NICE National Institute for Health and Care Excellence

# Diagnostics Assessment Programme

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

# Committee Papers

**NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE**

**Diagnostics Assessment Programme**

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

**Contents:**



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

**Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence Diagnostics Assessment Report**

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

| | |
|---|---|
| **Produced by** | Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health, University of Sheffield |
| **Authors** | Paul Tappenden, Professor of Health Economic Modelling, SCHARR, University of Sheffield, Sheffield, UK<br>Katy Cooper, Senior Research Fellow, SCHARR, University of Sheffield, Sheffield, UK<br>Jean Hamilton, Research Fellow, SCHARR, University of Sheffield, Sheffield, UK<br>Gamze Nalbant, Research Associate, SCHARR, University of Sheffield, Sheffield, UK<br>Munira Essat, Senior Research Fellow, SCHARR, University of Sheffield, Sheffield, UK<br>Annabel Rayner, Research Assistant, SCHARR, University of Sheffield, Sheffield, UK<br>Ruth Wong, Information Specialist, SCHARR, University of Sheffield, Sheffield, UK<br>Nicolò Matteo Luca Battisti, Consultant Medical Oncologist, The Royal Marsden NHS Foundation Trust, London, UK.<br>Lynda Wyld, Professor of Surgical Oncology, School of Medicine and Population Health, University of Sheffield, Sheffield, UK.<br>Uzma Asghar, Consultant Medical Oncologist, The Royal Marsden NHS Foundation Trust, London, UK. |
| **Correspondence author** | Paul Tappenden, Professor of Health Economic Modelling, SCHARR, University of Sheffield, Sheffield, UK |
| **Date completed** | 25th September 2023 |

**Source of funding**: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135822.

**Declared competing interests of the authors**

None of the authors have any conflicts of interest to declare.

**Acknowledgements**

We would like to thank Dr Sunhong Kwon and Dr Aline Navega Biz, SCHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, SCHARR, for providing administrative support and in preparing and formatting 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 Evidence Synthesis Programme. Any errors are the responsibility of the authors.

**This report should be referenced as follows:**

Tappenden P, Cooper K, Hamilton J, Nalbant G, Essat M, Rayner A, Wong R, Battisti N, Wyld L, Asghar U. Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer. Sheffield Centre for Health and Related Research (SCHARR), University of Sheffield, 2023.

**Contributions of authors**

Paul Tappenden led the project and was responsible for day-to-day management. Ruth Wong designed and ran the search strategy. Katy Cooper, Gamze Nalbant and Munira Essat conducted the systematic review of the clinical effectiveness data. Jean Hamilton provided statistical advice to support the interpretation of the studies included in the review. Paul Tappenden and Annabel Rayner undertook the systematic review of existing economic models, critiqued the test manufacturers' models and designed and implemented the EAG model. Nicolò Matteo Luca Battisti, Lynda Wyld and Uzma Asghar provided clinical advice throughout the course of the project. All authors were involved in drafting and commenting on the final report.

**Standard copyright statement**

Copyright belongs to The University of Sheffield.

**About SCHARR**

The Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health at the University of Sheffield. SCHARR brings together a wide range of medical- and health-related disciplines including public health, general practice, mental health, epidemiology, health economics, management sciences, medical statistics, operational research and information science. It includes the Sheffield unit of the Trent Institute for Health Services Research, which is funded by NHS R&D to facilitate high-quality health services research and capacity development.

The SCHARR Technology Assessment Group (SCHARR-TAG) synthesises research on the clinical effectiveness and cost-effectiveness of healthcare interventions for the NIHR Health Technology Assessment Programme on behalf of a range of policy makers, including the National Institute for Health and Care Excellence (NICE). SCHARR-TAG is part of a wider collaboration of a number of units from other regions including Southampton Health Technology Assessment Centre (SHTAC), University of Southampton; Aberdeen Health Technology Assessment Group (Aberdeen HTA Group), University of Aberdeen; Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; Peninsular Technology Assessment Group (PenTAG), University of Exeter; the NHS Centre for Reviews and Dissemination, University of York; Warwick Evidence, The University of Warwick; the BMJ Group, Kleijnen Systematic Reviews, Bristol Technology Assessment Group (Bristol TAG), University of Bristol and the Newcastle NIHR TAR Team, Newcastle University.

## SUMMARY OF CHANGES MADE TO THE REPORT FOLLOWING THE CONSULTATION PROCESS

| Section / location | Change |
|---|---|
| Section 1.4 | Clarified the explanation for why prediction of chemotherapy benefit could not be determined from MINDACT (i.e., because all patients in the clinical high-risk, MammaPrint high-risk group were offered chemotherapy). |
| Section 1.5 | Sentence amended as follows (additional text shown in bold): "**where data permit**, risk classification probabilities and DRFI estimates for each test have been taken from same source" |
| Section 3.2.2 | Wording changed to more accurately reflect Mook 2009 study ("reanalyses of studies" and "reanalysis of two cohorts").<br>Added information on RxPONDER in terms of number of positive nodes and number of patients screened and randomised.<br>Clarified the explanation for why prediction of chemotherapy benefit could not be determined from MINDACT. |
| Section 3.3 | Risk of bias: added information on RxPONDER in terms of possible selection bias. |
| Section 3.4.2<br>Table 5 | For the table summarising prognostic data for 10-year distant recurrence: updated table to present each study on a separate row for clarity |
| Section 3.4.4 and Table 6 | Added text on distant recurrence data in RxPONDER compared with that in RCT reanalyses.<br>For the table of prognostic data from RxPONDER: added recurrence risk data (which was already included in tables in Section 3.5, prediction) for comparison with other studies in this section. Also added Ns with 1, 2 or 3 positive nodes. |
| Section 3.5.1 | Wording changed to more accurately reflect Mook 2009 study ("reanalysis of two cohorts") |
| Section 3.5.3<br>Table 10 | Table of chemotherapy benefit in RxPONDER: Added sub-headings to more clearly differentiate between pre- and post-menopausal populations. Also added Ns with 1, 2 or 3 positive nodes. |
| Section 3.5.4 | Wording changed to more accurately reflect Mook 2009 study ("reanalysis of two cohorts") |
| Section 3.5.5 and Table 12 | Clarified the explanation for why prediction of chemotherapy benefit could not be determined from MINDACT (i.e., because all patients in the clinical high-risk, MammaPrint high-risk group were offered chemotherapy).<br>For the table of chemotherapy benefit in MINDACT: added separate column headers for low and high MMP chemotherapy HRs, for clarity |
| Section 3.6.3<br>Table 16 | Decision impact table: Omitted last line of table which related to an excluded study and was included in error. |
| Section 4.2.2.2 | The sentence detailing the EAG's concerns regarding the indirect comparisons contained in the Agendia model has been amended to read "The EAG believes that the comparisons against the other tumour profiling tests included in the Agendia model are problematic as they assume that the characteristics of the patient populations enrolled in TransATAC[19] and MINDACT[29] are identical with respect to prognostic factors and treatment effect modifiers." |
| Section 4.2.2.3 | The EAG has updated the re-analysis of the Agendia model using the version of the model provided as part of the company's consultation response. This includes correction of the errors relating to the half-cycle correction and programming errors in the supportive care cost calculations. Table 28 and the related text have been updated. |
| Section 4.3.1 | Section 4.3.1. The text regarding mAOL has been amended to read "Within the MINDACT trial,[29, 149] clinical high-risk was defined using mAOL. During the appraisal consultation process, the company highlighted that within the HR+, HER2-, LN1-3 population, mAOL high-risk is equivalent to NPI>3.4." |

| | |
|---|---|
| Section 4.4 | The following sentence was amended as follows (changes in bold):<br>"RxPONDER indicates that chemotherapy is not beneficial to post-menopausal women who have an RS of 0-25**. The** test for interaction between the treatment group and the continuous RS in RxPONDER, when adjusted for the continuous RS, **menopausal status, and treatment group, was not statistically significant within the range RS 0-25 (*p*=0.35).** |
| Section 4.4 | The sentence regarding the use of the same data source for risk classification probabilities and DRFI estimates has been amended to read "Where data permit, risk classification probabilities and DRFI estimates for each test have been taken from same source. This approach maintains correlation between these parameters and avoids the potential for spectrum bias." A similar amendment has been made to Section 5.2.2. |
| Section 5.1.1 | For RxPONDER, added Ns with 1, 2 or 3 positive nodes.<br>For MINDACT, clarified the explanation for why prediction of chemotherapy benefit could not be determined from MINDACT (i.e., because all patients in the clinical high-risk, MammaPrint high-risk group were offered chemotherapy). |
| Multiple sections | The EAG's criticisms regarding the use of mAOL have been deleted from Section 1.5 (page 18), Section 4.2.2.1 (page 100), Table 23 (page 100), Box 1 (page 104), Section 4.2.2.2 (page 105), Table 29 (page 111), Section 4.4 (page 145), Section 5.2.2 (page 151). The EAG has clarified that in the HR+, HER2-, LN1-3 population, mAOL high-risk is equivalent to NPI>3.4. |

## TABLE OF CONTENTS

**LIST OF TABLES**

**LIST OF FIGURES**



**LIST OF BOXES**

## LIST OF ABBREVIATIONS

| | |
|---|---|
| A2LA | American Association for Laboratory Accreditation |
| AACR | American Association for Cancer Research |
| AE | Adverse event |
| AI | Aromatase inhibitor |
| AJCC | American Joint Committee on Cancer |
| AML | Acute myeloid leukaemia |
| AOL | Adjuvant! Online |
| ASCO | American Society of Clinical Oncology |
| BC | Base case |
| BCSS | Breast cancer-specific survival |
| BNF | British National Formulary |
| BRCA | Breast cancer gene |
| BSA | Body surface area |
| CAP | College of American Pathologists |
| C-D | Carboplatin plus docetaxel |
| CDK4/6 | Cyclin-dependent kinase 4 and 6 |
| CE | Conformité Européene |
| CEA | Cost-effectiveness analysis |
| CEAC | Cost-effectiveness acceptability curve |
| CET | Chemotherapy plus endocrine therapy |
| CHF | Congestive heart failure |
| CHM | Commission on Human Medicines |
| CI | Confidence interval |
| CINV | Chemotherapy-induced nausea and vomiting |
| CML | Chronic myeloid leukaemia |
| CMML | Chronic myelomonocytic leukaemia |
| CPCI | Conference Proceedings Citation Index |
| CS | Company's submission |
| CTS | Clinical Treatment Score |
| DCIS | ductal carcinoma *in situ* |
| DFS | Disease-free survival |
| DG | Diagnostics Guidance |
| DM | Distant metastases |
| DMFI | Distant metastasis-free interval |
| DMFS | Distant metastasis-free survival |
| DRFI | Distant recurrence-free interval |
| DRFR | Distant recurrence-free rate |
| DRFS | Distant recurrence-free survival |
| DSA | Deterministic sensitivity analysis |
| EAG | External Assessment Group |
| EBC | Early breast cancer |
| EBCTCG | Early Breast Cancer Trialists' Collaborative Group |
| EC90 | Epirubicin and cyclophosphamide |
| EC90/P | Epirubicin, cyclophosphamide followed by paclitaxel |
| EC90/T75 | Epirubicin and cyclophosphamide followed by docetaxel |
| ECCO | European Cancer Organization |
| ECG | Electrocardiogram |
| EFS | Event-free survival |
| EMBASE | Excerpta Medica Database |
| eMIT | Electronic Market Information Tool |
| EORTC | European Organisation for Research and Treatment of Cancer |
| EQ-5D | Euroqol 5 Dimensions |
| ER | Oestrogen receptor |

| | |
|---|---|
| ERBB | Erythroblastic oncogene B |
| ESBC | Early-stage breast cancer |
| ESMO | European Society for Medical Oncology |
| ET | Endocrine therapy |
| FACT-B | Functional Assessment of Cancer Therapy – Breast cancer |
| FACT-G | Functional Assessment of Cancer Therapy – General |
| FBC | Full blood count |
| FEC100-T | Fluorouracil, epirubicin, cyclophosphamide and docetaxel |
| FEC75 | Fluorouracil, epirubicin and cyclophosphamide |
| FFPE | Formalin-fixed paraffin-embedded |
| FISH | Fluorescence *in situ* hybridisation |
| FN | Febrile neutropenia |
| G-CSF | Granulocyte-colony stimulating factor |
| GO | Gemtuzumab ozogamicin |
| GP | General practitioner |
| HCHS | Hospital and Community Health Services |
| HER2 | Human epidermal growth factor receptor 2 |
| HF | Heart failure |
| HR | Hormone receptor |
| HR | Hazard ratio |
| HRQoL | Health-related quality of life |
| HSCT | Haematopoietic stem cell transplantation |
| HTA | Health Technology Assessment |
| ICER | Incremental cost-effectiveness ratio |
| IDFS | Invasive disease-free survival |
| IHC | Immunohistochemistry |
| INAHTA | International Network of Agencies for Health Technology Assessment |
| Inc. | Incremental |
| Int. | Intermediate |
| IPD | Individual patient data |
| IV | Intravenous |
| LCIS | Lobular carcinoma *in situ* |
| LFT | Liver function test |
| LN | Lymph node |
| LN+ | Lymph node positive |
| LN0 | Lymph node negative |
| LR | Local recurrence |
| LRR | Locoregional recurrence |
| LYG | Life year gained |
| MAIC | Matching-adjusted indirect comparison |
| mAOL | Modified Adjuvant! Online |
| MDS | Myelodysplastic syndromes |
| MEDLINE | Medical Literature Analysis and Retrieval System Online |
| MeSH | Medical Subject Heading |
| MHRA | Medicines and Healthcare products Regulatory Agency |
| MIMS | Monthly Index of Medical Specialities |
| mRNA | Messenger ribonucleic acid |
| N/a | Not applicable |
| NCCN | National Comprehensive Cancer Network |
| NCDB | National Cancer Database |
| NCPE | National Centre for Pharmacoeconomics |
| NCRAS | National Cancer Registration and Analysis Service |
| NG | NICE guideline |
| NGS | Next-generation sequencing |
| NHB | Net health benefit |

| | |
|---|---|
| NHS | National Health Service |
| NHSCII | NHS Cost Inflation Indices |
| NHSE | National Health Service England |
| NICE | National Institute for Health and Care Excellence |
| NMB | Net monetary benefit |
| NPI | Nottingham Prognostic Index |
| NR | Not reported |
| ONS | Office for National Statistics |
| OS | Overall survival |
| P | Paclitaxel |
| PAM50 | Prediction Analysis of Microarray 50 |
| PARP | Poly-ADP ribose polymerase |
| PAS | Patient Access Scheme |
| PD-L1 | Programmed death ligand 1 |
| PR | Progesterone receptor |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| PROBAST | Prediction model study Risk Of Bias Assessment Tool |
| PSA | Probabilistic sensitivity analysis |
| PSS | Personal Social Services |
| PSSRU | Personal Social Services Research Unit |
| QALY | Quality-adjusted life year |
| RCT | Randomised controlled trial |
| RFI | Request for information |
| RNA | Ribonucleic acid |
| RoB2 | Risk of Bias tool version 2 |
| ROR | Risk of Recurrence |
| RR | Relative risk |
| RS | Recurrence Score |
| RT-PCR | Reverse transcriptase polymerase chain reaction |
| RT-qPCR | Reverse transcription-quantitative polymerase chain reaction |
| SAE | Serious adverse event |
| SC | Standard care |
| SCI-E | Science Citation Index – Expanded |
| SD | Standard deviation |
| SE | Standard error |
| SEER | Surveillance Epidemiology and End Results |
| SLR | Systematic literature review |
| SMR | Standardised mortality ratio |
| S-TAI | State-Trait Anxiety Inventory |
| SWOG | Southwest Oncology Group |
| SWQ | South-West quadrant |
| TA | Technology Appraisal |
| TAC | Docetaxel, doxorubicin and cyclophosphamide |
| TC | Docetaxel and cyclophosphamide |
| TNM | Tumour node metastasis |
| TTO | Time trade-off |
| U&E | Urea and electrolytes |
| UICC | Union International Contre le Cancer |
| UK | United Kingdom |
| UKBCG | UK Breast Cancer Group |
| US | United States |
| USA | United States of America |
| VAT | Value Added Tax |
| WHO ICTRP | World Health Organization International Clinical Trials Registry Platform |
| WTP | Willingness-to-pay |

# 1 EXECUTIVE SUMMARY

## 1.1 Background

Breast cancer is the most commonly diagnosed cancer and the fourth most common cause of cancer-related death in the UK. During the period 2016-2018, an average of 46,479 women and 319 men were diagnosed with breast cancer in England each year. Initial treatment for breast cancer usually involves surgery to remove the primary tumour and some or all of the axillary lymph nodes. This may be followed by one or more of the following treatments: radiotherapy, endocrine (hormone) therapy, targeted therapy, bisphosphonates and/or chemotherapy. A proportion of patients may also receive neoadjuvant therapy prior to surgery, although this is primarily aimed at women with triple negative or human epidermal growth factor receptor 2 (HER2) positive breast cancers. Chemotherapy can reduce the likelihood of cancer recurrence and death for women with early breast cancer. Due to their increased risk of recurrence, most women with lymph node positive (LN+) disease receive adjuvant chemotherapy. However, chemotherapy is associated with considerable adverse effects. Improved information on a patient's risk of recurrence (i.e., their prognostic risk) and/or their likely response to chemotherapy (i.e., predictive benefit) may help clinicians to target chemotherapy to those patients who will benefit the most from treatment. Avoiding chemotherapy in patients at low risk of recurrence, who would therefore obtain limited benefit, avoids the unpleasant side effects of chemotherapy and may reduce expenditure on both the chemotherapy itself and the treatment of adverse effects. Tumour profiling tests aim to improve the use of chemotherapy in breast cancer by improving the categorisation of patients according to risk and the identification of those patients who will gain the most benefit from chemotherapy.

In 2018, the National Institute for Health and Care Excellence (NICE) published Diagnostics Guidance (DG) No. 34. DG34 recommends the use of Oncotype DX, Prosigna and EndoPredict (EPclin) for guiding chemotherapy decisions in people with oestrogen receptor (ER)-positive, HER2-negative, lymph node negative (LN0) early breast cancer, including those with micrometastases. Two other tests which were assessed in DG34 – MammaPrint and IHC4 – were not recommended in the LN0 population. Whilst DG34 also included an assessment of the use of these tests in women with lymph node positive (LN+) early breast cancer, the Appraisal Committee did not make any specific recommendations on the use of any tumour profiling test within this subgroup. This assessment provides an updated systematic literature review (SLR) and economic analysis of four tumour profiling tests (Oncotype DX, Prosigna, EPclin and MammaPrint) compared to current decision-making for use in women with ER (and/or progesterone receptor [PR]) positive, HER2-negative, early breast cancer with 1 to 3 positive lymph nodes.

## 1.2 Objectives

The main research question to be addressed is: *"Do tumour profiling tests used for guiding adjuvant chemotherapy decisions in patients with ER-positive (and/or PR-positive), HER2-negative, early-stage breast cancer with 1 to 3 positive lymph nodes represent a clinically effective and cost-effective use of NHS resources?"*

The objectives of the assessment are as follows:

- To conduct a systematic review of the published evidence on the effectiveness and cost-effectiveness of the four tumour profiling tests (Oncotype DX, Prosigna, EPclin and MammaPrint).
- To develop a health economic model to assess the cost-effectiveness associated with the use of tumour profiling tests compared with current prognostic tools to guide the use of adjuvant chemotherapy in early breast cancer from the perspective of the NHS and Personal Social Services (PSS).

## 1.3 Methods

### *1.3.1 Clinical evidence review methods*

The External Assessment Group (EAG) undertook a systematic review of the clinical effectiveness of Oncotype DX, Prosigna, EPclin and MammaPrint for guiding adjuvant chemotherapy decisions in women with ER+, HER2-, early breast cancer where the study population was at least 80% LN+. Studies were identified from the previous review which informed NICE DG34 plus an updated search for studies published since 2017. The search covered MEDLINE, EMBASE, Cochrane, other databases, and manufacturer submissions to NICE. Eligible data types included prospective randomised controlled trials (RCTs) of the tests, and studies of prognostic ability, prediction of chemotherapy benefit, impact of tests on chemotherapy decisions (the latter were restricted to UK and European studies), and health-related quality of life (HRQoL) and anxiety associated with testing. Risk of bias was assessed using tools specific to the study type. Results were presented via a narrative synthesis and tabulation.

### *1.3.1 Cost-effectiveness methods*

The EAG undertook a systematic review of existing economic analyses of Oncotype DX, Prosigna, EPclin and MammaPrint for guiding adjuvant chemotherapy decisions in women with ER+, HER2-, LN+ early breast cancer. Studies included published analyses which were identified within the previous systematic review undertaken to inform NICE DG34 as well as economic analyses in LN+ populations published since 2017. The EAG also reviewed and critically appraised economic analyses of Oncotype DX and MammaPrint submitted to NICE by the test manufacturers.

