portal vein invasion and extrahepatic spread as separate variables. Subgroup analyses to explore use of subsequent anti-cancer procedures and medications are discussed in Section 4.3.6.

# National Institute for Health and Care Excellence
# Centre for Health Technology Evaluation

## Pro-forma Response

## ERG report

### Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma [ID1089]

You are asked to check the ERG report from BMJ Group to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Wednesday 16 May 2018** using the below proforma 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 proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

**Issue 1 CiC marking**

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current wording (Section 1.1 page 11 and Section 3.2, page 31)<br>The company of lenvatinib ([redacted]Eisai Co., Ltd) | All mention of branding to be marked CiC ([redacted] | This is pending EMA approval | Mark-up changes to be made my NICE. |

## Issue 2 Readability

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current wording (Section 1.1, page 11):<br>Lenvatinib's marketing authorisation covers its use in combination with everolimus for advanced renal cell carcinoma and differentiated thyroid carcinoma | Lenvatinib's marketing authorisation covers its use for differentiated thyroid carcinoma and in combination with everolimus for advanced renal cell carcinoma | The current wording implies that lenvatinib is licensed for use in combination with everolimus for differentiated thyroid carcinoma | Wording amended as described. |

## Issue 3 CiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current wording (Section 1.2, page 14):<br>Drug withdrawals due to TEAEs (19.7% vs 14.5% for lenvatinib and sorafenib, respectively), dose modifications due to TEAEs (61.8% vs 55.6%) | Drug withdrawals due to TEAEs ([redacted] for lenvatinib and sorafenib, respectively), dose modifications due to TEAEs ([redacted]) | These data were marked as CiC in the company submission | Mark-up changes to be made my NICE. |

### Issue 4 Transcription error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Table 7 (page 46)<br>Any anti-cancer procedure for the lenvatinib full population is listed as 122 (25.1) | The correct numbers should be 122 (25.5) | Error in the percentage reported | Correction made |

### Issue 5 Transcription error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Table 7 (page 46)<br>Current data for the immunostimulants row reads as follows<br>Immunostimulants (interferon) [redacted] | Immunostimulants (interferon) [redacted] | Error in percentages for the lenvatinib and sorafenib full population cells, and number and percentage for the lenvatinib Western population cell | Correction made |

### Issue 6 Transcription error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Table 7 (page 46<br>Current data for the antineoplastic and immunomodulating | Any antineoplastic and immunomodulating [redacted] | Error in percentage for | Correction made |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| agents row reads as follows<br>Any antineoplastic and immunomodulating agent [redacted] | | | | | agent | | | | | the sorafenib Western population cell | |

## Issue 7 CiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Table 8 (page 50)<br>Underlying cirrhosis results for the lenvatinib and sorafenib full populations are not marked as CiC | Underlying cirrhosis†, n (%) [redacted] | Data is marked as CiC in the company submission | Mark-up changes to be made my NICE. |

## Issue 8 Readability

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (Section 4.3.1, page 57) states:<br>At the primary data cut-off (13 November 2016), 351 patients in the lenvatinib group and 350 (73.4%) in the sorafenib group (73.5%) had died, | At the primary data cut-off (13 November 2016), 351 patients in the lenvatinib group (73.4%) and 350 in the sorafenib group (73.5%) had died, | The percentage of patients who died in the lenvatinib group appears to have been written in the wrong place | Correction made |

### Issue 9 Transcription error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current value (Table 13, page 61) for hazard ratio for IIR-assessed with mRECIST is 0.64 (0.55 to 0.77) | 0.64 (0.55 to 0.75) | Error in the upper CI limit reported | Correction made |

### Issue 10 CiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (Section 4.3.4, page 64):<br>Results using the per protocol set were similar to the primary analysis (OR 3.19, 95% CI: 2.18 to 4.66). | Results using the per protocol set were similar to the primary analysis [redacted] | Data is marked as CiC in the company submission | Mark-up changes to be made my NICE. |

### Issue 11 Percentage error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (Section 4.3.5, page 66) states:<br>Most patients in both groups had discontinued treatment by the data cut-off (92.4% of the lenvatinib group and 94.7% of the sorafenib group) | Most patients in both groups had discontinued treatment by the data cut-off (94.4% of the lenvatinib group and 94.7% of the sorafenib group) | At the data cut off, 94.4% of lenvatinib patients had discontinued treatment as presented in Company submission appendix D, Figure 2 (CONSORT diagram) | Correction made |

### Issue 12 CiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (section 4.3.7, page 71) states:<br>Drug withdrawals (19.7%), and dose modifications (61.8%, including dose reductions and interruptions) due to TEAEs in the lenvatinib [redacted] | [redacted], including dose reductions and interruptions) due to TEAEs in the lenvatinib [redacted] | Additional data is marked as CiC in the company submission | Mark-up changes to be made my NICE. |

### Issue 13 Transcription error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Table 19 p72 has errors in the following cells:<br>Fatal SAEs (lenvatinib episodes/PY cell): [redacted]<br>Non-fatal SAEs (lenvatinib episodes/PY cell): [redacted]<br>Dose reductions (sorafenib patients with ≥1 cell): [redacted] | The correct values should be as follows:<br>Fatal SAEs (lenvatinib episodes/PY cell): [redacted]<br>Non-fatal SAEs (lenvatinib episodes/PY cell): [redacted]<br>Dose reductions (sorafenib patients with ≥1 cell): [redacted] | Correction of values in line with those presented in the company submission | Correction made. |

## Issue 14 Error in reporting of model results

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| In four places in the document, current text states that lenvatinib is dominated by sorafenib:<br>• Section 5.2, Table 22, page 81<br>• Section 5.2, Table 23, page 82<br>• Section 5.5.1, Table 43, page 119<br>• Section 5.5.3, Table 45, page 122 | In each instance, the ICER value for lenvatinib should state 'dominant'. | In each instance, lenvatinib is associated with more QALYs and lower costs compared with sorafenib. | Correction made. |

## Issue 15 AiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Figure 12 (page 92) and Figure 13 (page 93) not marked AiC | Graphs to be marked as AiC | Data is marked as AiC in the Company submission | Mark-up changes to be made my NICE. |