The EAG also developed a *de novo* health economic model to assess the cost-effectiveness of Oncotype DX, MammaPrint, Prosigna, and EndoPredict (EPclin), each compared against current decision-making. The health economic analysis was undertaken from the perspective of the NHS and PSS and was largely based on the model developed to inform NICE DG34, with updates to reflect changes in the breast cancer treatment pathway and updated evidence on the tests identified from the clinical effectiveness review. The EAG model adopts a hybrid decision tree/Markov structure. The model parameters were informed by a number of sources, including the RxPONDER, TransATAC, SWOG-8814, and MINDACT trials, an unpublished UK decision impact study undertaken in women with LN+ early breast cancer, previous economic models, routine costing sources and other literature. All results presented in this report reflect the list prices of the tumour profiling tests; additional analyses which include price discounts for the tests and for downstream treatments are provided in a separate confidential appendix to this report.

## 1.4 Results

### *1.4.1 Clinical evidence results*

*Overview of available evidence*

The search identified 4,057 articles. In total, 54 articles were included, 42 relating to prognostic and predictive ability, and 12 relating to impact on chemotherapy decisions. Studies of prognostic and predictive ability included retrospective reanalyses of trials and cohorts, observational studies of prospective use of tests, and two prospective RCTs (RxPONDER and MINDACT). In RxPONDER, patients with an Oncotype DX Breast Recurrence Score (RS) of ≤25 were randomised to chemotherapy vs. no chemotherapy. In MINDACT, patients' genomic risk (via MammaPrint) and clinical risk (via modified Adjuvant! Online [mAOL]) were assessed; patients who were low-risk on both measures were allocated to no chemotherapy, those who were high-risk on both were allocated to chemotherapy, and patients with discordant risk were randomised to chemotherapy vs. no chemotherapy. The ongoing OPTIMA RCT compares Prosigna test-directed chemotherapy use vs. standard chemotherapy use; however, results are not yet available from this study.

*Prognostic ability*

The prognostic ability of a test describes its ability to differentiate between patients with good versus poor outcomes. For all four tests, within re-analyses of trials and cohorts and observational studies, the hazard ratios (HR) for distant recurrence between risk groups indicated statistically significant prognostic ability for most (though not all) analyses, both with and without adjustment for clinical factors. In RxPONDER, within the study population (RS 0-25), Oncotype DX was significantly prognostic for 5-year invasive disease-free survival (IDFS) after adjusting for clinical factors, overall and in the pre-menopausal and post-menopausal subgroups. In MINDACT, within clinical high-risk LN+ patients, the 8-year distant metastasis-free interval (DMFI) was 92.3% for MammaPrint low-risk

vs. 80.9% for MammaPrint high-risk, despite higher chemotherapy use for high-risk patients; however, no HRs or significance tests were reported for prognostic ability.

*Prediction of chemotherapy benefit: Summary*

Whether a test is predictive concerns whether the effect of chemotherapy vs. no chemotherapy on patient outcomes differs between test risk groups or ranges, and is generally assessed via a statistical interaction test. Some data assessing predictive ability were identified for Oncotype DX and MammaPrint. No predictive data in a LN+ population were identified for Prosigna or EPclin.

*Prediction of chemotherapy benefit: Oncotype DX*

In a reanalysis of the SWOG-8814 RCT, Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy vs. no chemotherapy. For 10-year disease-free survival (DFS), using cut-offs of RS <18 and >30, adjusted HRs indicated no effect of chemotherapy in the low-risk group (HR 1.02; 95% confidence interval [CI] 0.54 to 1.93; *p*=0.97); a non-significant effect in the intermediate-risk group (HR 0.72; 95% CI 0.39 to 1.31; *p*=0.48); and a borderline statistically significant effect in the high-risk group (HR 0.59; 95% CI 0.35 to 1.01; *p*=0.033). Interaction tests for chemotherapy effect and risk group were statistically significant in some analyses, but not others. The RxPONDER RCT reported no benefit of chemotherapy in post-menopausal patients with an RS of 0-25 (difference in 5-year distant recurrence-free interval [DRFI] of 0.8% favouring no chemotherapy; adjusted HR 1.12; 95% CI 0.82 to 1.52; *p*=0.49). Conversely, there was chemotherapy benefit in pre-menopausal patients with an RS of 0-25 (difference in DRFI of 2.4% favouring chemotherapy; adjusted HR 0.64; 95% CI 0.43 to 0.95; *p*=0.026). A test for interaction between RS (within the range 0-25) and effect of chemotherapy on IDFS was not statistically significant across all patients (HR 1.02; 95% 0.98 to 1.05; *p*=0.35) or in the pre-menopausal or post-menopausal subgroups, indicating no significant predictive effect within the range RS 0-25. The NCDB database reported 5-year OS within post-menopausal or older-age subgroups with RS ≤25; some of these analyses showed a statistically significant chemotherapy benefit while others did not; therefore, the results did not clearly either support or refute the RxPONDER findings.

*Prediction of chemotherapy benefit: MammaPrint*

A reanalysis of two cohorts from 2009 reported a non-significant interaction test between MammaPrint score and effect of chemotherapy on BCSS (*p*=0.95) indicating no predictive effect. In the MINDACT prospective RCT, within the clinical high-risk, MammaPrint low-risk, LN+, hormone receptor positive (HR+) HER2- subgroup, 8-year distant metastasis-free survival (DMFS) was 91.2% with chemotherapy vs. 89.9% with no chemotherapy, an absolute difference of 1.3% favouring chemotherapy, with a non-significant HR (HR 0.84; 95% CI 0.51 to 1.37; *p*=NR). Since all patients in the clinical high-risk,

MammaPrint high-risk group were offered chemotherapy, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit.

*Decision impact*

Evidence on chemotherapy decisions pre- and post-testing in LN+ populations included 12 studies of Oncotype DX (five in the UK and seven in other European countries). No decision impact studies were identified for EPclin, Prosigna or MammaPrint. The net change in the percentage of patients with a chemotherapy recommendation or decision (pre-test to post-test) was a reduction of 28% to 75% across five UK studies, and a reduction of 12% to 73% across seven European studies. Across five studies using the Oncotype DX cut-offs of RS <18 and >30, the net change in chemotherapy recommendations or decisions was: a reduction of 20% to 93% in the RS 0-17 risk group; a reduction of 19% to 54% in the RS 18-30 risk group; and a 17% reduction to 1.7% increase in the RS >30 risk group. In two studies using cut-offs of RS 11 and 25 or RS 13 and 25, the net change in chemotherapy recommendations or decisions was: a reduction of 52% to 67% in the RS <11 or RS <13 risk groups; a reduction of 18% to 56% in the RS 11-25 or RS 14-25 risk groups; and 0% change to 5% increase in the RS >26 risk group.

*HRQoL and anxiety*

No studies reported HRQoL or anxiety associated with use of tumour profiling tests in a LN+ population. Across studies in a LN0 or mixed population, some reported significant improvements in anxiety after testing, whilst others reported no significant change. Some studies reported a decrease in anxiety after a low-risk test result or when treatment was downgraded to no chemotherapy, but an increase in anxiety after a high-risk test result or when treatment was upgraded to chemotherapy.

### *1.4.2 Cost-effectiveness results*

The results of the EAG's probabilistic base case analyses are summarised below.

*Oncotype DX*

Within the pre-menopausal LN+ population, Oncotype DX is dominated by usual care. These results are driven by the estimated reduction in the use of adjuvant chemotherapy due to the test in women who would have benefitted from treatment.

Within the post-menopausal LN+ subgroup, Oncotype DX dominates current decision-making, provided the assumption of predictive benefit holds. These results are driven by an estimated reduction in the use of adjuvant chemotherapy in women who would not have benefitted from treatment. As was the case with the economic analyses in the LN+ subgroup undertaken to inform DG34, removing this assumption of predictive benefit results in a situation whereby Oncotype DX is dominated by current decision-making (based on the older RS cut-offs). This assumption of predictive benefit remains subject

to some uncertainty and it strongly influences the conclusions of the economic analysis in the post-menopausal LN+ subgroup.

*Prosigna*

The incremental cost-effectiveness ratio (ICER) for Prosigna versus current decision-making is expected to be £39,357 per quality-adjusted life year (QALY) gained. The model suggests that the use of Prosigna will result in a small decrease in the use of chemotherapy, a small reduction in the lifetime probability of developing distant metastases (DM) and additional net costs due to the cost of the test. The EAG's systematic review did not identify any evidence to support a predictive benefit for Prosigna in the LN+ population.

*EndoPredict (EPclin)*

The ICER for EPclin versus current decision-making is expected to be £4,113 per QALY gained. The model suggests that the use of EPclin will result in a small decrease in adjuvant chemotherapy use, a small reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. The EAG's systematic review did not identify any evidence to support a predictive benefit for EPclin in the LN+ population.

*MammaPrint*

MammaPrint is dominated by current decision-making. These results are driven by a large reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment, an increase in the lifetime probability of developing DM and additional net costs due to the cost of the test. The EAG's systematic review did not identify sufficient evidence to support a predictive benefit for MammaPrint in the LN+ population.

### 1.5 Discussion

*Strengths and limitations in the clinical evidence base*

Strengths of the clinical evidence base include the fact that there is fairly substantial evidence for prognostic ability of all four tests. A major limitation is that it is difficult to collect new data on predictive ability because it is not considered ethical to randomise patients who are high-risk on any of the tests to chemotherapy vs. no chemotherapy. Therefore, although there are prospective RCTs for the effect of chemotherapy in low- to intermediate-risk patients, data for high-risk patients are limited to retrospective reanalyses of trials, plus observational data in which test results may have influenced treatment. Decision impact data in a LN+ population were available for Oncotype DX, but not for the other three tests. Anxiety and HRQoL data were not identified in a LN+ population.

*Strengths and limitations relating to the health economic analysis*

The EAG's model has several strengths: the economic analysis is consistent with the NICE Reference Case and relates specifically to the LN+ population under consideration within this appraisal; the model structure is generally consistent with most published economic models of tumour profiling tests as well as the two economic models submitted by the test manufacturers; where data permit, risk classification probabilities and DRFI estimates for each test have been taken from same source, which avoids the potential for spectrum bias; the analysis uses a recent relevant UK decision impact study undertaken in LN+ women; and a broad assessment of uncertainty around all key model inputs has been presented, including testing assumptions around whether Oncotype DX is predictive of chemotherapy benefit. The EAG notes that under similar assumptions around the benefits of each tumour profiling test, the EAG's model and the Exact Sciences model suggest similar economic conclusions.

The EAG's economic analyses are subject to several weaknesses: the EAG's analyses of Oncotype DX based on RxPONDER indirectly assume a predictive benefit which reflects a plausible clinical assumption about the effect of chemotherapy in women who were excluded from the trial (external data from SWOG-8814 are used to inform the benefit of chemotherapy in women with an RS of >25), rather than a statistical test of interaction across the full RS spectrum; there are inconsistencies in Oncotype DX RS cut-offs between sources used in the model; the analyses rely on a decision impact study of Oncotype DX to estimate post-test probabilities for all 2- and 3-level tests, which is highly uncertain; and there is insufficient evidence to allow for the economic analyses of EPclin and MammaPrint in an exclusively pre-menopausal subgroup. There is uncertainty around the potential negative effects of chemotherapy on infertility which may not be fully captured in the analysis of Oncotype DX in the pre-menopausal LN+ subgroup. The EAG's analyses of net health benefit (NHB) provide a means for the Appraisal Committee to decide whether any missing health effects are likely to impact on the conclusions drawn from the economic analysis.

### 1.6 Implications for service provision

Oncotype DX, Prosigna and EPclin are already recommended for use in the NHS for women with ER+ (and/or PR+), HER2-, LN0 early breast cancer. Depending on the specific test and population under consideration, tumour profiling may result in fewer women receiving adjuvant chemotherapy (reducing costs and increasing capacity), but this may lead to more women requiring further treatment for DM (increasing costs and reducing capacity).

MammaPrint is not currently recommended for use in the NHS. MammaPrint testing can be undertaken either as an off-site service with samples sent to a laboratory in the US, or a through a decentralised testing service for laboratories with next-generation sequencing (NGS) capability. The per-sample pricing of MammaPrint remains the same regardless of where the testing is performed. Not all

laboratories will have NGS capabilities which has implications for how MammaPrint testing is organised and delivered. For the other tests, only one sample processing approach is available – for Oncotype DX, samples are processed centrally at the Exact Sciences laboratory in the US, whereas for Prosigna and EPclin, samples are processed in local laboratories.

### 1.7 Suggested research priorities

Research priorities include the following:

- There remains some uncertainty around whether Oncotype DX is predictive of chemotherapy benefit. Further studies demonstrating a statistical interaction between Oncotype DX RS and long-term chemotherapy benefit across the full range of RS would help to address this uncertainty. However, such studies would require significant time and resources. Such studies may not be considered ethical as they may require chemotherapy to be withheld from patients who are high-risk via Oncotype DX.
- The review of HRQoL studies did not identify any new relevant studies which quantify the negative impact of adjuvant chemotherapy. Future longer-term studies are required to estimate short-term toxicity as well as longer-term negative health effects, including temporary and permanent effects on fertility in pre-menopausal women. Such studies should include the use of a preference-based instrument.
- The review did not identify any relevant decision impact studies for the use of Prosigna, EPclin or MammaPrint in the LN+ population. Further UK studies assessing the impact of tumour profiling tests on recommendations for adjuvant chemotherapy may help to reduce uncertainty around the cost-effectiveness of these tests.

# 2. BACKGROUND AND DEFINITION OF THE DECISION PROBLEM

## 2.1 Condition and aetiology

Breast cancer is the most commonly diagnosed cancer and the fourth most common cause of cancer-related death in the UK. During the period 2016-2018, an average of 46,479 women and 319 men were diagnosed with breast cancer in England each year.[1] Initial treatment for breast cancer usually involves surgery to remove the primary tumour and some or all of the axillary lymph nodes. Depending on the breast cancer characteristics, this may be followed by one or more of the following treatments: radiotherapy, endocrine (hormone) therapy, targeted therapy, bisphosphonates and/or chemotherapy. A proportion of patients also receive neoadjuvant therapy prior to surgery, although this is primarily aimed at women with triple negative or HER2-positive breast cancers.

### *2.1.1 Aetiology, pathology and prognosis*

*Aetiology*

The causes of breast cancer are not completely understood but involve a complex interplay of inherited genetic and environmental factors on a range of oncogenes and tumour suppressor genes. Multiple risk factors have been identified including older age, early menopause, late menarche, family history, and genetic, hormonal and lifestyle factors such as obesity, smoking and alcohol consumption.[2]

*Pathology*

Breast carcinogenesis starts with genetic changes in a single or small group of cells in the epithelia of the ducts or the lobules of the breast. The genetic change allows cells to reproduce uncontrollably, and this, alongside numerous other cellular changes (summarised as the Hallmarks of Cancer), leads to cancer. Tumours that have not yet spread beyond the basement membrane of the milk ducts, into surrounding tissues are known as "carcinoma *in situ.*" Once the tumour begins to spread to the surrounding tissue, the tumour is known as "invasive". Once a blood supply is secured, more rapid growth and spread occurs. Cancer spreads by local infiltration and via the lymphatic system or the bloodstream. Lymphatic spread is usually first to the axillary lymph nodes in the armpit. Spread via the bloodstream can lead to distant metastases in the bone or viscera at which stage the disease is regarded as incurable.

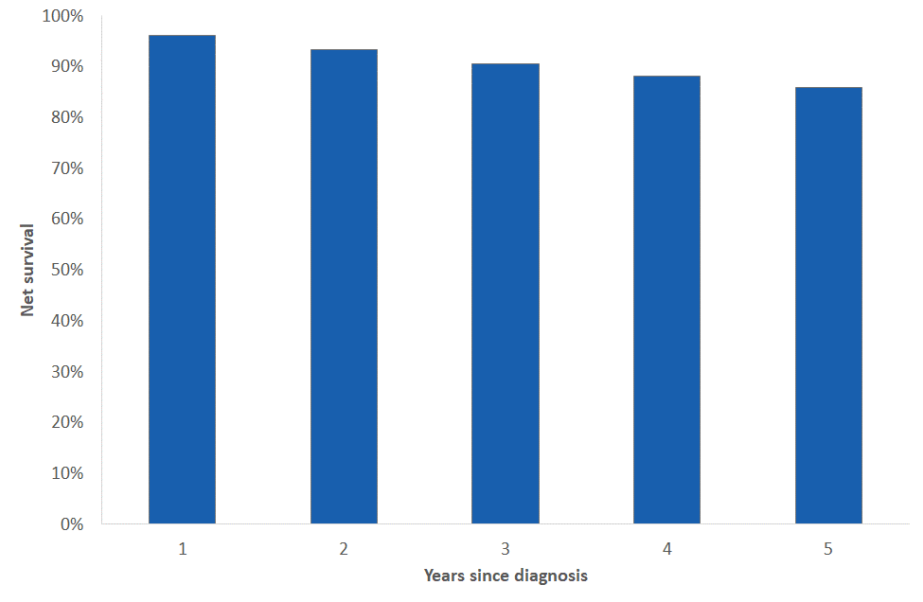
The presence or absence of axillary lymph node metastases is a key indicator of disease prognosis and adjuvant therapy is, in part, planned based on their presence and extent.[3] They are caused when a single or small numbers of cells detach from the main tumour, travel via the lymphatic system and establish themselves in the tissue of the axillary lymph nodes. Axillary metastases occur in approximately 41% of cases;[4] prognosis is better where there is no axillary spread. Nodal involvement is defined according

to both the number of affected nodes and the size of the disease focus in the node. Isolated tumour cells are not regarded as an indication for further surgery or use of adjuvant therapy and are largely ignored clinically (except in the post neoadjuvant setting). Larger nodal foci are classified as macro or micro metastases depending on whether they are greater than or less than 2mm. Micrometastases are used to guide chemotherapy decision-making but are not an indication for axillary clearance (again with the exception of post-neoadjuvant therapy). Macrometastases are both used to guide chemotherapy use and further axillary surgery. However, modern de-escalation paradigms now mean that axillary clearance is no longer mandatory if sentinel node biopsy yields macrometastases. Some women with a low disease burden may be offered axillary radiotherapy or even no further axillary treatment as an alternative. Where multiple or bulky nodal metastases are present, axillary clearance is still indicated to optimise local disease control.

*Prognosis*

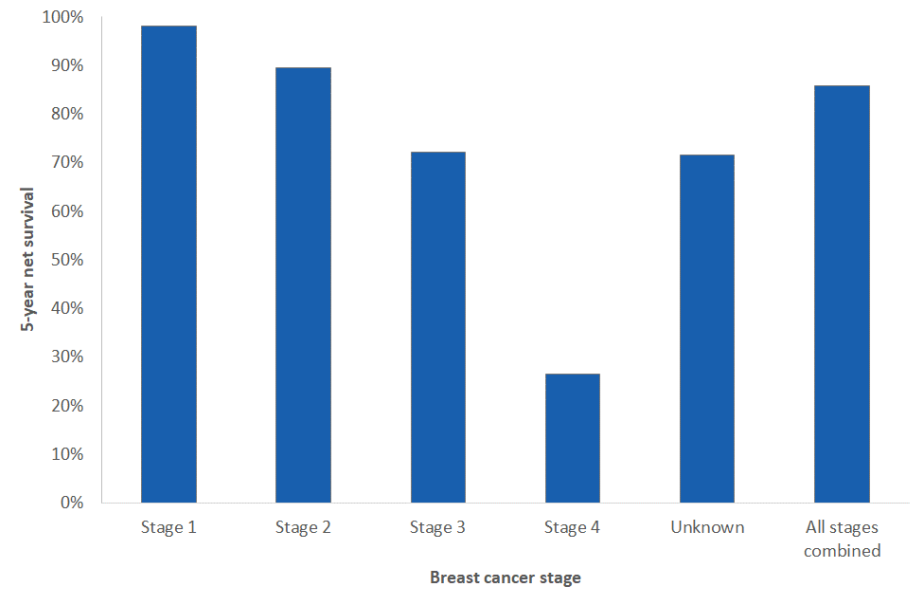
Age-standardised net survival according to time since breast cancer diagnosis is summarised in Figure 1, based on data for England published by NHS Digital.[5] The age-standardised 5-year net survival for women with breast cancer diagnosed between 2016 and 2020 is estimated to be around 86%. Net survival according to stage at diagnosis is shown in Figure 2.[5] The 5-year net survival for people with breast cancer varies by disease stage, with the highest survival in Stage 1 and the lowest survival in people with Stage 4 (metastatic) disease.

**Figure 1: 5-year net survival by time since diagnosis, adults with breast cancer diagnosed in 2016 to 2020, followed up to 2021**




*Source: NHS Digital, National Disease Registration Service*

**Figure 2: 5-year net survival by stage, adults with breast cancer diagnosed in 2016 to 2020, followed up to 2021**



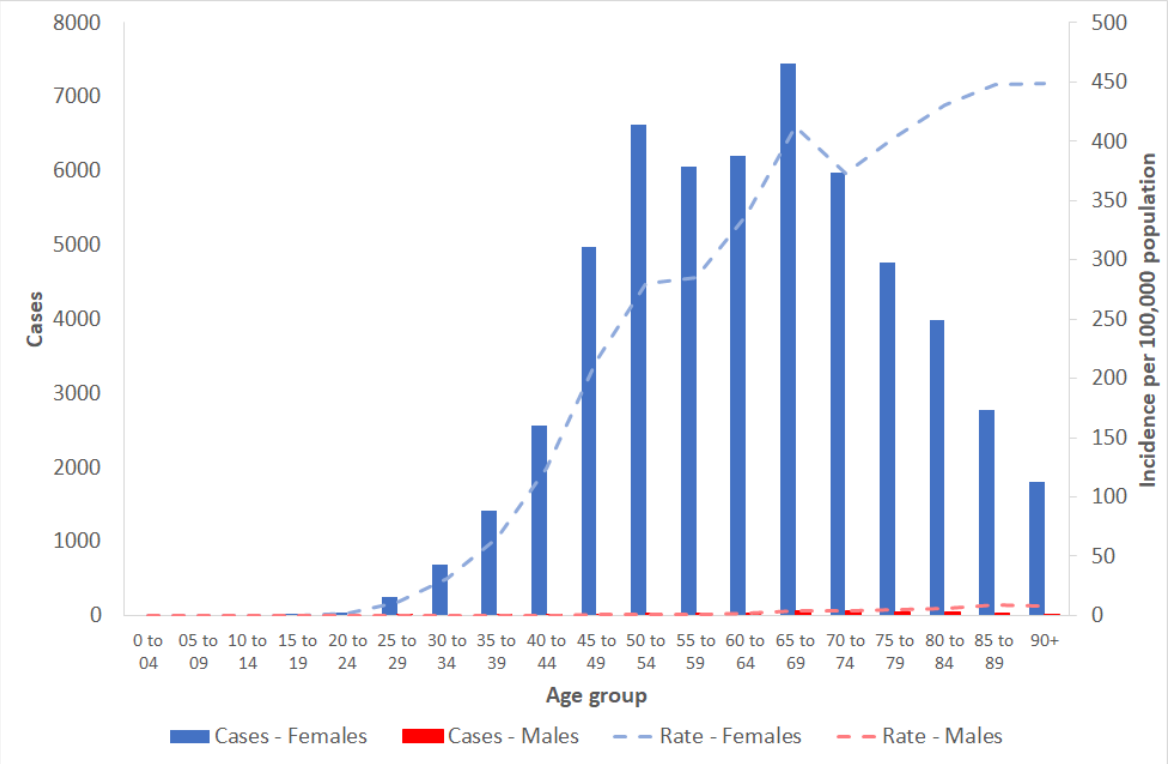

*Source: NHS Digital, National Disease Registration Service*

Several clinical and pathological factors affect prognosis. In general, good prognosis is associated with small tumour size, lymph node-negative (LN0) status, certain age groups (40-70 years), oestrogen receptor positive (ER+) and progesterone receptor positive (PR+) tumour biology. Overexpression of human epidermal growth factor receptor type 2 (HER2) is associated with poorer prognosis. The population under consideration within this appraisal relates specifically to people with ER+/PR+, HER2- early breast cancer and 1 to 3 positive lymph nodes (LN1-3).

### *2.1.2 Epidemiology and incidence*

Figure 3 presents estimates of breast cancer incidence by age and sex for the UK, based on data from 2016-2018 reported by Cancer Research UK.[6] Breast cancer incidence varies most according to gender. Women are considerably more likely to develop breast cancer than men. For both males and females, incidence generally increases with age. Over 82% of cases of breast cancer occur in people aged 50 years and over and approximately 24% of cases are in people aged 75 years and older.

**Figure 3: Breast cancer incidence by age and sex, 2016-2018, UK**




*Source: Cancer Research UK*

### *2.1.3 Significance in terms of ill-health (burden of disease)*

Breast cancer is the second largest cause of cancer death in women after lung cancer, with an age-standardised mortality rate of 32.8 per 100,000 women. The age-standardised mortality rate in men is substantially lower at 0.3 per 100,000 men. During the period 2017-2019, an average of 9,509 women and 69 men died from breast cancer in England each year.[7]

### *2.1.4 Current methods for staging of breast cancer*

Breast cancer staging takes into account three main factors: (i) tumour size; (ii) metastases to the regional lymph nodes, and (iii) the presence/absence of distant metastases.[8] The tumour/node/metastases (TNM) staging system was developed and is maintained by the American Joint Committee on Cancer (AJCC) and the Union International Contre le Cancer (UICC). Version 8 of the AJCC TNM staging system was published in 2018.[9] According to this staging system, T stage is classified according to size of the tumour and degree of local infiltration; N stage is classified according to the number and location of metastases to the lymph nodes in the axilla, between the ribs (internal mammary nodes) and above or below the collarbone (supraclavicular and infraclavicular nodes); and M stage is classified by the presence of metastases beyond the breast and regional lymph nodes (see Table 1). The overall TNM stage of the cancer is defined as shown in Table 2. Early breast cancer is generally defined as cancer which has not spread beyond the breast or the ipsilateral axillary lymph nodes, and is confined to Stages I, II or IIIA.

**Table 1: Breast cancer staging, AJCC, version 8**

| | |
|---|---|
| **Primary tumour (T)** | |
| Tx | Primary tumour cannot be assessed |
| T0 | No evidence of primary tumour |
| Tis | Carcinoma *in situ* |
| Tis (DCIS)* | Ductal carcinoma *in situ* |
| Tis (Paget) | Paget disease of the nipple NOT associated with invasive carcinoma and/or carcinoma *in situ* (DCIS) in the underlying breast parenchyma. Carcinomas in the breast parenchyma associated with Paget disease are categorized based on the size and characteristics of parenchymal disease, although the presence of Paget disease should still be noted. |
| T1 | Tumour ≤ 20 mm in greatest dimension |
| T1mi | Tumour ≤ 1 mm in greatest dimension |
| T1a | Tumour > 1 mm but ≤ 5 mm in greatest dimension (round any measurement >1.0-1.9 mm to 2 mm) |
| T1b | Tumour > 5 mm but ≤ 10 mm in greatest dimension |
| T1c | Tumour > 10 mm but ≤ 20 mm in greatest dimension |
| T2 | Tumour > 20 mm but ≤ 50 mm in greatest dimension |
| T3 | Tumour > 50 mm in greatest dimension |
| T4 | Tumour of any size with direct extension to the chest wall and/or to the skin (ulceration or macroscopic nodules); invasion of the dermis alone does not qualify as T4 |
| T4a | Extension to the chest wall; invasion or adherence to pectoralis muscle in the absence of invasion of chest wall structures does not quality as T4 |
| T4b | Ulceration and/or ipsilateral macroscopic satellite nodules and/or edema (including peau d'orange) of the skin that does not meet the criteria for inflammatory carcinoma |
| T4c | Both T4a and T4b |
| T4d | Inflammatory carcinoma |
| **Distant Metastases (M)** | |
| M0 | No clinical or radiographic evidence of distant metastases |
| cM0(i+)† | No clinical or radiographic evidence of distant metastases in the presence of tumor cells or deposits no larger than 0.2 mm detected microscopically or by molecular techniques in circulating blood, bone marrow, or other nonregional nodal tissue in a patient without symptoms or signs of metastases |
| cM1 | Distant detectable metastases detected by clinical and radiographic means |
| pM1 | Any histologically proven metastases in distant organs; or if in non-regional nodes, metastases greater than 0.2 mm |
| **Regional Lymph Nodes (N)** | |
| **Clinical** | |
| cNX‡ | Regional lymph nodes cannot be assessed (e.g., previously removed) |
| cN0 | No regional lymph node metastases (by imaging or clinical examination) |
| cN1 | Metastases to movable ipsilateral level I, II axillary lymph node(s) |
| cN1mi¶ | Micrometastases (approximately 200 cells, larger than 0.2 mm, but none larger than 2.0 mm) |
| cN2 | Metastases in ipsilateral level I, II axillary lymph nodes that are clinically fixed or matted; or in ipsilateral internal mammary nodes in the absence of axillary lymph node metastases |
| cN2a | Metastases in ipsilateral level I, II axillary lymph nodes fixed to one another (matted) or to other structures |
| cN2b | Metastases only in ipsilateral internal mammary nodes in the absence of axillary lymph node metastases |
| cN3 | Metastases in ipsilateral infraclavicular (Level III axillary) lymph node(s) with or without Level I, II axillary lymph node involvement; or in ipsilateral internal mammary node(s) with Level I, II axillary lymph node metastases; or metastases in ipsilateral supraclavicular lymph node(s) with or without axillary or internal mammary lymph node involvement |
| cN3a | Metastases in ipsilateral infraclavicular lymph node(s) |
| cN3b | Metastases in ipsilateral internal mammary lymph node(s) and axillary lymph node(s) |
| cN3c | Metastases in ipsilateral supraclavicular lymph node(s) |

| **Pathologic (PN)** | |
|---|---|
| pNX | Regional lymph nodes cannot be assessed (e.g., not removed for pathological study or previously removed) |
| pN0 | No regional lymph node metastasis identified or ITCs only |
| pN0(i+) | ITCs only (malignant cell clusters no larger than 0.2 mm) in regional lymph node(s) |
| pN0(mol+) | Positive molecular findings by reverse transcriptase polymerase chain reaction (RT-PCR); no ITCs detected |
| pN1 | Micrometastases; or metastases in 1–3 axillary lymph nodes; and/or clinically negative internal mammary nodes with micrometastases or macrometastases by sentinel lymph node biopsy |
| pN1mi | Micrometastases (approximately 200 cells, larger than 0.2 mm, but none larger than 2.0 mm) |
| pN1a | Metastases in 1–3 axillary lymph nodes, at least one metastasis larger than 2.0 mm |
| pN1b | Metastases in ipsilateral internal mammary sentinel nodes, excluding ITCs |
| pN1c | pN1a and pN1b combined |
| pN2 | Metastases in 4–9 axillary lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the absence of axillary lymph node metastases |
| pN2a | Metastases in 4–9 axillary lymph nodes (at least one tumor deposit larger than 2.0 mm) |
| pN2b | Metastases in clinically detected internal mammary lymph nodes with or without microscopic confirmation; with pathologically negative axillary nodes |
| pN3 | Metastases in 10 or more axillary lymph nodes; or in infraclavicular (Level III axillary) lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the presence of one or more positive Level I, II axillary lymph nodes; or in more than three axillary lymph nodes and micrometastases or macrometastases by sentinel lymph node biopsy in clinically negative ipsilateral internal mammary lymph nodes; or in ipsilateral supraclavicular lymph nodes |
| pN3a | Metastases in 10 or more axillary lymph nodes (at least one tumor deposit larger than 2.0 mm); or metastases to the infraclavicular (Level III axillary lymph) nodes |
| pN3b | pN1a or pN2a in the presence of cN2b (positive internal mammary nodes by imaging); or pN2a in the presence of pN1b |
| pN3c | Metastases in ipsilateral supraclavicular lymph nodes |

** Lobular carcinoma in situ (LCIS) is a benign entity and is removed from TNM staging in the AJCC Cancer Staging Manual, 8th Edition*

*† Imaging studies are not required to assign the cM9 category*

*‡ The cNX category is used sparingly in cases where regional lymph nodes have previously been surgically removed or where there is no documentation of physical examination of the axilla*

*¶ cN1mi is rarely used but may be appropriate in cases where sentinel node biopsy is performed before tumour resection, most likely to occur in cases treated with neoadjuvant therapy*

**Table 2: Summary of TNM stages**

| Stage | T | N | M |
|---|---|---|---|
| Stage 0 | Tis | N0 | M0 |
| Stage IA | T1 | N0 | M0 |
| Stage IB | T0 | N1mi | M0 |
| | T1 | N1mi | M0 |
| Stage IIA | T0 | N1 | M0 |
| | T1 | N1 | M0 |
| | T2 | N0 | M0 |
| Stage IIB | T2 | N1 | M0 |
| | T3 | N0 | M0 |
| Stage IIIA | T0 | N2 | M0 |
| | T1 | N2 | M0 |
| | T2 | N2 | M0 |
| | T3 | N1 | M0 |
| | T3 | N2 | M0 |
| Stage IIIB | T4 | N0 | M0 |
| | T4 | N1 | M0 |
| | T4 | N2 | M0 |
| Stage IIIC | Any T | N3 | M0 |
| Stage IV | Any T | Any N | M1 |

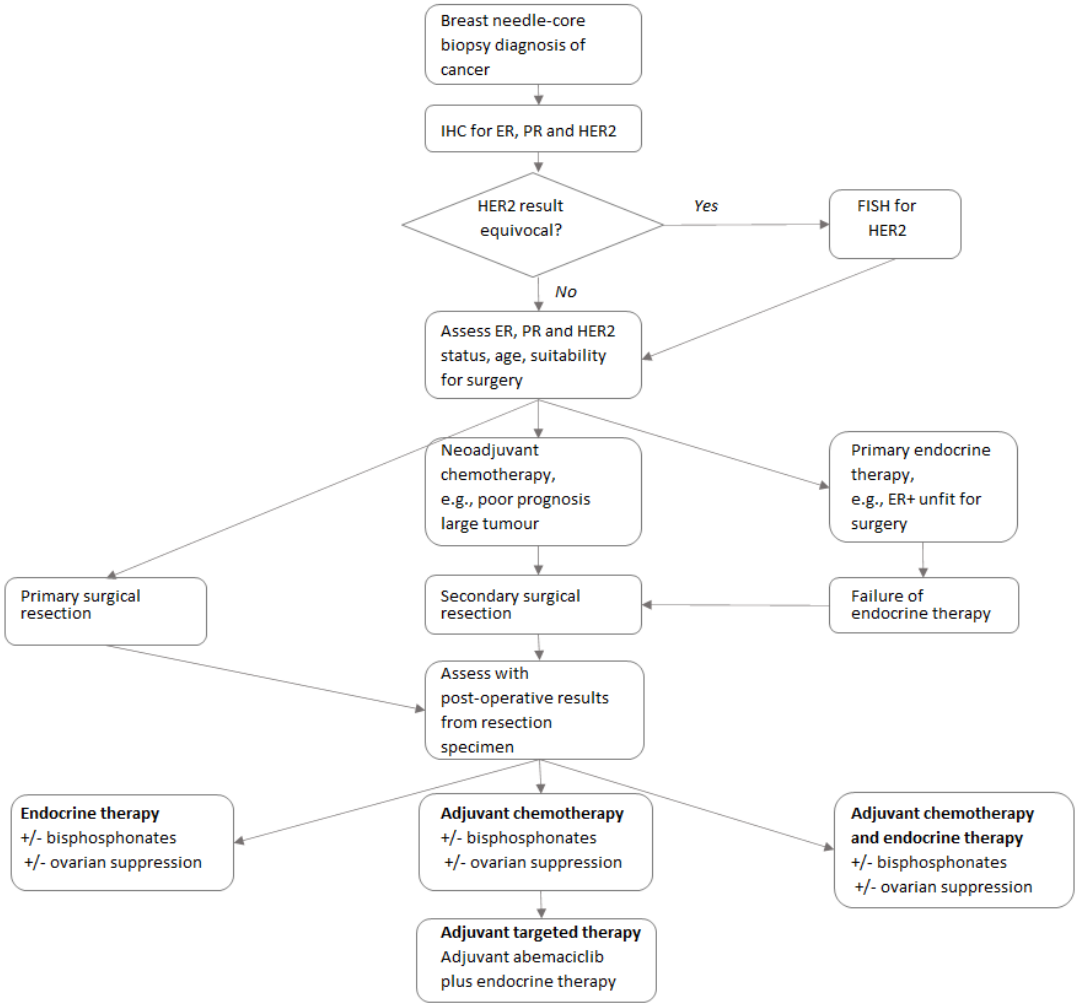
*T – tumour; N – node; M – metastasis; mi – micrometastases*

## 2.2 Current service provision

### *2.2.1 Management of early breast cancer*

NICE Guideline (NG) 101[3] provides recommendations on the diagnosis and management of early and locally advanced breast cancer. The guideline was first published in 2009 and was updated in 2018 and again in July 2023. The general treatment pathway for women with early breast cancer is summarised in Figure 4. Key recommendations for the diagnosis and management of early breast cancer are summarised in the subsequent sections, based on NG101,[3] Harnan *et al.*,[10] the summary provided in the NICE scope[11] and additional information provided by the clinical advisors to the External Assessment Group (EAG).

**Figure 4: Diagnosis and management pathway for early breast cancer**




***Footnotes:***
*Ki67 is tested for following biopsy in some centres, however, the methodology for this is not standardised*
*Extended endocrine therapy may be given for 10 years*
*If abemaciclib is given, this would normally be started after completion of a course of adjuvant chemotherapy*
*Ovarian suppression should be considered only in pre-menopausal women*
*Bisphosphonates are recommended for use only in post-menopausal women*
*Adjuvant and neoadjuvant treatment for HER2+ early breast cancer may include pertuzumab and trastuzumab or trastuzumab alone. In the neoadjuvant setting in poor responders, TDM1 may be given after surgery instead of continuing with the neoadjuvant regimen. This population is out of scope.*
*Women with triple negative breast cancer who have neoadjuvant chemotherapy and respond poorly may now be offered post-operative capecitabine chemotherapy.*
*Immunotherapy may be used in the neoadjuvant setting for women with triple receptor-negative breast cancer. This population is out of scope.*
*Women with known BRCA1 or 2 gene mutations are now eligible for adjuvant PARP inhibitors*
*Radiotherapy may also be offered depending on the type of surgery done and the patient's risk of recurrence*
*IHC – immunohistochemistry; ER – oestrogen receptor; PR – progesterone receptor; HER2 – human epidermal growth factor receptor 2; FISH – fluorescence in situ hybridisation*

### *Surgical resection and neoadjuvant treatments*

The initial treatment for early and some locally advanced breast cancers usually involves the surgical resection of the primary tumour. Surgical options to remove the disease in the breast include breast conserving surgery or mastectomy (where the whole breast is removed). If appropriate, women are offered the option to have reconstruction at the time of the initial surgery, or at a later date. Neoadjuvant systemic treatment may be given prior to surgery, with the aim of reducing the size of the tumour to

enable breast conserving surgery. Depending on whether clinical or ultrasound visible axillary disease is present, axillary surgery is also performed, involving a sentinel lymph node biopsy (SLNB) if the nodes are not thought to be involved and an axillary clearance if there is upfront nodal disease. Increasingly, for women with clinically involved nodes, where a good response to neoadjuvant chemotherapy is anticipated (triple negative or HER2+ breast cancer), chemotherapy will be given first to attempt to downstage the axilla.

In women who have a SLNB, if there is heavy nodal disease then a subsequent clearance is performed. If only micrometastases, isolated tumour cells or just 1 or 2 nodes are involved, then axillary radiotherapy or no formal axillary treatment is indicated. These strategies reduce the risk of adverse events (AEs) such as lymphoedema without a negative impact on survival.

*Adjuvant therapy planning*

After surgery, adjuvant treatment may be needed to treat residual micrometastatic disease following surgery and to reduce the risk of local and distant relapse. Adjuvant treatment may involve chemotherapy, endocrine therapy (ET), targeted therapy, radiotherapy or a combination of these treatments. The decision to offer, and the selection of, adjuvant therapy is made taking into account the patient's clinical history, the patient's fitness and health status, the stage of disease, the patient's likely prognosis, the molecular characteristics of the tumour and the patient's preferences. NG101[3] makes the following recommendations on adjuvant treatment planning:

- Consider adjuvant therapy after surgery for people with invasive breast cancer, and ensure that recommendations are recorded at the multidisciplinary team meeting.
- Base recommendations about adjuvant therapy on multidisciplinary team assessment of the prognostic and predictive factors, and the possible risks and benefits of the treatment. Make decisions with the person after discussing these factors.
- Use the PREDICT tool[12] (https://breast.predict.nhs.uk/tool) to estimate prognosis and the absolute benefits of adjuvant therapy for women with invasive breast cancer.
- When using version 2.0 of the PREDICT tool, be aware that:
    - it is less accurate for:
        - women under 30 with ER+ breast cancer
        - women aged 70 years and over
        - women with tumours larger than 5 cm
    - it has not been validated in men
    - the validation may have under-represented some ethnic groups.
- Note the potential limitations in versions of PREDICT after 2.0 may differ from those listed here.

The EAG's clinical advisors also commented that PREDICT version 2.0 has not been validated in pregnant women and that it may be less accurate for patients treated with neoadjuvant systemic therapy, patients aged 65 years and over, patients with a high comorbidity burden and patients with multifocal breast cancer, bilateral breast cancer, rare breast cancer subtypes or two different breast cancers. They also commented that PREDICT may be less accurate in the context of contemporary systemic treatment standards of care.

Whilst NG101[3] recommends the use of PREDICT to provide prognostic information on breast cancer recurrence and absolute chemotherapy benefit to guide decisions about the use of adjuvant chemotherapy, several other prognostic tools are also available which can help to predict the likelihood of breast cancer recurrence. These tools are described in Section 3.2.3.

NG101[3] also refers to recommendations from NICE Diagnostics Guidance (DG) 34 on the use of tumour profiling tests to guide adjuvant chemotherapy decisions.[13] Three tumour profiling tests (Oncotype DX, Prosigna and EndoPredict) are currently recommended for use in women with lymph node negative (LN0) early breast cancer, including those with micrometastases. DG34 did not make any specific recommendations on the use of these tests for women with lymph node positive (LN+) early breast cancer. The tumour profiling tests which are included as interventions within this appraisal are described in Section 3.3.