## Issue 16 AiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (section 5.4.8.2, page 102):<br>Nonetheless, the ERG would like to highlight that the size of the full | Nonetheless, the ERG would like to highlight that the size of the full population reduced from [redacted] and [redacted] for lenvatinib and sorafenib, respectively, at baseline, to [redacted] in both | Data is marked as AiC in the company submission | Mark-up changes to be made my NICE. |

| | | | |
|---|---|---|---|
| population reduced from 463 and 458 for lenvatinib and sorafenib, respectively, at baseline, to less than 25 in both treatment arms from cycle 26 to 43 | treatment arms from cycle 26 to 43 | | |

## Issue 17 Transcription error

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (section 5.4.9.6, page 118):<br><br>Within the company's Western subgroup analysis, the ERG would also like to note that the mean dose intensity was taken from the company's base case analysis informed by the full population (lenvatinib, 9.4 mg/day; sorafenib 669.1mg/day), rather than the mean dose intensity in the Western subgroup provided at clarification (lenvatinib [redacted]mg/day; sorafenib, [redacted]mg/day). However, given that the ERG prefers to include the costs of drug wastage in its base case analysis, this issue becomes somewhat redundant. | Within the company's Western subgroup analysis, the ERG would also like to note that the mean dose intensity was taken from the company's base case analysis informed by the full population (lenvatinib, 9.4 mg/day; sorafenib 663.8 mg/day), rather than the mean dose intensity in the Western subgroup provided at clarification (lenvatinib [redacted]mg/day; sorafenib, [redacted]mg/day). However, given that the ERG prefers to include the costs of drug wastage in its base case analysis, this issue becomes somewhat redundant. | The dose intensity for the sorafenib Western population was included as the value for the sorafenib full population, and vice versa. | Correction made. |

## Issue 18 Error in reporting of model results

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| In several places in the document, current text states that the company base case and ERG base case show lenvatinib to offer a life extension of 2.7 months and 3.6 months, respectively compared with sorafenib:<br>• Section, 5.5.1, page 119 (main body)<br>• Section 7, page 131 (main body)<br>• Section 7, page 131 (Table 50) | In all instances, text should be changed to state that lenvatinib offers a life extension of 3.1 months compared with sorafenib in the company base case and 4.1 months compared with sorafenib in the ERG base case. | The difference in *discounted* life years between lenvatinib and sorafenib is 2.7 months in the company base case and 3.6 months in the ERG base case. This value is used in Section 7 of the ERG report to quantify the life extension offered by lenvatinib over sorafenib.<br>The difference in *undiscounted* life years between lenvatinib and sorafenib is 3.1 months in the company base case and 4.1 months in the ERG base case, and these are the relevant values describing the life extension offered by lenvatinib. | Correction made. |

## Issue 19 CiC marking

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (section 5.5.3, page 122):<br>According to the company's analysis, lenvatinib was associated with mean cost-savings of £4,379 (95% CI: -£12,543 to £3,893) | According to the company's analysis, lenvatinib was associated with mean cost-savings of [redacted] | All cost data concerning lenvatinib to be marked CiC | Mark-up changes to be made my NICE. |

## Issue 20 Error in reporting of model methods

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (Section 1.6, page 21; emphasis added):<br><br>The cost difference was *less favourable* to lenvatinib after this change because of the removal of the subsequent treatment costs that were *greater in the sorafenib group*. | The cost difference was *more favourable* to lenvatinib after this change because of the removal of the subsequent treatment costs that were *greater in the lenvatinib group*. | Assuming list price for sorafenib, the costs of post-progression therapies are higher in the lenvatinib arm, irrespective of whether costed therapies include only sorafenib and regorafenib or a broader range of therapies. | Partial correction made. The text was amended to…<br><br>*"The overall cost difference was less favourable to lenvatinib after this change even though the greater subsequent treatment costs in the lenvatinib group were removed."* |

## Issue 21 Error in reporting of model methods

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| Current text (Section 1.6, page 21):<br><br>The ERG considered the company's approach, to take the difference in PFS between cycles, to be an overestimate of newly progressed patients as it did not factor in those who leave the PFS state because of death rather than progression.<br><br>Current text (Section 6.2, page | We propose that this critique be removed from both Section 1.6 and Section 6.2 in the report. | Due to ambiguous labelling in the model, the process by which the costs of post-progression therapies are applied has been misinterpreted.<br><br>An example of this process is as follows:<br><br>• 349 PFS events occurred in the lenvatinib arm of REFLECT (i.e. 349 individuals experienced either progression or death from the | Not a factual error. |

| | | | |
|---|---|---|---|
| 125):<br><br>The company's post-progression drug costs were applied to 'newly progressed' patients; however, the proportion of patients assumed to be newly progressed was taken as the difference in PFS between cycles. This is an overestimate of the proportion of newly progressed patients as some of those patients who leave the PFS state are people who have died. | | progression-free state).<br>• 106 patients in the lenvatinib arm used sorafenib post-progression.<br>• The proportion of lenvatinib patients leaving the progression-free state (due to either progression or death) who incur the cost of post-progression sorafenib use is therefore calculated as 106/349 = 30.37%<br><br>This process therefore already accounts for the fact that some patients leave the progression-free state due to death. | |

## Issue 22 Error in model implementation

| Description of problem | Description of proposed amendment | Justification for amendment | ERG response |
|---|---|---|---|
| The option to use Western subgroup utility data has not been implemented correctly. Only cell D15 on the 'Utility' sheet has been linked to the Western utility data in cells D48:D51 on the 'Model Parameters' sheet. | Cells C13, C15, D13 and D14 on the 'Utility' sheet should also be linked to the Western utility data in cells D48:D51 on the 'Model Parameters' sheet. | Results currently generated using this option will not be accurate. | Cells C13, C15, D13 and D14 on the 'Utility' sheet are reset when running a PSA and, therefore, the ERG applied these links in the 'Control' sheet defaults in order for the changes to be applied within the PSA. |

# Lenvatinib for advanced, unresectable, untreated hepatocellular carcinoma

ERRATUM

This report was commissioned by the NIHR HTA Programme as project number 16/56/18

BMJ Technology Assessment Group

This document contains errata in respect of the ERG report in response to the manufacturer's factual inaccuracy check.