*Endocrine therapy*

ET may be offered to people who have ER+ or PR+ breast cancer. ET stops the growth of the cancer by blocking the availability of hormones such as oestrogen and progesterone by reducing production (aromatase inhibitors [Ais]), receptor antagonism (tamoxifen) or degradation of the oestrogen receptor (fulvestrant). NG101[3] makes the following recommendations on the use of ET:

- Offer tamoxifen as the initial adjuvant ET for men and pre-menopausal women with ER+ invasive breast cancer unless, in a pre-menopausal woman she is also receiving ovarian suppression therapy when exemestane may be used.
- Offer an AI as the initial adjuvant ET for post-menopausal women with ER+ invasive breast cancer who are at medium or high risk of disease recurrence. Offer tamoxifen to women who are at low risk of disease recurrence, or if Ais are not tolerated or are contraindicated.
- Offer extended therapy (total duration of ET of more than 5 years) with an AI for post-menopausal women with ER+ invasive breast cancer who are at medium or high risk of disease recurrence and who have been taking tamoxifen for 2 to 5 years. Medium or high risk may include people who have LN+ breast cancer, with tumours that are T2 or greater and higher grade.

- Consider extended therapy (total duration of ET of more than 5 years) with an AI for post-menopausal women with ER+ invasive breast cancer who are at low risk of disease recurrence and who have been taking tamoxifen for 2 to 5 years. Low risk may include people with LN0 breast cancer, with smaller or lower-grade tumours.
- Consider extending the duration of tamoxifen therapy for longer than 5 years for both pre-menopausal and post-menopausal women with ER+ invasive breast cancer.
- Discuss the benefits and risks of extended ET with women.

*Adjuvant chemotherapy*

Adjuvant chemotherapy may be offered to women to reduce the risk of distant metastases, local recurrence and death. NG101[3] makes several recommendations on the use of adjuvant chemotherapy, including:

- For people with breast cancer of sufficient risk that chemotherapy is indicated, offer a third-generation regimen that contains both a taxane and an anthracycline. Please refer to the summaries of product characteristics for individual taxanes and anthracyclines because there are differences in their licensed indications.
- Discuss with people the benefits and risks of adding a taxane to anthracycline-containing regimens.
- Weekly and fortnightly paclitaxel should be available locally because these regimens may be tolerated better than 3-weekly docetaxel, particularly in people with comorbidities and older age.

*Bisphosphonates*

Bisphosphonates are used to slow down or prevent damage to bone and to prevent and treat osteoporosis. In women with breast cancer they have also been shown to reduce the risk of breast cancer recurrence, especially in the bones, in post-menopausal women. They are also used in women who are receiving AI therapy if they have reduced bone density. NG101[3] makes the following recommendations on adjuvant bisphosphonate therapy for people with LN+ breast cancer:

- Offer bisphosphonates (zoledronic acid or sodium clodronate) as adjuvant therapy to post-menopausal women with LN+ invasive breast cancer.
- Discuss the benefits and risks of bisphosphonate treatment with women, particularly the risk of osteonecrosis of the jaw, atypical femoral fractures and osteonecrosis of the external auditory canal. Follow the Medicines and Healthcare products Regulatory Agency/Commission on Human Medicines (MHRA/CHM) advice on bisphosphonates.

*Ovarian suppression*

Ovarian suppression treatment stops or reduces the amount of oestrogen made by the ovaries. NG101[3] makes the following recommendations regarding the use of ovarian suppression:

- Consider ovarian function suppression in addition to ET for pre-menopausal women with ER+ invasive breast cancer.
- Discuss the benefits and risks of ovarian function suppression in addition to ET with pre-menopausal women with ER+ invasive breast cancer. Explain to women that ovarian function suppression may be most beneficial for those women who are at sufficient risk of disease recurrence to have been offered chemotherapy.

*Adjuvant targeted therapy*

In the high-risk population, adjuvant targeted therapy may be used to reduce the risk of disease recurrence and is usually used in people who have previously completed a course of adjuvant chemotherapy. NICE Technology Appraisal (TA) 810[14] recommends abemaciclib in combination with ET as an option for the adjuvant treatment of hormone receptor positive (HR+), HER2-, LN+ early breast cancer in adults whose disease is at high risk of recurrence, defined by the following clinical and pathological features:

- at least 4 positive axillary lymph nodes, or
- 1 to 3 positive axillary lymph nodes, and at least one of the following criteria:
    - grade 3 disease (defined as at least 8 points on the modified Bloom–Richardson grading system or equivalent), or
    - primary tumour size of at least 5 cm.

Other targeted therapies such as trastuzumab, pertuzumab, neratinib, programmed death ligand 1 (PD-L1) inhibitors and capecitabine are only relevant to women with triple negative or HER2 positive breast cancer and as such are outside of the scope of this appraisal. NICE has recently issued a positive recommendation for the use of olaparib (alone or with ET) as an option for the adjuvant treatment of HER2-negative high-risk early breast cancer that has been treated with neoadjuvant or adjuvant chemotherapy in adults with germline breast cancer gene (BRCA) 1 or 2 mutations.

*Radiotherapy*

Radiotherapy to the breast and/or axilla may be used to reduce the risk of locoregional recurrence (LRR) following breast surgery. The specific radiation approach depends on the patient's age, their preferences, the location of the tumour, lymph node involvement, the type of surgery undertaken, and whether clear resection margins have been achieved. NG101[3] provides recommendations on the use of radiotherapy; however, these are not discussed here.

### *2.2.3 Prognostic risk prediction tools*

A number of prognostic risk prediction tools have been developed which estimate the risk of relapse and/or death conditional on clinical and pathological factors. These include the Nottingham Prognostic Index (NPI), Adjuvant! Online (AOL) and PREDICT. The factors included in the prediction algorithms and the outcomes predicted by these tools are summarised in Table 3.

**Table 3: Breast cancer risk prediction tools**

| Tool | NPI | AOL | PREDICT (Version 2.2) |
|---|---|---|---|
| **Factors included in the prediction algorithm** | • Tumour size<br>• Nodal status<br>• Tumour grade | • Age at diagnosis<br>• Comorbidity factors<br>• ER status<br>• Tumour size<br>• Tumour grade<br>• Nodal status | • Age at diagnosis<br>• Menopausal status<br>• Mode of detection<br>• Invasive tumour size<br>• Tumour grade<br>• Number of positive nodes<br>• ER status<br>• HER2/ERBB2 status<br>• Ki67 status<br>• Generation of chemotherapy regimen |
| **Outcome(s) predicted** | Mortality | Mortality or relapse | Mortality |

*NPI – Nottingham Prognostic Index; AOL – Adjuvant! Online; ER – oestrogen receptor; HER2 – human epidermal growth factor receptor 2; ERBB – erythroblastic oncogene B*

*Nottingham Prognostic Index (NPI)*

The NPI is a composite prognostic parameter involving both time-dependent factors and aspects of biological aggressiveness. The NPI score is based on a combination of tumour grade, lymph node involvement and tumour size. To calculate the score: ADD numerical grade (1, 2, or 3), lymph node score (LN0 = 1, 1 to 3 nodes = 2, >3 nodes = 3) and 0.2* tumour size in cm. Patients can be divided into three prognostic groups on the basis of the NPI: a good prognostic group (NPI < 3.4), a moderate prognostic group (3.4 < NPI < 5.4), and a poor prognostic group (NPI > 5.4). Most women with LN+ breast cancer fall into the NPI moderate and poor prognosis groups due to the presence of lymph node involvement.

*Adjuvant! Online (AOL)*

The AOL computer program was designed to provide estimates of the benefits of adjuvant ET and chemotherapy. The most recently available version of AOL did not include HER2 status and the potential benefit of trastuzumab. Patient and tumour characteristics are entered into the program and provide an estimate of the baseline risk of mortality or relapse for patients without adjuvant therapy. Information about the efficacy of different therapy options were derived from meta-analyses conducted by the Early Breast Cancer Trialists' Collaborative Group (EBCTCG)[15] in order to provide estimates of reduction in risk at 10-years of breast cancer related death or relapse for selected treatments. These estimates were then provided on printed sheets in simple graphical and text formats to be used during

clinical consultations. AOL has not been available since 2016. However, this tool has been used to determine clinical risk in some of the studies included in this assessment.

*PREDICT (Version 2.2)*

PREDICT is an online computer program designed to help women with breast cancer and their doctors make informed decisions about treatment with chemotherapy or ET following breast cancer surgery. PREDICT was developed using data from over 5,000 women with breast cancer from England and has been tested on data from another 23,000 women with breast cancer from around the world. Patient and tumour characteristics are entered into the program, which provides an estimate of the overall survival for patients with or without adjuvant hormone therapy, adjuvant chemotherapy and trastuzumab. The most recent version of PREDICT is Version 2.2, which includes an option for predicting 10- and 15-year outcomes and factors in the effect of receiving extended ET for 10 years.

The EAG's clinical advisors noted that there is variation in clinical practice in how breast cancer doctors decide whether to recommend adjuvant chemotherapy for women with LN+ early breast cancer, with some centres using risk prediction tools and others using clinical-pathological information without the use of a quantitative risk prediction tool.

## 2.3 Description of technologies under assessment

### *2.3.1 The potential value of tumour profiling tests to guide chemotherapy decisions for women with LN+ early breast cancer*

Meta-analyses of randomised clinical trials reported by the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) have indicated that the use of adjuvant chemotherapy is associated with a reduction in the risk of distant recurrence and death in women with early stage breast cancer.[15] Lymph node involvement is associated with an increased risk of recurrence; hence, the majority of women with LN+ early breast cancer in England currently receive adjuvant chemotherapy.[16, 17] However, chemotherapy is also associated with considerable AEs, including both short- and long-term effects. These AEs negatively impact on patients' health-related quality of life (HRQoL) and result in additional health care costs. Short-term toxicity that occurs during chemotherapy is usually temporary and reversible and commonly includes: nausea; vomiting; mouth soreness; diarrhoea; tiredness; liver damage; diarrhoea and constipation; skin rash and nail changes; hair loss and temporary lowering of the blood counts which can lead to hospitalisation due to neutropenic sepsis and death. Chemotherapy is also associated with a risk of late effects, including damage to the heart, temporary or permanent amenorrhea, peripheral neuropathy, and a small increase in the risk of secondary malignancies including leukaemia.[18] Adjuvant chemotherapy may prevent distant recurrence for some women with early breast cancer, whilst others will not obtain benefit from treatment, with many women remaining recurrence-free at 10 years without chemotherapy.[19] This presents a challenge for clinicians in estimating prognosis and

making the most appropriate therapeutic decisions regarding whether or not to offer adjuvant chemotherapy to women with early stage breast cancer. Improved information on a patient's risk of recurrence (i.e., prognostic risk) and/or likely response to chemotherapy (i.e., predictive benefit) may help target chemotherapy at those patients who will benefit the most from treatment. Avoiding chemotherapy in patients who have a lower risk of recurrence, who would therefore obtain limited benefit, avoids the unpleasant side effects of chemotherapy and reduces expenditure on both the chemotherapy itself and the treatment of AEs resulting from its use.

### *2.3.2 Summary of tumour profiling tests included in the assessment*

*Oncotype DX (Exact Sciences)*

Oncotype DX Breast Recurrence Score (Oncotype DX) is a Conformité Européene (CE) marked assay designed to quantify the 9-year risk of distant recurrence. The company claims that the test can also predict the likelihood of chemotherapy benefit. The test also reports the underlying tumour biology: ER, PR and HER2 status. The test is intended for use in people with early breast cancer that has the following clinical features:

- HR+
- HER2-
- LN0 or LN+ (up to 3 positive nodes).

Oncotype DX quantifies the expression of 21 genes. Of these, 16 are cancer-related genes correlated with distant recurrence-free survival (DRFS), and 5 are reference genes for normalising the expression of the cancer-related genes. This information is used to calculate the Breast Recurrence Score (RS).

Oncotype DX is offered as a test service to the NHS. Samples are processed centrally at the Exact Sciences centralised laboratory in the US, which is accredited by the American Association for Laboratory Accreditation (A2LA) and the College of American Pathologists (CAP). The test requires a formalin-fixed paraffin-embedded (FFPE) breast cancer tissue sample from a biopsy or surgical resection, which can be sent as a paraffin embedded block or as 15 unstained charged slides. The test process uses reverse transcription-quantitative polymerase chain reaction (RT-qPCR).

The test gives a recurrence score of between 0 and 100, which is used to estimate the 9-year risk of distant recurrence, assuming 5 years of standard ET. The company claims that the recurrence score also predicts the benefit of chemotherapy in terms of reducing the risk of distant recurrence. For LN+ disease (1 to 3 positive nodes), the Instructions For Use document state that a score below 18 predicts little to no chemotherapy benefit, a score between 18 and 30 predicts a potential chemotherapy benefit, and a score of 31 or more predicts a large benefit from chemotherapy. However, the company's website (accessed by NICE on the 27th February 2023), states that a recurrence score of 25 or less predicts no chemotherapy benefit for post-menopausal women and 2.9% benefit at 5 years for pre-menopausal

women. The company's website states that in both groups, a score of 26 to 100 is inferred to predict substantial chemotherapy benefit.

The Oncotype DX Breast Recurrence Score results are typically reported within 7 to 10 calendar days after the sample is received at the laboratory.

*Prosigna (Veracyte)*

Prosigna is a CE marked assay designed to provide information on breast cancer subtype and to predict DRFS at 10-years. The test is designed for use in post-menopausal women with early-stage breast cancer that is:

- HR+
- HER2- or HER2+
- LN0 or LN+ (up to 3 positive nodes, or 4 or more positive nodes).

Prosigna measures the expression of 50 genes used for intrinsic subtype classification, 8 housekeeping genes used for signal normalisation, 6 positive controls, and 8 negative controls. The test uses ribonucleic acid (RNA) extracted from an FFPE breast tumour tissue sample, and can be performed in local laboratories provided they have access to the nCounter Dx Analysis System. The company states that results are usually available within 3 days.

Prosigna classifies the risk of distant recurrence within 10 years, assuming 5 years of ET, based on the Prediction Analysis of Microarray 50 (PAM50) gene signature, breast cancer subtype, tumour size, nodal status and proliferation score. The proliferation score is determined by evaluating multiple genes associated with the proliferation pathway. The test gives an overall risk of recurrence (ROR) score between 0 and 100. Based on this score and the nodal status, samples are classified into risk categories. For LN+ disease (up to 3 positive nodes), a score of 0 to 15 indicates low risk, 16 to 40 indicates intermediate risk, and 41 to 100 indicates high risk. For 4 or more positive nodes, any score is assigned high risk. The EAG understands that most people with 4 or more positive nodes would be offered chemotherapy under current practice.

*EndoPredict (Myriad)*

EndoPredict is a CE marked assay that is designed to predict the likelihood of distant recurrence within 10 years of an initial diagnosis of breast cancer. The company claims that EndoPredict can also predict the absolute benefit of chemotherapy. The test is intended for use in pre- and post-menopausal people with early-stage breast cancer with all of the following clinical features:

- ER+
- HER2-
- LN0 or LN+ (up to 3 positive nodes).

EndoPredict measures the expression of 12 genes: 3 proliferation associated genes, 5 hormone receptor associated genes, 3 reference (normalisation) genes and 1 control gene. This information is used to calculate a 12-gene molecular score (or EP score).

EndoPredict requires RNA samples extracted from FFPE breast cancer tissue. The test can be performed in a local laboratory. It takes approximately 3 to 5 days to receive the test results after the sample has arrived at the laboratory.

The test process uses RT-qPCR. Online evaluation software (EndoPredict Report Generator) performs a quality check and calculates the EPclin score which is the final test result. The EPclin score is calculated by adding clinical data about tumour size and nodal status to the EP score. This can be used to estimate the likelihood of distant recurrence, assuming 5 years of ET. An EPclin score of less than 3.3 indicates low risk (less than 10%) of distant recurrence in the next 10 years. An EPclin score of 3.3 or more indicates high risk of distant recurrence in the next 10 years. The EPclin score can also be used to estimate absolute chemotherapy benefit; the company claims that people with an EPclin score of less than 3.3 are less likely to benefit from adjuvant chemotherapy.

*MammaPrint (Agendia)*

MammaPrint is a CE marked microarray that is designed to assess the risk of distant recurrence within 10 years. The company claims that the test also predicts whether a person would benefit from chemotherapy. The test is intended for use in pre- and post-menopausal women with stage I, II or operable stage III breast cancer with the following clinical features:

- HR+
- HER2-
- Tumour size up to 5cm
- LN0 or LN+ (up to 3 positive nodes).

MammaPrint measures the expression of 70 cancer-related genes, and 465 control genes.

The MammaPrint test is offered as an off-site service. In the UK, samples are sent for analysis at the Agendia laboratory in the US. A decentralised version of the test is also available for local laboratories with next-generation sequencing (NGS) capability. The test requires an FFPE breast cancer tissue sample. The company states that test results are typically reported within 10 days of receiving the sample at the laboratory and the average turnaround time is less than 5 days.

The test is based on diagnostic microarray. Software is used to calculate the MammaPrint result on a scale of −1 to +1. The score indicates the risk of developing distant metastases over the next 10 years without any adjuvant ET or chemotherapy. A MammaPrint result of 0 or less indicates high risk of

metastases in the next 10 years and a result of more than 0 indicates low risk (10% or less) of metastases in the next 10 years. A score of more than 0.355 can also be used to indicate ultra-low risk, which the company defines as more than 99% breast cancer-specific survival (BCSS) at 8 years and 97% BCSS at 20 years with 2 to 5 years of tamoxifen treatment.

**Table 4: Summary of tumour profiling tests**

| Test | Oncotype DX Breast Recurrence Score | Prosigna | EndoPredict EPclin score | MammaPrint |
|---|---|---|---|---|
| **Manufacturer** | Exact Sciences | Veracyte | Myriad | Agendia |
| **Purpose** | Recurrence risk and chemotherapy benefit | Intrinsic subtype and recurrence risk | Distant recurrence risk and chemotherapy benefit | Distant recurrence risk and chemotherapy benefit |
| **Description** | 21 gene assay (16 cancer genes; RT-qPCR) | 50 gene assay (50 cancer genes; direct mRNA counting) + clinical factors | 12 gene assay (8 cancer genes; RT-qPCR) + clinical factors | 70 gene assay (microarray) |
| **Testing location** | Test service (USA) | Local laboratory | Local laboratory | Local laboratory (NGS) or test service (USA) |
| **Stage** | Early-stage (Stage I to IIIa) | Early-stage (Stage I to IIIA) | Early-stage | Early-stage (Stage I, II or operable Stage III) |
| **Lymph node status** | LN0 or LN+ (up to 3 positive nodes) | LN0 and LN+ (up to 3 positive nodes, and 4+ nodes) | LN0 and LN+ (up to 3 positive nodes) | LN0 or LN+ (up to 3 positive nodes) |
| **Hormone receptor status** | HR+ | HR+ | ER+ | HR+ |
| **HER2 status** | HER2- | HER2- or HER2+ | HER2- | HER2- |
| **Menopausal status** | Pre- and post-menopausal | Post-menopausal only | Pre- and post-menopausal | Pre- and post-menopausal |
| **Test result** | Recurrence score<br>Chemotherapy benefit<br>Probability of distant recurrence (%) | Risk category (low, intermediate, high)<br>Intrinsic subtype<br>Probability of distant recurrence (%) | Risk category (low, high)<br>Chemotherapy benefit (%)<br>Probability of distant recurrence (%) | Risk category (low, ultra-low, high)<br>Chemotherapy benefit |
| **Assumptions** | Score assumes 5 years of endocrine treatment | Score assumes 5 years of endocrine treatment | Scores assume 5 years of endocrine treatment | Assumes no adjuvant therapy |

*ER – oestrogen receptor; HER2 – human epidermal growth factor; HR – hormone receptor; LN – lymph node; RT-qPCR – reverse transcription-quantitative polymerase chain reaction; NGS – next generation sequencing; mRNA – messenger ribonucleic acid; USA – United States of America*

### *2.3.3 Current usage of tumour profiling tests in the NHS*

NICE DG34 recommended the use of EndoPredict (EPclin score), Oncotype DX and Prosigna as options for guiding adjuvant chemotherapy decisions for people with ER+, HER2-, LN0 early breast cancer, including those with micrometastases, assessed to be at intermediate risk of recurrence of breast cancer after surgery.[13] Two tests – MammaPrint and IHC4 – were not recommended. DG34 did not make any specific recommendations on the use of any of these tumour profiling tests in people with LN+ early breast cancer. The current use of tumour profiling tests in guiding adjuvant chemotherapy decisions in women with LN+ early breast cancer in the NHS is limited: Oncotype DX is available in some UK centres in the private sector (if patients or insurers fund it), although this test is available in some NHS centres through early or compassionate access schemes or may be funded by local Trusts.

## 2.4 Description of decision problem

This assessment aims to evaluate whether tumour profiling tests used for guiding adjuvant chemotherapy decisions for people with ER+ (and/or PR+), HER2-, early-stage breast cancer with 1 to 3 positive lymph nodes (LN1-3) represent a clinically effective and cost-effective use of NHS resources.

This assessment represents an update to the systematic review and cost-effectiveness analysis (Harnan *et al.*[10]) which informed considerations for the LN+ subgroup within NICE DG34.[13]

### *2.4.1 Interventions*

The following tumour profiling tests are included in combination with current decision-making:

- EndoPredict (EPclin)
- MammaPrint
- Oncotype DX Breast Recurrence Score (RS)
- Prosigna (or ROR-PT, which is equivalent).

### *2.4.2 Comparators*

The comparator for this appraisal is current decision-making, which may include any tool, or clinical and pathological features, used to assess risk. Clinicopathological tools used in current practice include PREDICT and the NPI.

### *2.4.3 Population and important sub-groups*

The population of interest for this assessment relates to people with ER+ (and/or PR+), HER2-, early-stage breast cancer with 1 to 3 positive lymph nodes (LN1-3) who are deciding whether to have adjuvant chemotherapy.

The focus of this assessment is on patients with Stage I-IIIA disease.[20]

*Subgroups*

Where evidence allows, the following subgroups are considered:

- Pre-menopausal women and post-menopausal women
- People predicted to be in low-, intermediate- or high-risk groups using a risk assessment tool (such as PREDICT or NPI), or using clinical and pathological features
- Sex
- People of different ethnicities
- People with comorbidities which mean that they could be particularly affected by the side effects of chemotherapy.

### *3.4.4 Outcomes*

Relevant outcomes include the following:

*Intermediate measures:*

- Prognostic ability
- Ability to predict relative benefit from chemotherapy
- Impact of test results on decision-making.

*Clinical outcomes:*

- Distant recurrence-free survival (DRFS), distant recurrence-free interval (DRFI), distant metastasis-free survival (DMFS) and distant metastasis-free interval (DMFI)
- Disease-free survival (DFS) and breast cancer-specific survival (BCSS)
- Overall survival (OS)
- Disease-related morbidity and mortality
- Chemotherapy-related morbidity and mortality.

*Patient-reported outcomes:*

- HRQoL
- Anxiety.

Costs are considered from an NHS and Personal Social Services (PSS) perspective. The cost-effectiveness of interventions is expressed in terms of incremental cost per quality-adjusted life year (QALY) gained. Costs for consideration include:

- Costs of treating breast cancer, including: drug costs, administration costs, outpatient appointments, supportive care costs and costs associated with treating AEs
- Costs of the tests, including equipment costs and reagents, where applicable
- Costs of staff and associated training, where applicable.

### 2.5 Aims and objectives of the assessment

The main research question to be addressed is: *"Do tumour profiling tests used for guiding adjuvant chemotherapy decisions in patients with ER-positive (and/or PR positive), HER2-negative, early-stage breast cancer with 1 to 3 positive lymph nodes represent a clinically effective and cost-effective use of NHS resources?"*

The objectives of the assessment are as follows:

- To conduct a systematic review of the published evidence on the effectiveness and cost-effectiveness of the four tumour profiling tests (Oncotype DX, Prosigna, EPclin and MammaPrint).
- To develop a health economic model to assess the cost-effectiveness of these tumour profiling tests compared with current prognostic tools to guide the use of adjuvant chemotherapy in early breast cancer from the perspective of the NHS and PSS.

# 3 CLINICAL EFFECTIVENESS

This chapter presents the methods and results of a systematic review of clinical evidence for the effectiveness of tumour profiling tests to guide treatment decisions in people with ER+, HER2-, LN+ early breast cancer.

## 3.1 Methods for clinical review

### *3.1.1 Overview of systematic review methodology*

A systematic review was undertaken to update the previous systematic review (Harnan *et al.*, 2019[10]) conducted for the LN+ subgroup within NICE DG34.[13]

A protocol of this systematic review (CRD42023425638) is available on the PROSPERO website at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=425638 (accessed 21st August 2023). The review was conducted following the general principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[21]

### *3.1.2 Inclusion criteria*

#### 3.1.2.1 Population and subgroups

The relevant population is people with ER-positive (and/or PR-positive), HER2-negative, early stage breast cancer (Stage I, II or IIIA) with 1 to 3 positive lymph nodes (excluding those patients with micrometastases, who were included in the recommendations for LN0 patients in DG34[13]). In general, where studies included patients who were out of scope, if ≤20% were out of scope then the study was included (and heterogeneity was considered), whilst if >20% were out of scope then the study was excluded. Exceptions to this were that some studies did not report HER2 status, whilst some studies included LN+ patients but >20% had >3 positive nodes; these studies were included to ensure inclusion of sufficient relevant evidence, but these limitations were noted. Data for subgroups listed in Section 2.4 were included, where available.

#### 3.1.2.2 Interventions

The following interventions were included: Oncotype DX Breast Recurrence Score (RS); MammaPrint; Prosigna and EndoPredict (EPclin score). Only studies using the commercial versions of the tests were included. The review excluded *in silico* studies which use algorithms for the genes within a test and apply these to electronic (*in silico*) databases of genetic profiles generated from microarray techniques. Although the PAM50 score is a part of Prosigna, PAM50 intrinsic subtypes were not included, only the Prosigna ROR score. Magee equations (which approximate the Oncotype DX score) were not included.

#### 3.1.2.3 Comparators

The relevant comparator is current decision-making, which includes clinical and pathological features used to assess risk, and clinicopathological tools outlined in Section 2.2 (current tools include PREDICT and NPI whilst older tools include AOL). Due to the lack of availability of end-to-end studies comparing decision-making based on the test versus current tools, different evidence types were sought and are linked via the EAG's health economic model (see Section 4.3).

#### 3.1.2.4 Outcomes

The clinical review aimed to identify the following types of data:

- End-to-end studies comparing the tests versus current decision-making (if available)
- Prognostic ability
- Ability to predict benefit from chemotherapy
- Impact of test results on chemotherapy decisions (restricted to studies conducted in the UK or Europe, due to differing rates of chemotherapy use worldwide)
- HRQoL and anxiety associated with use of the tests.

The data on prognostic and predictive ability included the following clinical outcomes:

- Distant recurrence-free survival (DRFS), distant recurrence-free interval (DRFI), distant metastasis-free survival (DMFS) and distant metastasis-free interval (DMFI)
- Disease-free survival (DFS)
- Overall survival (OS) and breast cancer-specific survival (BCSS)

Studies only reporting local recurrence (LR) or LRR were excluded.

The different study types are linked via the EAG's health economic model in order to inform the clinical effectiveness and cost-effectiveness of the tests (see Section 4.3).

#### 3.1.2.5 Date and language limits

As noted above, this review updates a previous systematic review (Harnan *et al.*[10]). Relevant studies from all dates were included. Studies published prior to 2017 were identified and extracted from the previous review (search date February 2017), whilst studies published from 2017 onwards were identified via an update search. Studies not published in English would have been includable if sufficient data could be extracted; however, no relevant studies published in non-English languages were identified.

### *3.1.3 Search strategy*

The search strategy for the systematic review comprised the following main elements: searching of electronic databases, registers and websites; contact with experts in the field; review of bibliographies of retrieved papers and existing systematic reviews; review of Request For Information (RFI) documents and manufacturer submissions to NICE.[22-25] The databases, trial registers and websites searched included the following:

- MEDLINE and MEDLINE in Process (via Ovid)
- EMBASE (via Ovid)
- Cochrane Database of Systematic Reviews (via Wiley)
- Cochrane Central Register of Controlled Trials (via Wiley)
- HTA Database of the International Network of Agencies for Health Technology Assessment (INAHTA)
- Web of Science Citation Index Expanded (via Clarivate)
- Web of Science Conference Proceedings Citation Index (via Clarivate)
- World Health Organization International Clinical Trials Registry Platform (WHO ICTRP)
- Clinicaltrials.gov (National Library of Medicine)
- American Society of Clinical Oncology (ASCO)
- European Society for Medical Oncology (ESMO)
- American Association for Cancer Research (AACR)
- European Cancer Organization (ECO).

Search terms included free-text test names (EndoPredict, MammaPrint, Oncotype and Prosigna) and their related synonyms, combined with terms for breast cancer. The MEDLINE search strategy is included in Appendix 1. The searches were limited by date from 2017 to present, as the searches for the previous review[10] were conducted in February 2017.

### *3.1.4 Study selection and data extraction strategy*

Titles and abstracts of retrieved records were assessed for relevance. Early in the process, a 10% sample of records were checked between reviewers and any discrepancies were discussed to inform the remaining study selection process. The full texts of remaining records were obtained and assessed against the inclusion criteria (see Section 3.1.2). Any studies causing uncertainty were checked by a second reviewer with involvement of a third reviewer when necessary. Data were extracted into Microsoft Excel® by one reviewer and checked by a second reviewer. Studies published prior to 2017 were extracted from the existing review.[10]

### *3.1.5 Quality assessment strategy*

Studies were assessed using quality assessment tools relevant to the study design. Prospective RCTs were assessed using the Cochrane Risk of Bias tool Version 2 (RoB2).[26] Prognostic and prediction studies were assessed using the Prediction model study Risk Of Bias Assessment Tool (PROBAST);[27] items from each domain were selected based on their relevance to this review, and definitions of high or low risk for each item specific to this review were defined *a priori* (see Appendix 3). Each study, cohort or registry was assessed once, rather than assessing each publication separately. Decision impact studies did not undergo formal quality assessment, but the design and relevance of these studies were considered narratively. The impact of the quality of studies on the evidence base was considered within the narrative synthesis.

### *3.1.6 Methods of analysis/synthesis*

Results of the review were analysed and presented via a narrative synthesis and tabulation.

## 3.2 Results of clinical review: overview

### *3.2.1 Quantity and type of included studies*

The database search for the clinical review identified 4,057 articles, of which 501 were checked as full texts, and 41 were includable. In addition, 13 further articles were included from the previous review by Harnan *et al*.[10] Therefore, in total, 54 articles were included in the clinical review. Of these, 42 articles related to patient outcomes (prognostic, predictive and prospective use of test), whilst 12 related to decision impact studies. No studies were identified which assessed HRQoL or anxiety associated with use of tumour profiling tests in a LN+ population; therefore, a short summary of such studies in a LN0 or mixed population was provided (these were not counted as included studies for the purposes of the PRISMA flow chart; see Appendix 2).

Section 3.2.2 provides an overview of the identified evidence for each data type, together with a description of some of the key studies informing the clinical evidence and the cost-effectiveness analysis. The remainder of the clinical chapter presents the data on risk of bias in included studies (Section 3.3), prognostic ability of the tests (Section 3.4), prediction of chemotherapy benefit (Section 3.5), decision impact (Section 3.6), and HRQoL and anxiety (Section 3.7).

### *3.2.2 Summary of evidence identified for each outcome type*

*Prognostic ability: Summary of evidence*

The prognostic ability of a genomic test describes its ability to differentiate between patients with good versus poor outcomes. The evidence on prognostic ability in this review includes the following types of evidence and key studies:

- Prospective RCTs reporting recurrence/survival outcomes for patients within a particular test

risk group (or range). Two prospective RCTs reported data (RxPONDER[28] for Oncotype DX and MINDACT[29] for MammaPrint); these are described below;

- Re-analyses of clinical trials or cohorts with long-term follow-up, where the tests are used on stored tumour samples, and recurrence/survival outcomes are compared between risk groups;
- Observational studies of the use of the test in practice and recurrence/survival data by risk group. These studies have the limitation that test results may have influenced chemotherapy use.

*Prediction of chemotherapy benefit: Summary of evidence*

Whether a test is predictive for chemotherapy benefit is determined by whether the effect of chemotherapy (i.e., the hazard ratio [HR] for chemotherapy vs. no chemotherapy for recurrence/survival) differs between test risk groups or ranges. This is generally assessed via a statistical test for interaction.[30] The main study designs for this evidence are:

- Prospective RCTs which randomise patients within a particular test risk group (or range) to chemotherapy versus no chemotherapy. These studies can only provide data within that risk group/range. Two prospective RCTs reported data (RxPONDER[28] for Oncotype DX and MINDACT[29] for MammaPrint) and are described below.
- Re-analyses of studies of chemotherapy versus no chemotherapy, with long-term follow-up, where the tests are used on stored tumour samples, and HRs for chemotherapy versus no chemotherapy for recurrence/survival outcomes can be calculated per test risk group. Such data were available for Oncotype DX from a reanalysis of the SWOG-8814 RCT (Albain *et al.*, 2010)[31] and for MammaPrint from a reanalysis of two cohorts (Mook *et al.*, 2009).[32]
- Observational studies of the use of the test in practice and recurrence/survival data for chemotherapy versus no chemotherapy within each test risk group. These studies have the limitation that patients are not randomised to chemotherapy and so there may be confounding.

*Prospective RCT of Oncotype DX: RxPONDER*

The RxPONDER[28] study of Oncotype DX randomised patients with Oncotype DX RS ≤25 to chemo-endocrine therapy (CET) vs. endocrine monotherapy. Some prognostic data were reported, assessing whether RS as a continuous score (within the range RS 0-25) was related to patient outcomes (invasive disease-free survival [IDFS] only). The study also provided data on the effect of chemotherapy vs. no chemotherapy and whether RS was predictive of chemotherapy benefit. This consisted of data on outcomes (IDFS, DRFS and DRFI) for patients with and without chemotherapy, for the full study population (RS 0-25), as well as for narrower RS ranges, and for patient subgroups such as pre- and post-menopausal patients. In terms of prediction of chemotherapy benefit, a test for interaction was reported between RS (within the range 0-25) and effect of chemotherapy on IDFS (no interaction test

was reported for distant recurrence outcomes). A limitation of this study was that it could not provide prognostic or prediction data for patients with an RS outside the study range i.e., for patients with an RS of 26-100. In addition, the majority of patients in RxPONDER had only 1 positive node (65% had 1 positive node, 25% had 2 positive nodes, 9% had 3 positive nodes). Furthermore, patients had knowledge of their RS result before agreeing to be randomised, which may have resulted in selection bias (of 9,383 women screened, 4,300 were excluded before randomisation, of which 1,035 had RS >25 but the remaining 3,265 did not participate for other reasons). Results of RxPONDER are described in this report in Sections 3.4 and 3.5. RxPONDER also informs the EAG's economic analyses of Oncotype DX (see Section 4.3).

*Prospective RCT of MammaPrint: MINDACT*

The MINDACT[29] study of MammaPrint assessed patients' genomic risk (via MammaPrint) and clinical risk (via modified Adjuvant! Online [mAOL]). Patients who were low-risk on both MammaPrint and mAOL were allocated to no chemotherapy, those who were high-risk on both were allocated to chemotherapy, and patients with discordant risk were randomised to chemotherapy vs. no chemotherapy. Outcomes (DMFS, DMFI, DFS, OS) are presented for patients in the four subgroups according to high/low clinical risk and high/low MammaPrint risk. There are limitations in using MINDACT to assess prognostic ability because, due to the study design, MammaPrint results influenced chemotherapy use (more patients in the MammaPrint high-risk group received chemotherapy compared with the MammaPrint low-risk group), and no HRs or significance tests were reported for the difference in outcomes between test risk groups. The study also provided data on the effect of chemotherapy vs. no chemotherapy on patient outcomes (DMFS, DMFI, DFS, OS). Results for chemotherapy vs. no chemotherapy were presented for the clinical high, MammaPrint low group; however, data were not analysed for the clinical low, MammaPrint high group due to small numbers of LN+ patients. The study therefore provided data on chemotherapy benefit only for patients with clinical high, MammaPrint low risk. However, since all patients in the clinical high-risk, MammaPrint high-risk group were offered chemotherapy, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit. MINDACT informs the EAG's economic analysis of MammaPrint (see Section 4.3).

*Ongoing prospective RCT of Prosigna: OPTIMA*

The ongoing OPTIMA study[33] is an RCT of test-directed chemotherapy use vs. standard chemotherapy use. Included patients have high clinical risk of recurrence and are largely node-positive (1-9 positive nodes). Patients randomised to test-directed treatment receive a Prosigna test, then receive CET if high-risk on Prosigna, and ET alone if low-risk on Prosigna, whilst the standard care arm all receive CET. Pre-menopausal patients receive ovarian function suppression, to control for chemotherapy-induced

menopause. OPTIMA uses a non-inferiority design to assess IDFS, DRFI, BCSS, OS. This study is still in the recruitment phase and the review did not identify any published results of OPTIMA so far.

*Impact of test results on chemotherapy decisions: summary of evidence*

Evidence on pre-test and post-test decisions/recommendations for receiving chemotherapy was identified for Oncotype DX; this included five UK studies[34-38] and seven other European studies.[39-45] Of these, two UK studies and four other European studies reported data by Oncotype DX risk groups. No decision impact studies were identified which assessed EPclin, Prosigna or MammaPrint.

*HRQoL and anxiety associated with use of the tests: Summary of evidence*

No studies (or subgroups) reporting HRQoL or anxiety associated with use of tumour profiling tests were identified in a mainly LN+ population. Therefore, a brief summary of the evidence in a LN0 or mixed population is provided.

### 3.3 Risk of bias in included studies

A summary of risk of bias in the included studies is provided here, with further details in Appendix 3.

The two prospective RCTs (RxPONDER[28] and MINDACT[29]), assessed using the Cochrane RoB2 tool,[26] scored low risk of bias on all domains, and low risk of bias overall. As noted in Section 3.2.2, there may have been selection bias in RxPONDER since patients had knowledge of their RS result before agreeing to be randomised.

Risk of bias in prognostic and predictive studies was assessed using the PROBAST tool.[27] For prognostic studies, the following factors may have affected results to some extent. Studies varied in terms of whether people received chemotherapy or not; studies are therefore reported separately according to chemotherapy use in the section on prognostic ability (Section 3.4). In some studies, some participants did not match the review question (either not ER+, not HER2- or not LN1-3); these factors were taken into account when selecting studies for use in the economic model. Most studies excluded a proportion of patients for various reasons including insufficient tissue, missing data, failed tests and others, which may have influenced results to some extent, though the impact is difficult to assess. In terms of outcomes, chemotherapy decisions were not influenced by the test result in studies of retrospective use of the test (i.e., reanalyses of RCTs and cohorts), whereas in observational studies in which the test was used prospectively, chemotherapy decisions may have been influenced by the test result; therefore, observational studies are reported separately in the section on prognostic ability (Section 3.4).

For predictive studies, the following factors may have affected results to some extent. Only the SWOG-8814 study[31] was a reanalysis of an RCT in which chemotherapy use was randomised; in the remaining studies, chemotherapy use was not randomised. This limitation is reflected in the section on prediction of chemotherapy benefit (Section 3.5). In some studies, some participants did not match the review question (either not ER+, not HER2- or not LN1-3). Most studies excluded a proportion of patients for various reasons including insufficient tissue, missing data, failed tests and others, which may have influenced results to some extent, though the impact is difficult to assess. In terms of outcomes, chemotherapy decisions were not influenced by the test result in two studies of retrospective use of the test, whereas in the three observational registries in which the test was used prospectively, chemotherapy decisions may have been influenced by the test result; therefore, observational studies are reported separately in the section on prediction of chemotherapy benefit (Section 3.5).

## 3.4 Results: Prognostic ability

### *3.4.1 Overview of prognostic data in this report*

The prognostic ability of a genomic test describes its ability to differentiate between patients with good versus poor outcomes. Studies of prognostic ability provide risk classification probabilities i.e., the proportion of patients allocated to each risk group. They also provide the risk of distant metastases (DM) or OS per risk group, and HRs for the difference in outcomes between risk groups (both unadjusted and after adjustment for clinical and pathological factors). The evidence on prognostic ability in this review includes the following types of evidence and key studies:

- Prospective RCTs reporting recurrence/survival outcomes for patients within a particular test risk group (or range). These studies can only provide data within that risk group/range. Two prospective RCTs reported data (RxPONDER[28] for Oncotype DX and MINDACT[29] for MammaPrint);
- Re-analyses of clinical trials or cohorts with long-term follow-up, where the tests are used on stored tumour samples, and recurrence/survival outcomes are compared between risk groups. Such studies were identified for all four tests and are the main source of data on prognostic ability for distant recurrence (a reanalysis of the TransATAC study[19] provided data for three of the four tests). In total, 23 publications relating to 18 studies provided data on prognostic ability (some reported on more than one test): 5 studies of Oncotype DX,[19, 28, 31, 46-49] 5 studies of MammaPrint,[29, 32, 50-53] 6 studies of Prosigna[19, 46, 54-60] and 5 studies of EPclin.[19, 46, 57, 58, 61-63]
- Observational studies of the use of the test in practice and recurrence/survival data by risk group. These studies have the limitation that test results may have influenced chemotherapy use. These studies were identified for Oncotype DX only, and include the Clalit registry[64] in Israel, the Surveillance Epidemiology and End Results (SEER) registry[65-67] in the US, the National Cancer Database (NCDB)[68-72] in the US and a few smaller prospective studies.[73-75] These analyses provide real-world outcomes data for patients in different test risk groups, but have the limitation that test results likely influenced chemotherapy use.

A summary of prognostic data for distant recurrence across the four tests, based on reanalyses of trials or cohorts, is provided in Table 5. Full details of the prognostic data are provided in Appendix 4, which includes additional outcomes (such as DFS, OS and BCSS). Data for the two prospective RCTs (RxPONDER[28, 49] and MINDACT[29, 50]) are presented in Table 6 and Table 7. Data on observational studies of prospective use of Oncotype DX are provided in Table 8.