The table below lists the page to be replaced in the original document and the nature of the change:

| Page No. | Change |
|---|---|
| 11 | Existing marketing authorisation wording clarified (Issue 2) |
| 21 | Correction to state that lenvatinib group had greater costs of subsequent treatment, and text amended to clarify the overall cost difference discussion. |
| 45 | Data corrections in Table 7 (Issue 4, 5 and 6) |
| 55 | Amended syntax of sentence about number of patients alive (Issue 8) |
| 59 | Upper CI of hazard ratio for IIR PFS (mRECIST) corrected in Table 13 (Issue 9) |
| 64 | Percentage of patients who discontinued treatment corrected (Issue 11) |
| 70 | Corrections to treatment-adjusted rate of fatal and non-fatal SAEs, and dose reductions |
| 79-80 | Text in results tables changed from "Dominated" to "Dominant" |
| 116 | Dose intensities for sorafenib corrected. Section header moved to page 117. |
| 117 | Text in results tables changed from "Dominated" to "Dominant". Section header added from page 116.<br>Text edited to clarify that improved survival discussion is based on discounted values. |
| 120 | Text in results tables changed from "Dominated" to "Dominant". |
| 129 | OS benefits changed to undiscounted values, and discussion of end of life criteria altered to reflect these changes. |
| Abbreviations: CI, confidence interval; IIR, independent imaging review; mRECIST, modified Response Evaluation Criteria in Solid Tumours; OS, overall survival; PFS, progression-free survival. | |

# 1 SUMMARY

## *1.1 Critique of the decision problem in the company's submission*

The company of lenvatinib ([redacted] Eisai Co., Ltd) submitted to the National Institute for Health and Care Excellence (NICE) clinical and economic evidence in support of the effectiveness of lenvatinib in the treatment of adults with unresectable hepatocellular carcinoma (HCC) who have not previously received systemic treatment.

The company provides a reasonable overview of the disease area and current service provision. Briefly, HCC is the most common primary liver cancer in England and is associated with a poor prognosis (5-year survival less than 15%). HCC commonly occurs in the presence of liver cirrhosis, and major risk factors include chronic alcohol consumption, hepatitis B or hepatitis C infection, and non-alcoholic fatty liver disease. Severity and prognosis of HCC is commonly categorised with the Barcelona Clinic Liver Cancer (BCLC) system, which helps to guide choice of treatment. Sorafenib is the only systemic therapy approved by NICE as an option for treating HCC, and is only recommended for patients with advanced disease (BCLC stage C), and well-preserved liver function (Child–Pugh grade A).

Lenvatinib, which the company positions as an alternative to sorafenib, inhibits the activity of multiple vascular endothelial growth factor (VEGF) and other tyrosine kinase receptors implicated in tumour growth and spread. Lenvatinib's marketing authorisation covers its use for differentiated thyroid carcinoma and in combination with everolimus for advanced renal cell carcinoma but, at the time of writing, the extension for use in HCC had not been approved; the draft indication is for the treatment of adult patients who have received no prior systemic therapy for HCC.

The clinical evidence presented in the company's submission (CS) is derived from REFLECT, an international phase III, open-label randomised controlled trial (RCT). REFLECT enrolled patients with previously untreated, advanced or unresectable HCC, which is in line with the final scope issued by NICE. The comparator (sorafenib) and outcomes in REFLECT were also aligned with the final scope, and are relevant to UK clinical practice. While also listed in the scope, the company does not provide evidence comparing lenvatinib with best supportive care (BSC), which the ERG's clinical experts considered appropriate; BSC is not a relevant comparator because it is only given when a patient refuses, or is considered unfit for, systemic treatment. The Evidence Review Group's (ERG's) clinical experts consider REFLECT to provide evidence that is relevant to the decision problem, but the trial's inclusion criteria mean results may not be generalisable to patients with compromised liver function (Child–Pugh stage B or worse) or poor performance status.

post-progression health state, the potentially overestimated PFS could have an important impact on the difference in costs as well as QALYs. The ERG attempted to assess the potential impact of this by reducing the PFS benefit of lenvatinib by arbitrarily reducing the scale parameter of the generalised gamma curve in the ERG's base case analysis. The results of this are given in Section **Error! Reference source not found.**.

## 1.6 Summary of exploratory and sensitivity analyses undertaken by the ERG

***Economic***

The ERG made a minor correction to the company's base case analysis to ensure that all costs were incorporated within the half-cycle correction in the model. This had only a small impact on the company's ICER.

The ERG then performed a range of scenario analyses around the company's corrected base case including two key changes that formed the ERG's preferred base case. These were to adjust the OS models for subsequent anti-cancer interventions instead of just applying the costs of them, and the second was to use the generalised gamma function to model PFS (including a correction to prevent the curves for each treatment group crossing).

The impact of the first change increased the OS benefit in favour of lenvatinib because there was a greater use of subsequent treatments in the sorafenib group causing OS to be overestimated more than for the lenvatinib group. The overall cost difference was less favourable to lenvatinib after this change even though the greater subsequent treatment costs in the lenvatinib group were removed. This was a result of reduced post-progression resource use in the sorafenib group, which was overestimated through the extended survival. The results of the ERG's base case are presented in Table A with changes incorporated cumulatively.

The ERG also conducted further scenario analyses around the company's corrected base case; a key one being the ERG's application of a mortality adjustment to the company's estimation of newly progressed. The ERG considered the company's approach, to take the difference in PFS between cycles, to be an overestimate of newly progressed patients as it did not factor in those who leave the PFS state because of death rather than progression. This, therefore, potentially overestimated subsequent treatment costs. This change did not apply to the ERG's base case as subsequent treatment costs were removed.

extent and type of subsequent anti-cancer procedures and medication received in REFLECT do not reflect UK clinical practice because there are currently no drugs approved for use after first-line systemic therapies for advanced HCC in the National Health Service (NHS).

Table 7. Summary of subsequent procedures and medications (≥1% of patiets in either group; adapted from CS, Table 12 and company's response to clarification, Tables 2, 3, 4, 5)

| **Subsequent interventions received during survival follow-up, n (%)** | **Lenvatinib** | | **Sorafenib** | |
|---|---|---|---|---|
| | **Full population** | **Western** | **Full population** | **Western** |
| **Any anti-cancer procedure or medication** | 206 (43.1) | 44 (28.0) | 243 (51.1) | 71 (45.2) |
| **Any anti-cancer procedure** | 122 (25.5) | 11 (7.0) | 130 (27.3) | 18 (11.5) |
| High frequency ablation | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to bone | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to brain | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to liver | [redacted] | [redacted] | [redacted] | [redacted] |
| Radiotherapy to lymph nodes | [redacted] | [redacted] | [redacted] | [redacted] |
| Regional chemotherapy | [redacted] | [redacted] | [redacted] | [redacted] |
| Transcatheter arterial chemoembolisation | [redacted] | [redacted] | [redacted] | [redacted] |
| **Any anti-cancer medication** | 156 (32.6) | 41 (26.1) | 184 (38.7) | 61 (38.9) |
| Any antineoplastic and immunomodulating agent | [redacted] | [redacted] | [redacted] | [redacted] |
| Antimetabolites* | [redacted] | [redacted] | [redacted] | [redacted] |
| Cytotoxic antibiotics and related substances<br>Doxorubicin | [redacted] | [redacted] | [redacted] | [redacted] |
| Immunostimulants (interferon) | [redacted] | [redacted] | [redacted] | [redacted] |
| Other neoplastic agents<br>Cabozantinib<br>Cisplatin<br>Ipilimumab<br>Nivolumab<br>Oxaliplatin<br>Protein kinase inhibitors<br>Regorafenib<br>Sorafenib<br>Sunitinib<br>Tivantinib | [redacted] | [redacted] | [redacted] | [redacted] |
| Various<br>Folinic acid<br>Investigational drugs | [redacted] | [redacted] | [redacted] | [redacted] |

Abbreviations: n, number of patients.
*Antimetabolites received by more than 1% of patients in either group were capecitabine, flurouracil, gemcitabine and uftoral.

Outcomes measured in REFLECT were considered clinically relevant by the ERG's clinical experts, and in line with the decision problem outlined in the NICE final scope1 (Section 3.4). Full details of the definitions used, reasons for censoring, and methods of analysis are provided in Table 9, but briefly:

was uncertain about trial conduct or considered there a risk of bias, for example with rules for censoring for PFS and TTP, the company provided adequate explanation and supplementary analyses at the clarification stage for the ERG to assess the potential impact on relative treatment effects.

The ERG requested additional information at the clarification stage to assess whether the Western subgroup or full population forms the most appropriate evidence base for this STA. Some differences between the full population and Western subgroup were noted, particularly in terms of baseline characteristics, but evidence from separate Western and Asia–Pacific trials for sorafenib in HCC (SHARP[31] and the Asia–Pacific study[32]) suggest that this does not necessarily impact relative treatment effects (Table 4). Neither the full population nor the Western subgroup were considered reflective of UK clinical practice with regards to the extend and type of subsequent treatments received, and there was more imbalance in the Western subgroup. Overall, the ERG does not consider there to be sufficient evidence to justify the inevitable loss of precision by focusing on the Western subgroup.

The ERG considers the company's statistical approach appropriate and well described and, where post-hoc analyses were conducted to test the robustness of treatment effects, these were justified fully with reference to the SAP. In cases where the ERG was uncertain about methods or considered there a risk of bias, for example with rules for censoring for PFS and TTP, the company provided adequate explanation and supplementary analyses at the clarification stage for the ERG to assess the potential impact on relative treatment effects.

## *1.7 Clinical effectiveness results*

### 1.7.1 Overall survival

At the primary data cut-off (13 November 2016), 351 patients in the lenvatinib group (73.4%) and 350 in the sorafenib group (73.5%) had died, at which point median survival follow-up was 27.7 months in the lenvatinib group and 27.2 months in the sorafenib group (Table 11). In the primary analysis, median OS was 13.6 months for patients treated with lenvatinib and 12.3 months for patients treated with sorafenib; the HR for lenvatinib versus sorafenib is 0.92 (95% CI: 0.79 to 1.06; Figure 3). The upper 95% CI falls below the predefined limit to declare non-inferiority (Section 4.2.3), but criteria for superiority were not met. While the company found no statistical deviation from the PH assumption for OS, the cumulative hazard plots were neither entirely straight nor parallel, so the ERG interprets the HRs with caution.

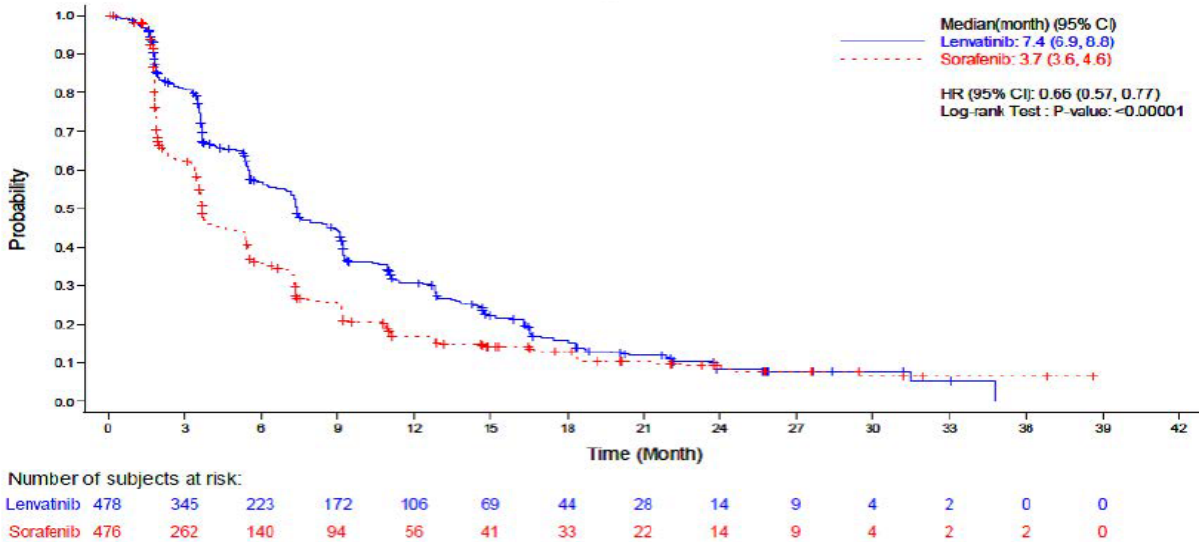
Table 13. Investigator- and IIR-assessed progression-free survival in REFLECT