### *3.4.2 Summary of distribution of genomic risk groups and distant recurrence risk*

Table 5 summarises prognostic data from studies reporting 10-year distant recurrence outcomes. For patients receiving endocrine monotherapy, the review identified 1 study of Oncotype DX,[19] no studies of MammaPrint, 2 studies of Prosigna[19, 54] and 3 studies of EPclin.[19, 61, 63] A third study of Prosigna[56] which used different cut-offs is shown in Table 5 for completeness, but is not included in this textual summary. In terms of distribution, the study of Oncotype DX[19] (which used cut-offs of RS <18 and RS 30) assigned more patients to the low-risk group (57%) than the studies of Prosigna (4-8% low-risk)[19, 54] or EPclin (19-35% low-risk).[19, 61, 63] Freedom from distant recurrence at 10 years in the low-risk group was 81% (1 study of Oncotype DX),[19] 100% (2 studies of Prosigna)[19, 54] and 94-100% (3 studies of EPclin).[19, 61, 63] Freedom from distant recurrence at 10 years in the high-risk group was 62% (1 study of Oncotype DX),[19] 69-76% (2 studies of Prosigna)[19, 54] and 70-81% (3 studies of EPclin).[19, 61, 63] Further details of the prognostic data can be found in Appendix 4.

Table 5 also presents 10-year distance recurrence data from further studies in which some or all patients received chemotherapy, including 1 study of Oncotype DX,[47] 3 studies of MammaPrint,[32, 51, 52] 1 study of Prosigna[57, 58] and 1 study of EPclin.[57, 58] The distributions and 10-year distant recurrence data in these studies follow a similar pattern to the studies of ET monotherapy. MammaPrint, for which there were no studies of ET monotherapy, assigned 38-48% of patients to the low-risk group, whilst freedom from distant recurrence at 10 years ranged from 79-95% in the low-risk group and 54-81% in the high-risk group.[32, 51, 52] Further details of the prognostic data can be found in Appendix 4.

### *3.4.3 Summary of prognostic ability across tests*

Table 5 (last two columns) also provides a summary of whether tests were significantly prognostic for 10-year distant recurrence. This is generally based on an HR for distant recurrence between risk groups or an HR per unit change in test score; full details of HRs are included in Appendix 4. Prognostic significance is summarised for unadjusted analyses, as well as for adjusted analyses which indicate whether tests remain prognostic after adjustment for clinical factors. For all four tests, the HR for prognostic ability was statistically significant for most, though not all, analyses.

**Table 5: Summary of prognostic data for 10-year distant recurrence (all four tests)**

| Test | ET/CT | Reference Study | Design | N pts | Outcome | Nodal status | HR, HER2 | Meno status | Test cut-offs | Distribution % | | | DR free 0-10yr % | | | [a]Sig prog 10y? | [a]Sig prog 10y adj? |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Low | Int | High | Low | Int | High | | |
| **Oncotype DX** | ET alone | Sestak 2018[19] (TransATAC) | RCT-R | N=183 | DRFI | LN1-3 | HR+ HER2- | Post | 18, 30 | 57 | 32 | 11 | 81 | 71 | 62 | N | N |
| | All CT+ET | Mamounas 2018[47] (NSABP-28) | RCT-R | N=722 | DRFI | LN1-3 | ER+ NR HER2 | Pre/post | 18, 30 | 37 | 34 | 28 | 85 | 72 | 63 | Y | Y |
| **Mamma Print** | Variable ET/CT | Drukker 2014[51] | Cohort-R | N=144 | DMFS | 74% LN1-3 26% LN4+ | 77% ER+ NR HER2 | Pre/post (age <53) | 0.4 | 38 | - | 62 | 79 | - | 54 | Y | - |
| | | Mook 2009[32] | Cohort-R | N=241 | DMFS | LN1-3 + Lnmicro | 79% ER+ 84% HER2- | Pre/post | NR | 41 | - | 59 | 91 | - | 76 | Y | N |
| | | Vliek 2017[52] (RASTER) | Cohort-R | N=134 | DRFI | LN1-3 | 83% ER+ 85% HER | Pre/post | NR | 48 | - | 52 | 95 | - | 81 | Y | - |
| **Prosigna** | ET alone | Sestak 2018[19] (TransATAC) | RCT-R | N=183 | DRFI | LN1-3 | HR+ HER2- | Post | 16, 40 | 8 | 32 | 60 | 100 | 79 | 69 | N | Y |
| | | Gnant 2014[54] / Filipits 2014[55] (ABCSG-8) | RCT-R | N=413 | DMFS | 89% LN1-3 11% LN4+ | ER+ HER2- | Post | 16, 40 | 4 | 34 | 62 | 100 | 94 | 76 | - | Y |
| | | Laenkholm 2018[56] (DBCG) | Cohort-R | N=1,395 | DRFS | LN1-3 | HR+ HER2- | Post | Varies by N nodes | 26 | 28 | 46 | 97 | 89 | 78 | Y | Y |
| | All CT+ET | Martin 2016/14[57, 58] (GEICAM 9906) | RCT-R | N=536 | DMFS | 64% LN1-3 36% LN4+ | ER+ HER2- | 54% pre 46% post | 18, 65 | 19 | 56 | 26 | 92 | 74 | 66 | Y | N |
| **EPclin** | ET alone | Sestak 2018[19] (TransATAC) | RCT-R | N=183 | DRFI | LN1-3 | HR+ HER2- | Post | 3.3 | 23 | - | 77 | 94 | - | 70 | Y | Y |
| | | Filipits 2019[61] (ABCSG-6/8) | RCT-R | N=453 | DRFR | LN1-3 | ER+ HER2- | Post | 3.3 | 35 | - | 65 | 96 | - | 81 | Y | Y |
| | | Constantinidou 2022[63] | Cohort-R | N=62 | DRFS | LN1-3 | ER+ HER2- | Pre | 3.3 | 19 | - | 81 | 100 | - | 75 | N | Y |
| | All CT+ET | Martin 2016/14[57, 58] (GEICAM 9906) | RCT-R | N=555 | DMFS | 64% LN1-3 36% LN4+ | ER+ HER2- | 54% pre 46% post | 3.3 | 13 | - | 87 | 100 | - | 72 | Y | Y |

*[a]The last two columns indicate how many studies report an HR between test risk groups which is statistically significant at the 5% level (unadjusted or adjusted for clinical factors).*
*Adj – adjusted; cohort-R – cohort reanalysis; CT – chemotherapy; DMFS – distant metastasis-free survival; DR, distant recurrence; DRFI – distant recurrence-free interval; DRFR – distant recurrence-free rate; DRFS – distant recurrence-free survival; ET – endocrine therapy; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR – hormone receptor; int – intermediate; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; prog – prognostic; RCT – randomised controlled trial; RCT-R – RCT reanalysis; sig – significant; var – variable; yr – year.*

### *3.4.4 Prognostic data from prospective RCT of Oncotype DX (RxPONDER)*

The prospective RCT of Oncotype DX (RxPONDER)[28] randomised patients with an Oncotype DX RS of ≤25 to chemotherapy plus ET vs. ET monotherapy. The publication mainly focusses on prediction of chemotherapy benefit; this is discussed in Section 3.5. RxPONDER also reports some prognostic data which are presented in Table 6. These data are not included in summary Table 5 because prognostic data were not reported for distant recurrence, only for IDFS. Prognostic ability in RxPONDER could only be analysed within the study population (those with an RS of 0-25) so there are no prognostic data covering patients with Oncotype DX RS 26-100. Within the range RS 0-25, Oncotype DX was significantly prognostic for 5-year IDFS after adjusting for clinical factors, both in the overall population (HR per unit-RS 1.05; 95% confidence interval [CI] 1.04 to 1.07; $p<0.001$) and in the pre-menopausal and post-menopausal subgroups (similar HRs to the overall population, see Table 6).[28] A further RxPONDER publication[49] reported IDFS results by ethnicity; 5-year IDFS within RS 0-25 was slightly worse in black patients (87.0%) and slightly better in Asian patients (93.9%) compared with white patients (91.5%), but overall rates were similar, and no data were reported for prognostic ability by ethnicity (see Table 6).

Distant recurrence data in RxPONDER are also shown in Table 6. Across all patients (all RS 0-25), the 5-year DRFI was 94-96%, both in pre-menopausal and post-menopausal groups, with or without chemotherapy. For comparison, in two RCT reanalyses, 5-year DRFI in the RS 0-17 group was 96% and 94%, while 5-year DRFI in the RS 18-30 group was 85% and 87% (TransATAC[19] and Penault-Llorca 2018,[48] Appendix 4).

**Table 6: Prognostic data from prospective RCT of Oncotype DX (RxPONDER)**

| Reference Study | Out-come | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Risk 0-5yr % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low RS≤25 | | High | | |
| | | | | | | CT | No | | | |
| **Oncotype: Prospective RCT: Distant recurrence** | | | | | | | | | | |
| Kalinsky 2021[28] RxPONDER | DRFI (0-5y) | CT v none Prosp RCT n=3,353 | LN1-3 (65% 1 node, 25% 2 nodes, 9% 3 nodes) 100% HR+ 100% HER2- | Post-meno | All ≤25 | 94.9 | 93.9 | - | - | - |
| | | n=1,665 | | Pre-meno | All ≤25 | 96.3 | 93.9 | - | - | - |
| **Oncotype: Prospective RCT: IDFS** | | | | | | | | | | |
| Kalinsky 2021[28] RxPONDER | IDFS (0-5y) | n=5,018 | LN1-3 100% HR+ 100% HER2- | All meno (67% post) | All ≤25 | 92.2 | 91.0 | - | *0-5yr: HR per unit-RS (within RS 0-25, adj for meno and CT): 1.05 (1.04 to 1.07), *p*<0.001 | Y* |
| | | | | Post-meno | All ≤25 | 91.2 | 91.9 | - | *0-5yr: HR per unit-RS (within RS 0-25, adj for CT, nodes, grade, tumour size, age): 1.05 (1.03 to 1.07), *p*<0.001 | Y* |
| | | | | Pre-meno | All ≤25 | 93.9 | 89.0 | - | *0-5yr: HR per unit-RS (within RS 0-25, adj for CT, nodes, grade, tumour size, age): 1.06 (1.02 to 1.09), *p*=0.001 | Y* |
| **Oncotype: Prospective RCT: IDFS by ethnicity** | | | | | | | | | | |
| Abdou 2023[49] RxPONDER | IDFS | n=4,015 CT+ET vs. ET Prosp RCT | LN1-3 100% HR+ 100% HER2- | White (n=2,833) | All ≤25 | 91.5 | | - | - | - |
| | | | | Black (n=248) | All ≤25 | 87.0 | | - | - | - |
| | | | | Asian (n=324) | All ≤25 | 93.9 | | - | - | - |
| | | | | Hispanic (n=610) | All ≤25 | 91.4 | | - | - | - |

*[a]The last column indicates whether each HR between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*

*Adj – adjusted; CI – confidence interval; CT – chemotherapy; ET – endocrine therapy; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR – hormone receptor; IDFS – invasive disease-free survival; int – intermediate; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; prosp – prospective; RCT – randomised controlled trial; RS – Recurrence Score (Oncotype DX); sig – significant; y/yr – year*

### *3.4.5 Prognostic data from prospective RCT of MammaPrint (MINDACT)*

The prospective RCT of MammaPrint (MINDACT)[29] assessed patients' genomic risk via MammaPrint and clinical risk via mAOL. Patients who were low-risk on both MammaPrint and mAOL were allocated to no chemotherapy, those who were high-risk on both were allocated to chemotherapy, and patients with discordant risk were randomised to chemotherapy vs. no chemotherapy. All the MINDACT data presented in Chapter 3 of this report refer to the LN+ subgroup, unless stated otherwise. In terms of distribution, within the clinical high-risk group, 69% were MammaPrint low-risk and 31% were MammaPrint high-risk, whilst within the clinical low-risk group, 92% were MammaPrint low-risk and 8% were MammaPrint high-risk (see Table 7).

Outcome data from the MINDACT LN+ subgroup are presented for patients in the different risk groups (Table 7).[29] However, it is difficult to compare outcomes for the MammaPrint low-risk and high-risk groups because, due to the study design, MammaPrint results influenced chemotherapy use (more patients in the MammaPrint high-risk group received chemotherapy compared with the MammaPrint low-risk group) which confounds the analysis of prognostic ability. Therefore, MINDACT data are not included in summary Table 5. Within clinical high-risk patients, outcomes were generally better for MammaPrint low-risk than MammaPrint high-risk groups, despite the fact that only 50% of low-risk patients but all high-risk patients were allocated chemotherapy. For example, 8-year DMFI was 92.3% for MammaPrint low-risk vs. 80.9% for MammaPrint high-risk, with other outcomes showing a similar pattern (see Table 7). However, no HRs or significance tests were reported for differences in outcomes between test risk groups (i.e., prognostic ability). Within clinical low-risk patients, 8-year DMFI was 95.2% for MammaPrint low-risk patients (allocated no chemotherapy), but the MammaPrint high-risk group was not analysed due to small numbers of LN+ patients (n=15). A further MINDACT publication[50] assesses an ultra-low-risk MammaPrint group, which incorporates 15% of the LN+ subgroup, with an 8-year DMFI of 95.2% (presumably across clinical low-risk and high-risk groups). The effect of chemotherapy vs. No chemotherapy within each group is discussed in Section 3.5.

**Table 7: Prognostic data from prospective RCT of MammaPrint (MINDACT)**

| Reference Study | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | Risk 0-8yr % | | HR between groups | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | High | Low | High | | |
| **MammaPrint: Prospective RCT: Distant recurrence** | | | | | | | | | | | |
| Piccart 2021[29c] MINDACT | DMFI | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | 31 | 92.3 (50% CT) | 80.9 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | 8 | 95.2 (no CT) | -[b] | - | - |
| | DMFS | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | 31 | 91.0 (50% CT) | 79.1 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | 8 | 94.0 (no CT) | -[b] | - | - |
| Lopes Cardozo 2022[50] MINDACT | DMFI | N=201 (ultra-low) Var ET/CT Prosp-RCT | LN1-3 99% ER+ 97% HER2- | - | >0.355 ultra-low | Ultra-low: 15 | - | Ultra-low: 95.2 | - | - | - |
| **MammaPrint: Prospective RCT: DFS** | | | | | | | | | | | |
| Piccart 2021[29c] MINDACT | DFS | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | 31 | 84.5 (50% CT) | 74.5 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | 8 | 85.6 (no CT) | - | - | - |
| **MammaPrint: Prospective RCT: OS** | | | | | | | | | | | |
| Piccart 2021[29c] MINDACT | OS | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | 31 | 95.1 (50% CT) | 89.1 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | 8 | 98.1 (no CT) | - | - | - |

*[a]The last column indicates whether each HR between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*
*[b]The mAOL low-risk, MammaPrint high-risk group was not analysed due to small numbers of LN+ patients (n=15).*
*[c]Piccart 2021 data are from the Piccart et al.[29] (2021) supplement, Table S10.*
*Adj – adjusted; CI – confidence interval; CT – chemotherapy; DFS – disease-free survival; DMFI – distant metastasis-free interval; DMFS – distant metastasis-free survival; ER – oestrogen receptor; ET – endocrine therapy; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR – hormone receptor; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; OS – overall survival; prosp – prospective; RCT – randomised controlled trial; sig – significant; y/yr – year.*

### *3.4.6 Ongoing prospective RCT of Prosigna: OPTIMA*

As described in Section 3.2, the ongoing OPTIMA study[33] is an RCT of Prosigna test-directed chemotherapy use vs. standard chemotherapy use. The review did not identify any published results of OPTIMA so far.

### *3.4.7 Observational data: Prospective use of Oncotype DX*

Several publications report observational studies or registry data for the prospective use of Oncotype DX in clinical practice. In these studies, the Oncotype DX result likely influenced the use of chemotherapy and therefore outcomes. As such, these data have limited use in comparing outcomes between test groups (prognostic ability), though they do provide large-sample data on real-world outcomes. These studies included the Clalit registry[64] in Israel (n=709), the SEER registry[65-67] in the US

(n=6,483), the National Cancer Database (NCDB)[68-72] in the US (n=25,029) and a few smaller prospective studies.[73-75] An overview of results is described here, with full results in Appendix 5, whilst data on distant recurrence are shown in Table 8.

In terms of distribution using cut-offs of RS <18 and >30 (see Appendix 5), across the Clalit[64] and SEER[65, 66] registries, 53-58% were low-risk (RS 0-17), 35-36% intermediate-risk (RS 18-30) and 7-10% high-risk (RS ≥30), which is similar to the distribution in the TransATAC study[19] (57% low, 32% intermediate, 11% high). A study in younger patients[73] (age ≤40 years) reported a greater proportion of high-risk patients (33% low, 42% intermediate, 25% high). Using an RS cut-off of >25, across the Clalit[64] and NCDB[68] registries plus a German study,[75] the distribution ranged from 81-88% (RS 0-25) and 13-19% (RS ≥26). The NCDB also reports the distribution using RS cut-offs of <11 and >25 as follows: 24% (RS 0-10), 64% (RS 11-25) and 13% (RS ≥26).

Distant recurrence data from two sources (Clalit registry[64] and the Young Women's Breast Cancer Study[73]) are shown in Table 8. Within Clalit,[64] using the RS cut-offs of <18 and >30, the 5-year DRFI was 97% in low-risk patients (7% chemotherapy use), 94% in intermediate-risk (40% chemotherapy use) and 83% in high-risk (86% chemotherapy use), with Oncotype DX being significantly prognostic despite the greater chemotherapy use in higher-risk patients (see Table 8). Using the cut-offs of RS <11 and >25, 5-year DRFI was 96% (RS 0-10), 96% (RS 11-25) and 87% (RS ≥26), with Oncotype DX again being statistically significantly prognostic. In younger patients, both Clalit[64] and the Young Women's Breast Cancer Study[73] show a statistically significant prognostic effect (see Table 8). However, in older patients (≥70 years), there was no statistically significant prognostic effect on 5-year DRFI in Clalit[64] (see Table 8).

Data on other outcomes are shown in Appendix 5. For BCSS and OS, most analyses of the Clalit,[64] SEER[65, 66] and NCDB[68, 69, 72] registries showed a prognostic effect of Oncotype DX using both the cut-offs of RS <18 and >30 and RS <11 and >25. Subgroup analyses of SEER reported statistically significant prognostic ability in white patients but non-significant results in black or other ethnicities,[65] whilst statistically significant prognostic ability was reported in both men and women,[67] though these subgroups were based on small numbers. Analyses of NCBD reported statistically significant prognostic ability in patients aged 40-50 years[70] and in patients with lobular cancer.[71]

### *3.4.8 Conclusions for prognostic data*

For all four tests, within re-analyses of trials and cohorts, the HR for distant recurrence between risk groups indicated statistically significant prognostic ability for most (though not all) analyses, both with and without adjustment for clinical factors. An analysis of the Clalit registry[64] reported that Oncotype DX was significantly prognostic for distant recurrence using both the cut-offs of RS <18 and >30 and

RS <11 and >25, despite greater chemotherapy use in higher-risk patients. In the RxPONDER prospective RCT,[28] within the study population (RS 0-25), Oncotype DX was significantly prognostic for 5-year IDFS after adjusting for clinical factors, overall and in the pre-menopausal and post-menopausal subgroups. In the MINDACT RCT,[29] within LN+ patients at high clinical risk, 8-year DMFI was 92.3% for MammaPrint low-risk vs. 80.9% for MammaPrint high-risk, despite higher chemotherapy use for high-risk patients; however, no HRs or significance tests were reported for prognostic ability.

**Table 8: Observational data for Oncotype DX (distant recurrence)**

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N ET/CT | Meno Age Clin | Test cut-offs | Distribution % | | | % risk of outcome | | | HR between test risk groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | Int | High | Low | Int | High | | |
| **Oncotype: Distant recurrence** | | | | | | | | | | | | | | |
| Clalit, Israel | Stemmer 2017[64] | LN1mic: 42%<br>LN1-3: 58%<br>100% ER+<br>100% HER2- | DRFI (0-5yr) | n=709 Var ET/CT | All meno | 18, 30 | 53 | 36 | 10 | 96.8 (7% CT) | 93.7 (40% CT) | 83.1 (86% CT) | 0-5yr: Low vs high: HR 0.19 (0.09 to 0.40)<br>0-5yr: Int vs. high: HR 0.39 (0.20 to 0.79), $p$<0.001<br>*0-5yr: Adj HR: Low vs high: HR 0.23 (0.11 to 0.50)<br>*0-5yr: Adj HR: Int vs. high: HR 0.42 (0.20 to 0.86), $p$=0.001 | Y<br>Y<br>Y*<br>Y* |
| | | | | | | 11, 25 | ≤25: 81 | | 19 | 95.7 (5% CT) | 96.0 (18% CT) | 86.9 (77% CT) | 0-5yr: $p$<0.001 | Y |
| | | | | | | ≤25, 26-30 | | | | 96.0 (15% CT) | | 91.5 (67% CT) | - | - |
| | | | | | | 18-25 | | | | | 94.4 (31% CT) | | - | - |
| | | | | n=109 Var ET/CT | Age <50 | 18, 30 | 48 | 37 | 16 | 96.2 (12% CT) | 100.0 (48% CT) | 64.2 (100% CT) | 0-5yr: $p$<0.001 | Y |
| | | | | n=464 Var ET/CT | Age 50-69 | 18, 30 | 54 | 37 | 9 | 97.6 (6% CT) | 93.5 (42% CT) | 87.8 (90% CT) | 0-5yr: $p$=0.017 | Y |
| | | | | n=136 Var ET/CT | Age ≥70 | 18, 30 | 57 | 33 | 10 | 94.7 (7% CT) | 88.7 (22% CT) | 92.9 (57% CT) | 0-5yr: $p$=0.458 | N |
| Young Women's Breast Cancer Study | Poorvu 2020[73] | Lnmic, LN1-3<br>100% ER+<br>100% HER2- | DRFS (0-6yr) | n=163 Var ET/CT | Age ≤40 | 18, 30 | 33 | 42 | 25 | 0-6yr: 85.9 (83% CT) | 0-6yr: 87.3 (97% CT) | 0-6yr: 62.8 (98% CT) | 0-6yr: $p$=0.004 | Y |
| | | | | | | 11, 25 | 9 | 54 | 37 | 0-6yr: 92.3 (79% CT) | 0-6yr: 85.2 (92% CT) | 0-6yr: 71.3 (97% CT) | 0-6yr: $p$=0.10 | N |

*[a]The last column indicates whether each HR between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*

*Adj – adjusted; CI – confidence interval; CT – chemotherapy; DRFI – distant recurrence-free interval; DRFS – distant recurrence-free survival; ER – oestrogen receptor; ET – endocrine therapy; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR – hormone receptor; int – intermediate; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; sig – significant; var – variable; y/yr – year*

## 3.5 Results: Prediction of chemotherapy benefit

### *3.5.1 Overview of predictive data in this report*

This section summarises two types of data: (a) the effect of chemotherapy vs. no chemotherapy on patient outcomes within a test risk group or range; and (b) whether this effect of chemotherapy vs. no chemotherapy differs significantly between test risk groups or ranges, i.e., whether the test is predictive of chemotherapy benefit, generally assessed via a test for interaction between chemotherapy effect and risk score.[30] Data of the above types for the LN+ population were only identified for Oncotype DX and MammaPrint. No data on predictive benefit were identified for Prosigna or EPclin in the LN+ population. In total, 14 publications[28, 31, 64, 69-72, 76-82] relating to 5 studies of Oncotype, and 2 publications[29, 32] relating to 2 studies of MammaPrint, provided data on prediction and/or effect of chemotherapy.