| | Lenvatinib (N = 478) | Sorafenib (N = 476) |
|---|---|---|
| **Investigator-assessed with mRECIST (primary analysis)** | | |
| Patients with events, n (%)<br>Progressive disease<br>Death | [redacted] | [redacted] |
| Median (95% CI), months* | 7.4 (6.9 to 8.8) | 3.7 (3.6 to 4.6) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.66 (0.57 to 0.77) | |
| PFS rate, % (95% CI)† 6 months<br>12 months<br>18 months<br>24 months | [redacted] | [redacted] |
| Median follow-up (95% CI), months | [redacted] | [redacted] |
| ***Post-hoc* analysis - investigator-assessed (mRECIST) with amended censoring rules** | | |
| Patients with events, n (%)<br>Progressive disease<br>Death | [redacted] | [redacted] |
| Median (95% CI), months* | [redacted] | [redacted] |
| Stratified Cox model hazard ratio (95% CI)‡,§ | [redacted] | |
| **IIR-assessed with mRECIST[41]** | | |
| Median PFS (95% CI), months* | 7.3 (5.6 to 7.5) | 3.6 (3.6 to 3.7) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.64 (0.55 to 0.75) | |
| **IIR-assessed with RECIST v1.1[41]** | | |
| Median (95% CI), months* | 7.3 (5.6 to 7.5) | 3.6 (3.6 to 3.9) |
| Stratified Cox model hazard ratio (95% CI)‡,§ | 0.65 (0.56 to 0.77) | |

Abbreviations: CI, confidence interval; IIR, independent imaging review; n = number of patients; mRECIST, modified Response Evaluation Criteria for Solid Tumours; PFS, progression-free survival; RECIST v1.1, Response Evaluation Criteria for Solid Tumours, version 1.1

* 95% confidence intervals are estimated with a generalised Brookmeyer and Crowley method; †PFS rate and 95% CI were calculated using the Kaplan–Meier product-limit method and the Greenwood Formula; ‡Hazard ratio is for lenvatinib vs sorafenib, based on a Cox model including treatment group as a factor; §Stratified by region (Region 1: Asia–Pacific; Region 2: Western regions), macroscopic portal vein invasion or extrahepatic spread or both (yes, no), ECOG PS (0, 1) and body weight (<60 kg, ≥60 kg)

Figure 4. Kaplan–Meier curve for investigator-assessed progression-free survival (reproduced from CS, Figure 4)



### 1.7.2 Health-related quality of life

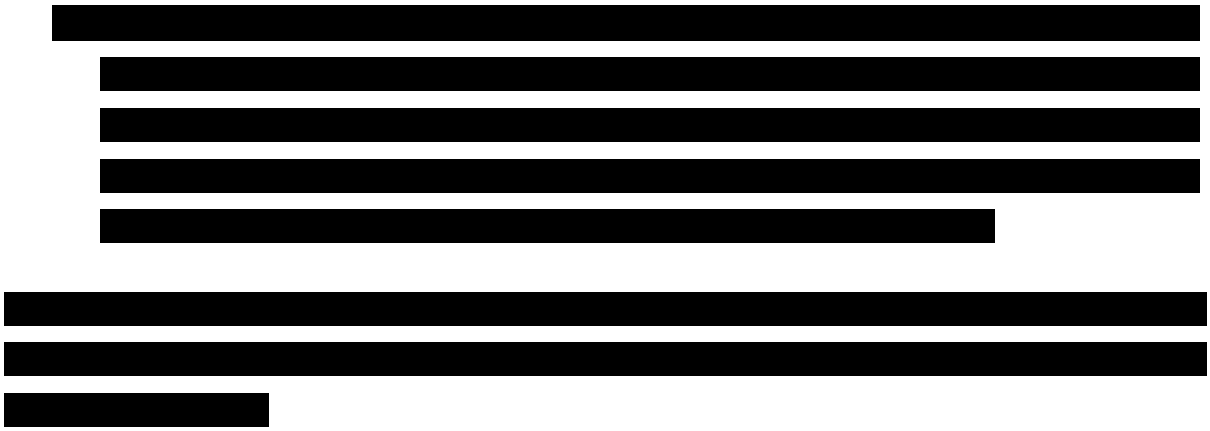
HRQoL data from REFLECT had not been published at the time this report was written, meaning the CS mainly included narrative summaries of selected results. HRQoL was assessed with three separate measures (Table 16) administered at baseline, day one of each treatment cycle and at the off-treatment visit. Results from the EORTC scales are discussed here. EQ-5D-3L data were incorporated in the economic model and are discussed in 5.4.8.

Table 16. Summary of HRQoL measures used in REFLECT

| Scale | Scope | Scoring |
|---|---|---|
| EORTC QLQ-HCC18 | HCC specific | 18 items across 8 scales.<br>1 ("not at all") to 4 ("very much"). |
| EORTC QLQ-C30 | Cancer generic | 30 items across 5 functional and 9 symptom domains<br>1 ("not at all") to 4 ("very much") or 1 ("very poor") to 7 ("excellent").<br>Summary score (1 to 100) and 1 to 7 global health status score (higher scores on both indicate better HRQoL). |
| EQ-5D-3L | Generic | 5 domains rated 1 ("no problems") to 3 ("extreme problems"): mobility, self-Care, usual activities, pain/discomfort, and anxiety/depression. Scores give a single health utility index.<br>EQ-VAS from 0 to 100 (worst to best imaginable health state). |

Abbreviations: EORTC-QLQ-HCC18, European Organisation for Research and Treatment of Cancer quality of life scale – hepatocellular carcinoma 18 items; EORTC QLQ-C30, EORTC QLQ cancer 30 items; EQ-5D-3L, EuroQol 5-dimensions, 3-levels; EQ-VAS, EuroQol visual analogue scale. HRQoL, health-related quality of life.

Completion of questionnaires remained above 90% throughout the Randomisation Phase, but the company considered attrition after cycle 18 too high to yield meaningful cross-sectional results (CS, Table 18). Most patients in both groups had discontinued treatment by the data cut-off (94.4% of the lenvatinib group and 94.7% of the sorafenib group) but a table of baseline, cross-sectional, and end of treatment scores for each domain and scale was not presented for comparison between groups. The ERG notes from the CSR that full HRQoL findings from REFLECT were presented in a separate report, which was not provided to the ERG. The following results are based on narrative and graphical summaries in the CS and CSR:



increased, proteinuria and hepatic failure. Without information about when withdrawals and modifications occurred in each group, it is unknown how much of the difference can be explained by imbalance in treatment exposure.

Table 19. Summary of TEAEs by patient incidence and treatment exposure (adapted from CS, Tables 20, 21, 22 and 23)

| Adverse event | Lenvatinib (N = 476) | | Sorafenib (N = 475) | |
|---|---|---|---|---|
| | N patients with ≥1 (%) | Episodes/PY | N patients with ≥1 (%) | Episodes/PY |
| TEAE | 470 (98.7) | 6,124 (18.89) | 472 (99.4) | 4,718 (19.73 |
| Treatment-related TEAE | 447 (93.9) | 3,546 (10.94) | 452 (95.2) | 2,865 (11.98) |
| SAE total | 205 (43.1) | 409 (1.26) | 144 (30.3) | 232 (0.97) |
| Fatal | 61 (12.8) | 61 (0.19) | 36 (7.6) | 36 (0.15) |
| Non-fatal | 189 (39.7) | 185 (38.9) | 128 (26.9) | 207 (0.87) |
| Drug withdrawal due to TEAE | 94 (19.7) | - | 69 (14.5) | - |
| Dose modification due to TEAE (all) | 294 (61.8) | - | 264 (55.6) | - |
| Reductions | 184 (38.7) | - | 185 (38.9) | - |
| Interruptions | 248 (52.1) | - | 193 (40.6) | - |
| **Specific TEAEs occurring in ≥10% of patients by preferred term (safety analysis set)** | | | | |
| Hypertension | 201 (42.2) | 253 (0.78) | 144 (30.3) | 166 (0.69) |
| Diarrhoea | 184 (38.7) | 327 (1.01) | 220 (46.3) | 351 (1.47) |
| Decreased appetite | 162 (34.0) | 196 (0.60) | 127 (26.7) | 139 (0.58) |
| Weight decreased | 147 (30.9) | 165 (0.51) | 106 (22.3) | 113 (0.47) |
| Fatigue | 141 (29.6) | 156 (0.48) | 119 (25.1) | 130 (0.54) |
| Palmar-plantar erythrodysaesthesia syndrome | 128 (26.9) | 142 (0.44) | 249 (52.4) | 289 (1.21) |
| Proteinuria | 117 (24.6) | 163 (0.50) | 54 (11.4) | 74 (0.31) |
| Dysphonia | 113 (23.7) | 131 (0.40) | 57 (12.0) | 66 (0.28) |
| Nausea | 93 (19.5) | 113 (0.35) | 68 (14.3) | 77 (0.32) |
| Platelet count decreased | 87 (18.3) | 117 (0.36) | 58 (12.2) | 67 (0.28) |
| Abdominal pain | 81 (17.0) | 108 (0.33) | 87 (18.3) | 106 (0.44) |
| Hypothyroidism | 78 (16.4) | 79 (0.24) | 8 (1.7) | 8 (0.03) |
| Vomiting | 77 (16.2) | 106 (0.33) | 36 (7.6) | 57 (0.24) |
| Constipation | 76 (16.0) | 89 (0.27) | 52 (10.9) | 60 (0.25) |
| Blood bilirubin increased | 71 (14.9) | 97 (0.30) | 63 (13.3) | 75 (0.31) |
| Pyrexia | 69 (14.5) | 80 (0.25) | 63 (13.3) | 72 (0.30) |
| Ascites | 68 (14.3) | 77 (0.24) | 44 (9.3) | 51 (0.21) |
| Oedema peripheral | 66 (13.9) | 85 (0.26) | 33 (6.9) | 40 (0.17) |
| Aspartate aminotransferase increased | 65 (13.7) | 82 (0.25) | 80 (16.8) | 100 (0.42) |
| Abdominal pain upper | 58 (12.2) | 73 (0.23) | 40 (8.4) | 48 (0.20) |
| Asthenia | 54 (11.3) | 76 (0.23) | 48 (10.1) | 54 (0.23) |
| Alanine aminotransferase increased | 53 (11.1) | 67 (0.21) | 52 (10.9) | 66 (0.28) |
| Back pain | 50 (10.5) | 55 (0.17) | 31 (6.5) | 32 (0.13) |
| Rash | 46 (9.7) | 51 (0.16) | 76 (16.0) | 87 (0.36) |
| Stomatitis | 45 (9.5) | 55 (0.17) | 56 (11.8) | 67 (0.28) |
| Alopecia | 14 (2.9) | 15 (0.05) | 119 (25.1) | 123 (0.51) |

Abbreviations: CS, company submission; PY, patient year; SAE, serious adverse event; TEAE, treatment-emergent adverse event.

# 5 COST EFFECTIVENESS

## 5.1 Introduction

This section provides a structured description and critique of the systematic literature review and *de novo* economic evaluation submitted by the company. The company provided a written submission of the economic evidence along with an electronic version of the Microsoft® Excel based economic model. Table 1 summarises the location of the key economic information within the company's submission (CS).

Table 1. Summary of key information within the company's submission

| Information | Section (CS) |
|---|---|
| Details of the systematic review of the economic literature | B.3.1 |
| Model structure | B.3.2.2 |
| Technology | B.3.2.3 |
| Clinical parameters and variables | B.3.3 |
| Measurement and valuation of health effects and adverse events | B.3.4 |
| Resource identification, valuation and measurement | B.3.5 |
| Summary of base-case analysis inputs and assumptions | B.3.6 |
| Results | B.3.7 |
| Sensitivity analysis | B.3.8 |
| Subgroup analysis | B.3.9 |
| Validation | B.3.10 |
| Interpretation and conclusions | B.3.11 |
| Abbreviations used in table: CS, company submission. | |

## 5.2 Summary of the company's key results

The results of the company's deterministic base case analysis are presented in Table 2, and the results of the company's probabilistic sensitivity analysis are given in Table 3.