For Oncotype DX, the following data on the effect of chemotherapy were identified:

- A reanalysis of the SWOG-8814 RCT (Albain *et al.*, 2010),[31] in which Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy vs. no chemotherapy. This study did not report distant recurrence but did report data for DFS, BCSS and OS. HRs for chemotherapy vs. no chemotherapy were reported for Oncotype DX low-, intermediate- and high-risk groups using the cut-offs of RS 18 and 30, and interaction tests were conducted to assess whether these HRs were statistically significantly different between risk groups;
- The RxPONDER prospective RCT,[28, 76] which reported the effect of chemotherapy vs. no chemotherapy among patients with an RS of 0-25, as well as a test for interaction between RS (within the range 0-25) and effect of chemotherapy on IDFS.
- Registry data from the Clalit,[64] SEER[65-67] and NCDB[68-72] registries, reporting outcomes per risk group for patients with and without chemotherapy. A limitation is that the use or non-use of chemotherapy was not randomised, and may correlate with clinical factors which affect outcomes; therefore, data on the effect of chemotherapy from these studies should be treated with caution. No interaction tests were reported for risk group and effect of chemotherapy.

For MammaPrint, the following data on the effect of chemotherapy were identified:

- A reanalysis of two cohorts (Mook *et al.*, 2009)[32] which only reported a *p*-value for an interaction test for BCSS.
- The MINDACT prospective RCT,[29] which reported the effect of chemotherapy vs. no chemotherapy on 8-year DMFS within the mAOL high-risk, MammaPrint low-risk, LN+, HR+ HER2- subgroup. However, since no data were available for the LN+ MammaPrint high-risk group and no interaction tests were presented, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit.

### *3.5.2 Prediction of chemotherapy benefit: RCT reanalysis (Oncotype DX)*

Albain *et al.* (2010)[31] reported a reanalysis of the SWOG-8814 RCT, in which Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy vs. no chemotherapy. This study did not report on outcomes relating to distant recurrence, but did report DFS, BCSS and OS (see Table 9). HRs for chemotherapy vs. no chemotherapy were reported for Oncotype DX low-, intermediate- and high-risk groups using the cut-offs of RS <18 and >30, and interaction tests were conducted to assess whether these HRs were statistically significantly different between risk groups.

For 10-year DFS, the adjusted HR for chemotherapy vs. no chemotherapy indicated no effect of chemotherapy in the Oncotype DX low-risk group (HR 1.02; 95% CI 0.54 to 1.93; $p$=0.97); a non-significant effect of chemotherapy in the intermediate-risk group with a point estimate favouring chemotherapy (HR 0.72; 95% CI 0.39 to 1.31; $p$=0.48); and a borderline statistically significant effect of chemotherapy in the high-risk group (HR 0.59; 95% CI 0.35 to 1.01; $p$=0.033; see Table 9).[31] Similar data are presented in Table 9 for DFS at different timepoints and for BCSS and OS.

**Table 9: Prediction of chemotherapy benefit: RCT reanalysis (Oncotype DX)**

| Study Ref Design | Nodal status HR, HER2 | Outcome | N | Meno | Test cut-offs | % risk of outcome | | | | | | Abs diff CT vs no CT | | | HR for CT vs. no CT (95% CI) | | | | Interaction | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | Adj | | |
| | | | | | | CT | No | CT | No | CT | No | | | | | | | | | |
| **Oncotype DX: RCT-reanalysis: Distant recurrence** | | | | | | | | | | | | | | | | | | | | |
| No data | | | | | | | | | | | | | | | | | | | - | - |
| **Oncotype DX: RCT-reanalysis: DFS** | | | | | | | | | | | | | | | | | | | | |
| SWOG-8814 Albain 2010[31] RCT-R | LN1-3: 62% LN4+: 38% 100% HR+ 88% HER2- | DFS 0-5yr | n=367 | Post-meno | 18, 30 | - | - | - | - | - | - | - | - | - | 1.34 (0.47 to 3.82) | 0.95 (0.43 to 2.14) | 0.59 (0.32 to 1.11) | Y | *p*=0.029 (adj nodes) | Y* |
| | | DFS 0-10yr | n=367 | Post-meno | 18, 30 | 64 | 60 | - | - | 55 | 43 | 4 | - | 12 | 1.02 (0.54 to 1.93) *p*=0.97 | 0.72 (0.39 to 1.31) *p*=0.48 | 0.59 (0.35 to 1.01) *p*=0.033 | Y | *p*=0.053 (adj nodes) *p*=sig (NR) (adj various) p=0.15 (adj Allred-ER) | N* Y* N* |
| | | DFS 5-10yr | n=367 | Post-meno | 18, 30 | - | - | - | - | - | - | - | - | - | 0.88 (0.38 to 1.92) | 0.52 (0.21 to 1.27) | 0.60 (0.22 to 1.62) | Y | *p*=0.58 (cont RS, adj nodes) | N* |
| **Oncotype DX: RCT-reanalysis: BCSS and OS** | | | | | | | | | | | | | | | | | | | | |
| SWOG-8814 Albain 2010[31] RCT-R | LN1-3: 62% LN4+: 38% 100% HR+ 88% HER2- | BCSS 0-10yr | n=367 | Post-meno | 18, 30 | - | - | - | - | 73 | 54 | - | - | 19 | *p*=0.56 | *p*=0.89 | *p*=0.033 | Y | - | - |
| | | OS 0-10yr | n=367 | Post-meno | 18, 30 | - | - | - | - | 68 | 51 | - | - | 17 | 1.18 (0.55 to 2.54); *p*=0.68 Log-rank *p*=0.63 | 0.84 (0.40 to 1.78); *p*=0.65 Log-rank *p*=0.85 | 0.56 (0.31 to 1.02); *p*=0.057 Log-rank *p*=0.027 | Y | Int (adj nod): 0-10yr: *p*=0.026 0-5yr: *p*=0.016 5-10yr: *p*=0.87 | Y* Y* N* |

[a]*The last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors.*

*Abs diff – absolute difference; adj – adjusted; BCSS – breast cancer-specific survival; CI – confidence interval; CT – chemotherapy; DFS – disease-free survival; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR – hormone receptor; int – intermediate; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; OS – overall survival; prosp – prospective; pred – predictive of CT benefit; RCT – randomised controlled trial; RCT-R – RCT reanalysis; RS – Recurrence Score (Oncotype DX); sig – significant; unadj – unadjusted; yr – year*

Interaction tests were conducted for chemotherapy effect and risk group: some were statistically significant whilst others were not (see Table 9). For 5-year DFS, the interaction test was statistically significant ($p$=0.029, adjusted for N positive nodes). For 10-year DFS, the interaction test did not quite reach statistical significance when adjusted for N positive nodes ($p$=0.053), and was stated to be statistically significant when adjusted for various clinical factors ($p$-value not reported [NR]), but was no longer significant when adjusted for Allred-scored ER status ($p$=0.15). The interaction test for late DFS events (5-10 years) was not statistically significant ($p$=0.58). An interaction test was also conducted for OS (adjusted for N positive nodes); this was statistically significant at 0-5 years ($p$=0.016) and 0-10 years ($p$=0.026) but not for late events (5-10 years; $p$=0.87).[31]

### *3.5.3 Prediction of chemotherapy benefit: prospective RCT of Oncotype DX (RxPONDER)*

The RxPONDER[28] prospective RCT of Oncotype DX randomised patients with an RS of 0-25 to chemotherapy plus ET vs. ET monotherapy (see Table 10). In terms of distant recurrence, the results indicated that chemotherapy had little benefit in post-menopausal patients with an RS of 0-25; the 5-year DRFI was 95.8% with chemotherapy vs. 96.6% with no chemotherapy, an absolute difference of 0.8% favouring no chemotherapy (adjusted HR 1.12; 95% CI 0.82 to 1.52; $p$=0.49). Conversely, there was a benefit of chemotherapy in pre-menopausal patients with an RS of 0-25; the 5-year DRFI was 96.3% with chemotherapy vs. 93.9% with no chemotherapy, an absolute difference of 2.4% favouring chemotherapy (adjusted HR 0.64; 95% CI 0.43 to 0.95; $p$=0.026).[28] Similar data are presented for 5-year DRFS and IDFS, again showing a statistically significant benefit of chemotherapy in pre-menopausal patients, but not in post-menopausal patients (see Table 10).

A test for interaction was reported between RS (within the range 0-25) and the effect of chemotherapy on IDFS; no interaction test was reported for distant recurrence. The test did not show a statistically significant interaction across all patients (HR for interaction 1.02; 95% 0.98 to 1.05; $p$=0.35), with similar non-significant results in the pre-menopausal and post-menopausal subgroups; see Table 10).[28] Separate data on the effect of chemotherapy on IDFS were also presented within smaller RS ranges (RS 0-10, 11-15, 16-20, 21-25; and 0-13 and 14-25). However, there was no clear pattern or trend in the HR for the effect of chemotherapy, within either the pre-menopausal or post-menopausal groups (see Table 10). This indicates no statistically significant predictive effect within the RS 0-25 group, though RxPONDER cannot provide data on whether there is a predictive effect between the RS 0-25 and RS 26-100 groups.

**Table 10: Prediction of chemotherapy benefit: Prospective RCT of Oncotype DX (RxPONDER)**

| Study Ref Design | Nodal status HR, HER2 | Outcome | N | Meno Clin risk | Test cut-offs | % risk of outcome | | | | Abs diff CT vs no CT | | HR for CT vs. no CT (95% CI) | Interaction | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low RS≤25 | | High | | Low RS≤25 | High | | | |
| | | | | | | CT | No | CT | No | | | | | |
| **Distant recurrence: Full population** | | | | | | | | | | | | | | |
| RxPONDER Kalinsky 2021[28] ([b]Kalinsky SABCS 2021 slides[76]) Prosp RCT | LN1-3 (65% 1 node, 25% 2 nodes, 9% 3 nodes) 100% HR+ 100% HER2- | DRFS (0-5yr) | n=5,018 | All (67% post) | All ≤25 | 94.9 | 93.9 | - | - | 1.0 | - | RS ≤25: 0.88 (0.71 to 1.09), *p*=0.25 | - | - |
| **Distant recurrence: Post-menopausal** | | | | | | | | | | | | | | |
| RxPONDER | LN1-3 100% HR+ 100% HER2- | DRFS (0-5yr) | n=3,353 | Post-meno | All ≤25 | 94.4<br>94.3[b] | 94.4<br>94.8[b] | - | - | 0.1<br>-0.5[b] | - | RS ≤25: HR 1.05 (0.81 to 1.37), *p*=0.70<br>[b]RS ≤25: Adj HR 1.12 (0.88 to 1.44), *p*=0.35 | - | - |
| | | DRFI (0-5yr) | n=3,353 | Post-meno | All ≤25 | 95.8[b] | 96.6[b] | - | - | -0.8[b] | - | [b]RS ≤25: Adj HR 1.12 (0.82 to 1.52), *p*=0.49 | - | - |
| **Distant recurrence: Pre-menopausal** | | | | | | | | | | | | | | |
| RxPONDER | LN1-3 100% HR+ 100% HER2- | DRFS (0-5yr) | n=1,665 | Pre-meno | All ≤25 | 96.1<br>95.9[b] | 92.8<br>93.4[b] | - | - | 3.3<br>2.5[b] | - | RS ≤25: HR 0.58 (0.39 to 0.87), *p*=0.009<br>[b]RS ≤25: Adj HR 0.66 (0.45 to 0.97), *p*=0.033 | - | - |
| | | DRFI (0-5yr) | n=1,665 | Pre-meno | All ≤25 | 96.3[b] | 93.9[b] | - | - | 2.4[b] | - | [b]RS ≤25: Adj HR 0.64 (0.43 to 0.95), *p*=0.026 | - | - |
| | | | NR | Pre-meno | 0-13 | - | - | - | - | 2.3 | - | - | - | - |
| | | | | | 14-25 | - | - | - | - | 2.8 | - | - | - | - |
| **IDFS: Full population** | | | | | | | | | | | | | | |
| RxPONDER | LN1-3 100% HR+ 100% HER2- | IDFS (0-5yr) | n=5,018 | All (67% post) | All ≤25 | 92.2 | 91.0 | - | - | 1.2 | - | RS ≤25: 0.86 (0.72 to 1.03), *p*=0.10 | HR 1.02 (0.98 to 1.05), *p*=0.35 (adj meno) | N |

| Study Ref Design | Nodal status HR, HER2 | Outcome | N | Meno Clin risk | Test cut-offs | % risk of outcome | | | | Abs diff CT vs no CT | | HR for CT vs. no CT (95% CI) | Interaction | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low RS≤25 | | High | | Low RS≤25 | High | | | |
| | | | | | | CT | No | CT | No | | | | | |
| **IDFS: Post-menopausal** | | | | | | | | | | | | | | |
| RxPONDER | LN1-3 100% HR+ 100% HER2- | IDFS (0-5yr) | n=3,353 | Post-meno | All ≤25 | 91.3<br>91.2[b] | 91.9<br>91.9[b] | - | - | -0.6<br>-0.7[b] | - | RS ≤25: HR 1.02 (0.82 to 1.26), *p*=0.89<br>[b]RS ≤25: Adj HR 1.06 (0.87 to 1.30), *p*=0.55 | HR 1.01 (0.97 to 1.06), *p*=0.48 | N |
| | | | NR | Post-meno | 0-10 | 92.7 | 92.7 | - | - | 0.0 | - | RS 0-10: 0.72 (0.44 to 1.18) | - | - |
| | | | | | 11-15 | 93.5 | 95.8 | - | - | -2.3 | - | RS 11-15: 1.30 (0.88 to 1.92) | - | - |
| | | | | | 16-20 | 93.2 | 90.8 | - | - | 2.4 | - | RS 16-20: 0.91 (0.57 to 1.43) | - | - |
| | | | | | 21-25 | 84.8 | 93.2 | - | - | -8.4 | - | RS 21-25: 1.13 (0.75 to 1.70) | - | - |
| | | | | | 0-13 | - | - | - | - | - | - | RS 0-13: 1.01 (0.71 to 1.44) | - | - |
| | | | | | 14-25 | - | - | - | - | - | - | RS 14-25: 1.01 (0.77 to 1.33) | - | - |
| **IDFS: Pre-menopausal** | | | | | | | | | | | | | | |
| RxPONDER | LN1-3 100% HR+ 100% HER2- | IDFS (0-5yr) | n=1,665 | Pre-meno | All ≤25 | 93.9<br>93.9[b] | 89.0<br>89.0[b] | - | - | 4.9<br>4.9[b] | - | RS ≤25: HR 0.60 (0.43 to 0.83), *p*=0.002<br>[b]RS ≤25: Adj HR 0.64 (0.47 to 0.87), *p*=0.004 | HR 1.04 (0.97 to 1.12), *p*=0.26 | N |
| | | | NR | Pre-meno | 0-10 | 96.6 | 92.4 | - | - | 4.2 | - | RS 0-10: 0.47 (0.18 to 1.20) | - | - |
| | | | | | 11-15 | 95.5 | 93.3 | - | - | 2.2 | - | RS 11-15: 0.68 (0.33 to 1.37) | - | - |
| | | | | | 16-20 | 91.5 | 83.8 | - | - | 7.7 | - | RS 16-20: 0.57 (0.35 to 0.94) | - | - |
| | | | | | 21-25 | 92.4 | 85.2 | - | - | 7.2 | - | RS 21-25: 0.63 (0.30 to 1.31) | - | - |
| | | | | | 0-13 | - | - | - | - | - | - | RS 0-13: 0.49 (0.24 to 0.99) | - | - |
| | | | | | 14-25 | - | - | - | - | - | - | RS 14-25: 0.63 (0.43 to 0.91) | - | - |

*[a]Last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors.*
*[b]Additional RxPONDER data from Kalinsky 2021 SABCS slides[76]*
*Abs diff – absolute difference; adj – adjusted; CI – confidence interval; CT – chemotherapy; DRFI – distant recurrence-free interval; DRFS – distant recurrence-free survival; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR -, hormone receptor positive; IDFS – invasive disease-free survival; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; prosp – prospective; pred – predictive of CT benefit; RCT – randomised controlled trial; RS – Recurrence Score (Oncotype DX); sig – significant; unadj – unadjusted; yr – year*

### *3.5.4 Prediction of chemotherapy benefit: reanalysis of cohort (MammaPrint)*

In terms of prediction of CT benefit, the only data identified for MammaPrint was a reanalysis of two cohorts (Mook *et al.*, 2009)[32] presenting an interaction test between MammaPrint score and effect of chemotherapy on BCSS. The adjusted interaction test had a non-significant *p*-value of 0.95 (see Table 11).

**Table 11: Prediction of chemotherapy benefit: Reanalysis of cohort (MammaPrint)**

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno | Test cut-offs | % risk of outcome | | | | Abs diff CT vs. no CT | | HR for CT vs. no CT | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | High | | Low | High | Low | High | | |
| | | | | | | | CT | No | CT | No | | | | | | |
| **MammaPrint: Cohort reanalysis: BCSS** | | | | | | | | | | | | | | | | |
| NKI, Italy, VdV Cohort-R | Mook 2009[32] | LN1micro to LN3<br>79% ER+<br>84% HER2- | BCSS 0-10yr | n=347 | All meno | NR | - | - | - | - | - | - | - | - | Int *p*=0.95 (adj) | N* |

*[a]The last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors.*

*Abs diff – absolute difference; adj – adjusted; BCSS – breast cancer-specific survival; CI – confidence interval; cohort-R – cohort reanalysis; CT – chemotherapy; ER – oestrogen receptor; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; HR – hormone receptor; LN – lymph nodes (number positive); meno – menopausal; NR – not reported; prosp – prospective; pred – predictive of CT benefit; RCT – randomised controlled trial; sig – significant; unadj – unadjusted; yr – year*

### *3.5.5 Chemotherapy effect within groups: prospective RCT of MammaPrint (MINDACT)*

As noted in Section 3.4, the prospective RCT of MammaPrint (MINDACT)[29] randomised patients to chemotherapy vs. no chemotherapy if they had a discordant genomic risk (via MammaPrint) and clinical risk (via mAOL). Data were presented for the effect of chemotherapy vs. no chemotherapy on outcomes for the clinical high-risk, MammaPrint low-risk, LN+, HR+ HER2- subgroup. However, data were not analysed for the clinical low-risk, MammaPrint high-risk group due to small numbers of LN+ patients.[29] This is consistent with the company's focus on the clinical high-risk group in the Agendia submission to NICE.[83]

Within the clinical high-risk, MammaPrint low-risk, LN+, HR+, HER2- subgroup, 8-year DMFS was 91.2% with chemotherapy vs. 89.9% with no chemotherapy, an absolute difference of 1.3% favouring chemotherapy,[29] with a non-significant HR (HR 0.84; 95% CI 0.51 to 1.37; *p*=NR; Table 12). Similar data for this subgroup are presented for 8-year DMFI, DFS and OS, though no HRs were presented for the effect of chemotherapy for these outcomes (see Table 12).[29]

The DMFS HR (above) indicates that the effect of chemotherapy in clinical high-risk, MammaPrint low-risk patients was not statistically significant, but the point estimate was in favour of chemotherapy. Since all patients in the clinical high-risk, MammaPrint high-risk group were offered chemotherapy, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit.