Table 2. Base-case results (taken from the company's response to clarification)

| Technologies | Total costs | Total LYG | Total QALYs | Incremental costs | Incremental LYG | Incremental QALYs | ICER |
|---|---|---|---|---|---|---|---|
| Sorafenib | £65,592 | 1.46 | 1.03 | - | - | - | - |
| Lenvatinib | [redacted] | 1.69 | [redacted] | [redacted] | 0.23 | [redacted] | Dominant |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. | | | | | | | |

Table 3. PSA results using 10,000 iterations (reproduced from the economic model provided at clarification)

| Technologies | Total costs | Total QALYs | Incremental costs | Incremental QALYs | ICER |
|---|---|---|---|---|---|
| Sorafenib | £65,688 | 1.03 | - | - | - |
| Lenvatinib | [redacted] | [redacted] | [redacted] | [redacted] | Dominant |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years | | | | | |

## 5.3 ERG comment on company's review of cost-effectiveness evidence

The company carried out a systematic literature review (SLR) to identify cost-effectiveness and resource use evidence in patients with treatment-naïve advanced, unresectable or metastatic hepatocellular carcinoma (HCC). Another SLR was carried out to identify studies reporting on the health-related quality of life (HRQoL) of patients with advanced, unresectable or metastatic HCC. All database searches were carried out in November 2017.

When conducting the SLRs, the company searched the MEDLINE and Embase databases, as well as the Cochrane Library, using the OVID platform. In addition, conference proceedings and grey literature, including relevant heath technology assessment (HTA) bodies, were also hand searched for evidence. Search strategies are provided in Appendix G for cost-effectiveness and resource use evidence and Appendix H for HRQOL evidence. In summary, search terms combined the population (patients with advanced, unresectable or metastatic HCC) with economic and quality of life terms, which the ERG considers to be inclusive.

Overall, a total of 46 cost-effectiveness and resource use studies met the pre-specified inclusion and exclusion criteria reported in Table 7 of Appendix G. Of the 46 included studies, 27 were cost-effectiveness analyses, while 19 were budget impact or cost analyses. A summary of the eight cost-effectiveness analyses conducted from a UK perspective is provided in Table 25 of the CS. However, none of the 19 included resource use studies were conducted in the UK and, therefore, data from those studies were not extracted by the company. However, it is important to clarify that the company considered the included cost-effectiveness studies (including TA189[34] and TA474[24]) to inform resource use inputs in their submission, although they were not reported again under resource use studies. Sources of resource and cost use data are described in greater detail in Section **Error! Reference source not found.**.

As for the HRQoL search, a total of 14 HRQoL studies met the pre-specified inclusion and exclusion criteria reported in Table 10 of Appendix H. A summary of the 14 HRQoL studies is provided in Table 11 of Appendix H. However, as described in Section **Error! Reference source not found.**, it was not necessary to incorporate data from any of the identified studies because EQ-5D-3L data directly from the REFLECT trial were used to populate the economic model.

Table 4. Investigational drugs by treatment arm and region

| Subsequent interventions received during survival follow-up, n (%) | Lenvatinib | | Sorafenib | |
|---|---|---|---|---|
| | Full | Western | Full | Western |
| Investigational drugs | [redacted] | [redacted] | [redacted] | [redacted] |

As for the Western subgroup, subsequent anti-cancer interventions (medications or procedures) were not costed in a scenario analysis because the Western subgroup analysis controls for subsequent anti-cancer interventions and, therefore, the company concluded it would be inconsistent to apply different subsequent intervention costs by treatment arm. However, the ERG notes that the adjustment was based only on the categorisation of patients into those who received any subsequent anti-cancer procedure or medication and those who received neither. As such, the adjustment does not control for any imbalances in the types of intervention received in each treatment arm. The ERG was unable to explore an analysis, which included different subsequent intervention costs by treatment arm, as the company did not provide the percentage of progressed patients receiving subsequent interventions or the mean duration of subsequent interventions for the Western subgroup. As such, the ERG recommends that the company's analysis of the Western subgroup is interpreted with caution.

Within the company's Western subgroup analysis, the ERG would also like to note that the mean dose intensity was taken from the company's base case analysis informed by the full population (lenvatinib, 9.4 mg/day; sorafenib 663.8mg/day), rather than the mean dose intensity in the Western subgroup provided at clarification (lenvatinib [redacted] mg/day; sorafenib, [redacted] mg/day). However, given that the ERG prefers to include the costs of drug wastage in its base case analysis, this issue becomes somewhat redundant.

Finally, as for the company's scenario which excludes subsequent anti-cancer interventions, the impact on the full population was notable and favoured lenvatinib (as lenvatinib was associated with higher subsequent anti-cancer medication costs than sorafenib in the base case analysis) with the difference in costs increasing from [redacted] to [redacted]. A scenario was not presented by the company for the Western subgroup as the multivariable models for PFS and OS estimated in the full population included region as a covariate and adjusted for the use of post-progression therapies in the company's Western subgroup analysis.

Overall, while there was variation in the type and extent of subsequent interventions, neither the full population nor the Western subgroup are considered reflective of UK clinical practice. Moreover, the company's results from the full population and Western subgroup are robust to all additional analyses which change the approach to cost subsequent anti-cancer interventions.

## 5.5 Results included in company's submission

The company presented deterministic and probabilistic results. The base case results were calculated deterministically (using mean parameter values) as well as probabilistically (assessing the simultaneous effect of parameter uncertainty).

The company also carried out a series of univariate sensitivity analyses and scenario analyses to test the robustness of model results to changes in model parameters and structural assumptions. Base case results are presented in Section 5.5.1, while the results of deterministic and probabilistic sensitivity analyses are presented in Sections 5.5.2 and 5.5.3, respectively. No subgroup analyses were performed in the company's original model, although a Western subgroup was provided at clarification as requested by the ERG, as discussed in Section **Error! Reference source not found.**.

### 5.5.1 Base case results

The original multivariable analyses contained Child–Pugh score as a continuous variable. However, during the clarification stage, Child–Pugh class as a categorical variable (A vs B) was considered more appropriate by the company. Consequently, all analyses used to generate results in the revised economic model were based on re-estimated statistical models. In response to clarification questions, the company also corrected end-of-life care costs in their revised analyses. All other aspects of the analyses remain the same.

The results of the company's revised base case analysis using PAS prices for lenvatinib are presented in Table 5. As described previously in Section **Error! Reference source not found.**, the base-case analysis was based on multivariable adjustments to the PFS and OS curves for imbalances in baseline characteristics (including AFP and HCC aetiology) but not for imbalances in subsequent treatments.

According to the company's base case analysis, lenvatinib increases (discounted) survival by around 2.7 months compared to sorafenib. This translates to an incremental average QALY gain for lenvatinib of [redacted] QALYs. Overall, sorafenib is dominated by lenvatinib, in that lenvatinib is less costly and more effective.

Table 5. Base-case results (taken from the company's response to clarification)

| Technologies | Total costs | Total LYG | Total QALYs | Incremental costs | Incremental LYG | Incremental QALYs | ICER |
|---|---|---|---|---|---|---|---|
| Sorafenib | £65,592 | 1.46 | 1.03 | - | - | - | - |
| Lenvatinib | [redacted] | 1.69 | [redacted] | [redacted] | 0.23 | [redacted] | Dominant |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. | | | | | | | |

### 5.5.2 Deterministic sensitivity analysis

The company carried out univariate sensitivity analyses to assess the impact of varying the values of parameters from their means over a plausible range determined by either the 95% confidence interval. According to the univariate sensitivity analysis, the main drivers of the model are the constant terms (baseline hazard of events) for the base-case PFS and OS models (lognormal and loglogistic) for each of the lenvatinib and sorafenib arms. As for the scenario analysis, results were most sensitive to a 60% discount to the sorafenib list price and shorter time horizons.

### 5.5.3 Probabilistic sensitivity analysis

The company performed probabilistic sensitivity analysis (PSA) to assess the joint parameter uncertainty around the base case results. The results are based on 10,000 PSA iterations.

The ERG considers the parameters and respective distributions chosen for PSA, outlined in Appendix N of the CS, to be generally sound. In summary, utilities, AE rates and the proportion of patients receiving disease management were varied using a beta distribution, while unit costs and the frequency of resource use were varied using a gamma distribution. Drug acquisition cost were kept constant. Where the covariance structure between parameters was known (survival curves), correlated random draws were sampled from a multivariate normal distribution.

The results of the company's PSA in the revised model using PAS prices for lenvatinib are presented in Table 6. According to the company's analysis, lenvatinib was associated with mean cost-savings of [redacted] and mean incremental QALYs of [redacted] which the ERG considers to be comparable to the deterministic base case results. Furthermore, the ERG could produce very similar PSA results when they replicated the analysis using 10,000 iterations.

Table 6. PSA results using 10,000 iterations (reproduced from the economic model provided at clarification)

| Technologies | Total costs | Total QALYs | Incremental costs | Incremental QALYs | ICER |
|---|---|---|---|---|---|
| Sorafenib | £65,688 | 1.03 | - | - | - |
| Lenvatinib | [redacted] | [redacted] | [redacted] | [redacted] | Dominant |
| Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years | | | | | |

The scatterplots and cost-effectiveness acceptability curves (CEACs) are presented in **Error! Reference source not found.** and **Error! Reference source not found.** using PAS prices for lenvatinib. The majority of PSA iterations lie in the south-east quadrant of the cost-effectiveness plane. Moreover, the probability that lenvatinib was cost-effective at a threshold of £50,000 per QALY was 100%.

# 7 END OF LIFE

The company propose that lenvatinib meets both criteria outlined by the National Institute for Health and Care Excellence (NICE) for an end of life treatment. The company's assessment and ERG's comments are provided in Table 7.

The company refers to the 2012 joint European Society for Medical Oncology and European Society of Digestive Oncology (ESMO–ESDO) clinical practice guidelines as evidence for the first criterion, which break down prognosis by Barcelona Clinic Liver Cancer (BCLC) stages. The guidelines state that median OS for those with advanced HCC (BCLC Stage C), the relevant population for this Single Technology Appraisal, is 4 to 8 months based on natural history, and 6 to 11 months for those treated with sorafenib (current standard of care).[21] Median OS in the sorafenib group of REFLECT (12.3 months), and mean survival for sorafenib from the company's updated base case (17.5 months) and the ERG's base case (16.2 months), all fall below the threshold of 24 months.

The second criterion was met based on the company's base case extrapolation of OS, with an undiscounted incremental mean OS benefit of 3.1 months. The benefits in OS were increased in the ERG's preferred base case analysis, which produced an undiscounted incremental mean OS benefit of 4.1 months, therefore meeting the second criteria. This increase in the ERG's preferred base case was caused by the adjustment for subsequent treatments.

Overall, there is little uncertainty that lenvatinib meets the first end-of-life criterion concerning life expectancy, and the second criterion is also likely to be met, with both the company's and ERG's preferred assumptions resulting in a survival benefit greater than 3 months.

Table 7. End of life considerations

| NICE criterion | Company assessment | ERG comment |
|---|---|---|
| The treatment is indicated for patients with a short life expectancy, normally less than 24 months | CS section B.1.3, page 12. The current life expectancy for patients with advanced HCC is less than 1 year.[21] | ESMO-ESDO guidelines give median OS based on natural history as 4 to 8 months, and 6 to 11 months with sorafenib, for people with BCLC Stage C HCC.<br>From the company submission:<br>• Sorafenib median OS in REFLECT (primary analysis) = 12.3 months (upper quartile 25.4 months).<br>• Sorafenib mean OS = 17.5 months in company base case, 16.2 months in ERG's base case. |
| 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 | CS section B.3.11, page 116. Results of the cost-effectiveness analysis show an incremental mean overall survival benefit for lenvatinib of 3.1 months compared with sorafenib. | Incremental mean OS benefit from the company's economic model:<br>• Company base case = 3.1 months;<br>• ERG base case = 4.1 months. |

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; CS, company submission; ERG, evidence review group; ESMO-ESDO, joint European Society for Medical Oncology and European Society of Digestive Oncology guidance; HCC, hepatocellular carcinoma; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; OS, overall survival.