**Table 12: Chemotherapy effect within risk groups: Prospective RCT of MammaPrint (MINDACT)**

| Study Ref Design | Nodal status HR, HER2 | Outcome | N | Meno Clin risk | Test cut-offs | % risk of outcome | | | | Abs diff CT vs no CT | | HR for CT vs. no CT (95% CI) | | Interaction | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low MMP | | High MMP | | Low MMP | High MMP | Low MMP | High MMP | | |
| | | | | | | CT | No | CT | No | | | | | | |
| **MammaPrint: Prospective RCT: Distant recurrence** | | | | | | | | | | | | | | | |
| MINDACT Piccart 2021[29b] Prosp RCT | LN1-3 100% HR+ 100% HER2- | DMFS (0-8yr) | N=658 | High mAOL[c] | >0 low, ≤0 high | 91.2 | 89.9 | - | - | 1.3 | - | 0.84 (0.51 to 1.37), p=NR | - | - | - |
| | | DMFI (0-8yr) | N=658 | High mAOL[c] | >0 low, ≤0 high | 92.3 | 90.9 | - | - | 1.4 | - | - | - | - | - |
| **MammaPrint: Prospective RCT: Other outcomes** | | | | | | | | | | | | | | | |
| MINDACT Piccart 2021[29b] Prosp RCT | LN1-3 100% HR+ 100% HER2- | DFS (0-8yr) | N=658 | High mAOL[c] | >0 low, ≤0 high | 85.3 | 82.8 | - | - | 2.5 | - | - | - | - | - |
| | | OS (0-8yr) | N=658 | High mAOL[c] | >0 low, ≤0 high | 95.5 | 94.9 | - | - | 0.6 | - | - | - | - | - |

*[a]The last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors.*

*[b]Piccart 2021 data are from the Piccart et al.[29] (2021) supplement, Table S10.*

*[c]The mAOL low-risk, MammaPrint high-risk group was not analysed due to small numbers of LN+ patients.*

*Abs diff - absolute difference; adj - adjusted; CI - confidence interval; CT - chemotherapy; DFS - disease-free survival; DMFI - distant metastasis-free interval; DMFS - distant metastasis-free survival; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor positive; LN - lymph nodes (number positive); mAOL - modified Adjuvant! Online; meno - menopausal; NR - not reported; OS - overall survival; prosp - prospective; pred - predictive of CT benefit; RCT - randomised controlled trial; sig - significant; unadj - unadjusted; yr - year*

### *3.5.6 Effect of chemotherapy within RS groups: registry data (Oncotype DX)*

Several publications report registry data for the prospective use of Oncotype DX in clinical practice, with outcomes per risk group for patients with and without chemotherapy. These studies included analyses of the Clalit registry[64, 77] in Israel (n=709), the SEER registry[78] in the US (n=2,588) and the NCDB[69-72, 79-82] in the US (n=28,591). However, use or non-use of chemotherapy was not randomised, and may correlate with clinical factors which affect outcomes; therefore, the interpretation of the data on effect of chemotherapy from these studies should be approached with caution. An overview of results is described here, with full results in Appendix 6. Data on distant recurrence are shown in Table 13, whilst Table 14 presents data on post-menopausal or older age groups, for comparison with the RxPONDER findings in post-menopausal patients.

Data on distant recurrence for chemotherapy vs. no chemotherapy are only reported for the Clalit registry[64, 77] (SEER only reports BCSS while NCBD only reports OS). Within Clalit (see Table 13), using the cut-offs of RS <18 and >30, the relationship between Oncotype DX risk group and effect of chemotherapy was unclear, with the absolute difference in 5-year DRFI favouring chemotherapy for the intermediate-risk group (difference 8.7%, $p$=0.019) but favouring no chemotherapy for the low-risk (4.8%, $p$=0.245) and high-risk (8.0%, $p$=NR) groups. However, using the cut-offs of RS 11 and 25, there appeared to be a trend towards a greater effect of chemotherapy in high-risk groups, with the absolute difference in 5-year DRFI favouring no chemotherapy in the low-risk group (13%, $p$=NR) but favouring chemotherapy in the intermediate-risk (3.4%, $p$=NR) and high-risk (17.8%, $p$=0.017) groups. Across all patients with an RS of ≤25 (irrespective of age or menopausal status), the difference in 5-year DRFI was 2.1% favouring chemotherapy, though this was not statistically significant ($p$=0.521). Data for all outcomes and subgroups are presented in Appendix 6.

### *3.5.7 Effect of chemotherapy for older patients with RS ≤25: registry data (Oncotype DX)*

Since a key finding of RxPONDER was a lack of chemotherapy benefit in post-menopausal patients with an RS of ≤25, results from registry studies for similar post-menopausal or older subgroups are presented in Table 14. All of these are analyses of 5-year OS from the NCDB database. Some analyses did show a statistically significant chemotherapy benefit in older patients with an RS of ≤25, contradicting the RxPONDER results, including analyses of patients aged 51-70 years with an RS of ≤25 ($p$=0.006),[72] patients aged >50 years with an RS of 11-25 ($p$=0.004)[69] and patients aged >50 years with an RS of 20-25 ($p$=0.019).[80] Conversely, some analyses did not show a statistically significant chemotherapy benefit in older patients, including analyses of patients with ductal carcinoma aged 50-75 years with an RS of ≤25 ($p$=NR)[71] and patients aged >70 years with an RS of ≤25 ($p$=0.69).[72] When also considering the limitations of these studies, the results do not clearly either support or refute the RxPONDER findings.

### *3.5.8 Conclusions for prediction of chemotherapy benefit data*

Some data assessing predictive ability were identified for Oncotype DX and MammaPrint. No predictive data in a LN+ population were identified for Prosigna or EPclin. In a reanalysis of the SWOG-8814 RCT,[31] using cut-offs of RS <18 and >30, adjusted HRs indicated no effect of chemotherapy on 10-year DFS in the low-risk group; a non-significant effect in the intermediate-risk group; and a borderline statistically significant effect in the high-risk group. Interaction tests for chemotherapy effect and risk group were statistically significant in some analyses but not others. The RxPONDER RCT[28] reported no benefit of chemotherapy in post-menopausal patients with an RS of 0-25. Conversely, there was chemotherapy benefit in pre-menopausal patients with an RS of 0-25. A test for interaction between RS (within the range 0-25) and effect of chemotherapy on IDFS was not statistically significant interaction across all patients or in the pre-menopausal or post-menopausal subgroups, indicating no significant predictive effect within the range RS 0-25. Within registry data for Oncotype DX, the relationship between Oncotype DX risk group and effect of chemotherapy was unclear, and no interaction tests were reported. The NCDB database[69, 71, 72, 80] reported 5-year OS within post-menopausal or older subgroups with an RS of ≤25; some studies reported a statistically significant chemotherapy benefit while others did not; therefore, the results did not clearly either support or refute the RxPONDER findings.

In terms of MammaPrint, a reanalysis of two cohorts from 2009[32] reported a non-significant interaction test between MammaPrint score and effect of chemotherapy on BCSS ($p$=0.95) indicating no predictive effect. In the MINDACT prospective RCT,[29] within the mAOL high-risk, MammaPrint low-risk, LN+, HR+ HER2- subgroup, 8-year DMFS was 91.2% with chemotherapy vs. 89.9% with no chemotherapy, an absolute difference of 1.3% favouring chemotherapy, with a non-significant HR (HR 0.84; 95% CI 0.51 to 1.37; $p$=NR). Since no data for the LN+ MammaPrint high-risk group and no interaction tests were presented, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit.

**Table 13: Effect of chemotherapy within risk groups: Registry data for Oncotype DX (distant recurrence)**

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| **Oncotype DX: Observational registry: Distant recurrence** | | | | | | | | | | | | | | | | | | | | |
| Clalit, Israel | Stemmer 2017[64] | LN1mic: 42%<br>LN1-3: 58%<br>100% ER+<br>100% HER2- | DRFI 0-5yr | n=709 | All meno | 18, 30 | 92.3 | 97.1 | 99 | 90.3 | 82 | 90 | -4.8 | 8.7 | -8.0 | *p*=0.245 | *p*=0.019 | - | - | - |
| | | | | | | 11, 25 | 83.3 | 96.3 | 98.8 | 95.4 | 97.5 | 79.7 | -13.0 | 3.4 | 17.8 | - | - | *p*=0.017 | - | - |
| | | | | | | All ≤25 | - | - | 97.7 | 95.6 | - | - | 2.1 | | - | *p*=0.521 | | - | - | - |
| | | | | | | All 18-25 | - | - | 100 | 91.8 | - | - | - | 8.2 | - | - | *p*=0.058 | - | - | - |
| | Rotem 2022 (abst)[77] | LN+<br>100% ER+<br>100% HER2- | DRFS 0-7yr | n=140 | All meno | All 26-30 | - | - | - | - | 89.4 | 78.0 | - | - | 11.4 | - | - | Not sig | - | - |

*[a]The last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors. Abs diff - absolute difference; adj - adjusted; CI - confidence interval; CT - chemotherapy; DRFI - distant recurrence-free interval; DRFS - distant recurrence-free survival; ER - oestrogen receptor; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor; int - intermediate; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; prosp - prospective; pred - predictive of CT benefit; RCT - randomised controlled trial; RS - Recurrence Score (Oncotype DX); sig - significant; unadj - unadjusted; yr - year*

**Table 14: Chemotherapy effect within risk groups: Registry data for Oncotype DX (post-menopausal or older age groups)**

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno Age Clin | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| **Oncotype: Observational registry: OS** | | | | | | | | | | | | | | | | | | | | |
| NCDB | Cao 2022 (abst)[80] | LN1-3 100% ER+ 100% HER2- | OS NR | n=NR | Age>50 | All 20-25 | - | - | - | - | - | - | - | - | - | - | Unadj: 0.521 (NR), *p*=0.019 | - | - | - |
| NCDB (cont) | Ibraheem 2019[69] | LN1-3: 97% LN4-9: 3% 100% HR+ 100% HER2- | OS 0-5yr | n=8,886 | Age>50 | All 11-25 | - | - | - | - | - | - | - | - | - | - | Adj: 0.64 (0.47 to 0.86), *p*=0.004 | - | - | - |
| NCDB (cont) | Weiser 2022[71] | LN1-3 100% HR+ 100% HER2- | OS 0-5yr | NR | Age 50-75 Ductal | All ≤25 | - | - | - | - | - | - | - | - | - | Adj: 1.12 (0.86 to 1.46) | - | - | - | - |
| NCDB (cont) | Weiser 2021[72] | LN1-3 100% HR+ 100% HER2- | OS 0-5yr | NR | Age 51-70 | All ≤25 | - | - | - | - | - | - | 1.6 | | - | Adj: 1.49 (1.12 to 1.97), *p*=0.006 | | | - | - |
| | | | | | | All 12-17 | - | - | - | - | - | - | - | 3.6 | - | - | Adj: 2.80 (1.45 to 5.24) | - | - | - |
| | | | | | | All 18-25 | - | - | - | - | - | - | - | 3.2 | - | - | Adj: 1.37 (0.92–2.05) | - | - | - |
| | | | | NR | Age>70 | All ≤25 | - | - | - | - | - | - | - | - | - | Adj: 1.1 (0.68 to 1.78), *p*=0.69 | | | - | - |

*[a]The last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors. Abs diff - absolute difference; adj - adjusted; CI - confidence interval; CT - chemotherapy; ER - oestrogen receptor; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor positive; int - intermediate; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; OS - overall survival; prosp - prospective; pred - predictive of CT benefit; RCT - randomised controlled trial; RS - Recurrence Score (Oncotype DX); sig - significant; unadj - unadjusted; yr - year*

## 3.6 Results: Decision impact

### *3.6.1 Decision impact: overview and study characteristics*

Decision impact studies assess how recommendations or decisions to use or not to use chemotherapy change before and after the test. Only decision impact studies from the UK and Europe were included because other countries may have different rates of chemotherapy use. In total, 12 publications[34-45] relating to 12 studies reported decision impact data for Oncotype DX in a LN+ population. These included five UK studies[34-38] and 7 other European (non-UK) studies[39-45] (see Table 15). All studies included a combination of patients in both pre-menopausal and post-menopausal stages, except for one study[36] which exclusively focused on post-menopausal patients. One UK study (Holt *et al.*, 2023) was reported only as an abstract[35] and as unpublished data submitted to NICE as part of the Exact Sciences submission and the Peony Breast Cancer Unit submission. No UK and European studies assessed the decision impact of MammaPrint, Prosigna or EPclin in a LN+ population.

Patients were allocated pre-test to either chemotherapy or no chemotherapy. This could be a recommendation (by a physician or multidisciplinary team) or an actual treatment decision (what the patient actually received). They were then split into four post-test groups: those whose decision/recommendation (1) remained chemotherapy, (2) remained no chemotherapy, (3) changed from no chemotherapy to chemotherapy or (4) changed from chemotherapy to no chemotherapy. These data can also be summarised in terms of the total proportion allocated to chemotherapy both pre- and post-test and the net change in chemotherapy use (see Table 15).

### *3.6.2 Decision impact results for Oncotype DX: all patients*

Across all test risk groups, the total proportion of patients allocated to pre-test chemotherapy ranged from 46% to 100% among 5 UK studies[34-38] and 38% to 100% among 7 European (non-UK) studies.[39-45] The total proportion allocated to post-test chemotherapy ranged from 13% to 31% among 5 UK studies[34-38] and 20% to 57% among 7 European (non-UK) studies[39-45] (see Table 15).

Among the 5 UK studies,[34-38] the net reduction in chemotherapy recommendations (pre-test to post-test) was 28%,[37] 37%[34] and 75%[38] across three studies, and the net reduction in chemotherapy decisions was 41%,[34] 52%[35] and 69%[36] across three studies (see Table 15). Two of these studies[36, 38] assessed only patients with an initial recommendation for chemotherapy and so it may be misleading to calculate the absolute change. Also, in two studies,[36, 37] the post-test decisions were based entirely on the test result and so their findings are less reliable. Across 7 European studies,[39-45] the net reduction in chemotherapy recommendations (pre-test to post-test) ranged from 12%[40] to 73%.[43] Two of these studies also reported changes from pre-test chemotherapy recommendation to post-test decision with a net reduction of 14%[40] and 29%[45] in chemotherapy use.

**Table 15: Decision impact: Oncotype DX (not split by test risk group)**

| Ref Country | Study, setting Years | HR, HER2 | Nodal status Clinical risk | Recom/ decision | Meno status | Test group | RxP | N pts | No CT | No CT to CT | CT | CT to no CT | Pre-test CT | Post-test CT | Net change CT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Battisti 2019 (abst)[34] **UK** | PONDx; 30 centres 2017-2018 | ER+ HER2- | LN1-3 | R-R | All (65% post) | All RS | - | 567 | - | - | - | - | 371 (65%) | 162 (29%) | -209 (-37%) |
| | | | | R-D | All (65% post) | All RS | - | 567 | - | - | - | - | 371 (65%) | 140 (25%) | -231 (-41%) |
| Holt 2023 (abst)[35] Holt 2023 [unpub] **UK** | 14 centres 2017-2022 | ER+ HER2- | LN1-3 | R-D | All (77% post) | All RS | - | 664 | 117 (18%) | 17 (3%) | 171 (26%) | 359 (54%) | 530 (80%) | 188 (28%) | -342 (-52%) |
| | | | | | Pre-meno | All RS | - | 152 | 23 (15%) | 6 (4%) | 65 (43%) | 58 (38%) | 123 (81%) | 71 (47%) | -52 (-34%) |
| | | | | | Post-meno | All RS | - | 512 | 94 (18%) | 11 (2%) | 106 (21%) | 301 (59%) | 407 (79%) | 117 (23%) | -290 (-57%) |
| Loncaster 2017[36] **UK** | Greater Manchester (NR centres) 2012-2015 | ER+ HER2- | LN+ CT indicated. Post-test decision based on RS | R-D | Post-meno | All RS | - | 65 | 0 (0%) | 0 (0%) | 20 (31%) | 45 (69%) | 65 (100%) | 20 (31%) | -45 (-69%) |
| Malam 2022[37] **UK** | Norfolk and Norwich (1 centre) 2014-2020 | ER+ HER2- | LN1-3 Post-test decision based on RS | R-R | All meno | All RS | - | 69 | 36 (52%) | 1 (1.4%) | 8 (12%) | 24 (35%) | 32 (46%) | 9 (13%) | -19 (-28%) |
| Nanda 2021 (abst)[38] **UK** | Oxford + Swansea (2 centres) 2013-2019 | ER+ HER2- | LN1-3 (inc micromets) CT indicated | R-R | All meno | All RS | - | 173 | 0 (0%) | 0 (0%) | 44 (25%) | 129 (75%) | 173 (100%) | 44 (25%) | -129 (-75%) |
| Eiermann 2013[45] **Germany** | 15 centres 2010-2011 | ER+ HER2- | LN1-3 | R-R | All meno | All RS | - | 122 | 18 (15%) | 12 (10%) | 58 (46%) | 34 (28%) | 92 (75%) | 70 (57%) | -22 (-18%) |
| | | | | R-D | All meno | All RS | - | 122 | - | - | - | - | 92 (75%) | 57 (47%) | -35 (-29%) |
| Cognetti 2021[39] **Italy** | PONDx; 27 centres 2016-2017 | ER+ HER2- | LN1-3 | R-R | All (55% post) | All RS | - | 414 | - | - | - | - | 258 (62%) | 110 (28%) | -148 (-55%) |
| Dieci 2019[41] **Italy** | ROXANE; 9 centres 2017-2018 | HR+ HER2- | LN1-3 94% high clin risk (mAOL) | R-R | All (55% post) | All RS | - | 99 | 42 (42%) | 3 (3%) | 24 (24%) | 30 (30%) | 54 (55%) | 27 (27%) | -27 (-27%) |
| Dieci 2018[40] | | | LN1-3 | R-R | All (55% post) | All RS | - | 126 | 49 (39%) | 5 (4%) | 52 (41%) | 20 (16%) | 72 (57%) | 57 (45%) | -15 (-12%) |

| Ref Country | Study, setting Years | HR, HER2 | Nodal status Clinical risk | Recom/ decision | Meno status | Test group | RxP | N pts | No CT | No CT to CT | CT | CT to no CT | Pre-test CT | Post-test CT | Net change CT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Italy** | Breast DX, 9 centres 2014-2016 | ER+ HER2- | Int clin risk | R-D | All (55% post) | All RS | - | 126 | - | - | - | - | 72 (57%) | 54 (43%) | -18 (-14%) |
| Zambelli 2020[44] **Italy** | BONDX (4 centres, Lombardy) 2017-2018 | ER+ HER2- | LN1-3 Int clin risk | R-R | All meno | All RS | - | 127 | 79 (62%) | 0 (0%) | 25 (20%) | 23 (18%) | 48 (38%) | 25 (20%) | -23 (-18%) |
| Fernandez-Perez 2021 (abst)[42] **Spain** | 9 centres (Galicia) 2013-2018 | HR+ HER2- | LN1-3 (inc micromets) | R-R | All (50% post) | All RS | - | 229 | - | - | - | - | 159 (69%) | 59 (26%) | -100 (-44%) |
| Llombart-Cussac 2023[43] **Spain** | KARMA Dx (8 centres) 2016-2017 | ER+ HER2- | LN1-3 High clin risk CT indicated | R-R | All meno | All RS | - | 150 | 0 (0%) | 0 (0%) | 41 (27%) | 109 (73%) | 150 (100%) | 41 (27%) | -109 (-73%) |

*CT - chemotherapy; D - decision; ER - oestrogen receptor; HER2 - human epidermal growth factor receptor 2; HR - hormone receptor positive; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; Pre/post-RxP - Pre/post publication of RxPONDER; R - recommendation; RS - Recurrence Score (Oncotype DX)*

### *3.6.3 Decision impact results for Oncotype DX: by risk group*

Of the 12 Oncotype DX studies, 2 UK studies[35, 36] and 4 European studies[41, 43-45] presented data by Oncotype DX risk groups (see Table 16). Five studies[35, 36, 43-45] used RS 18 and 30 cut-offs, whilst two studies used the newer cut-offs of RS 11 and 25 (one study[41]) or RS 13 and 25 (one study[35]).

Among the studies that used the cut-offs of RS 18 and 30,[35, 36, 43-45] the net change in chemotherapy recommendations or decisions (pre-test to post-test) was as follows: a decrease of 20%,[44] 68%,[35] 91%[43] and 93%[36] in the RS 0-17 risk group; a decrease of 19%,[44] 35%,[35] 37%[36] and 54%[43] in the RS 18-30 risk group; and either a 17%[36] decrease, no change,[43, 44] or a 1.7%[35] increase in the RS >30 risk group.

In the study that used cut-offs of RS 11 and 25,[41] the net change in chemotherapy recommendations (pre-test to post-test) was as follows: 52% decrease in the RS <11 risk group; 18% decrease in the RS 11-25 risk group; and 0% change in the RS >26 risk group. In the study that used cut-offs of RS 13 and 25 (UK),[35] the net change was as follows: 67% decrease in the RS 0-13 risk group; 56% decrease in the RS 14-25 risk group; and 5% increase in the RS 26-100 risk group. This study also reported results for pre- and post-menopausal subgroups and, within the post-menopausal subgroup, pre- and post-publication of the RxPONDER results.

### *3.6.4 Conclusions for decision impact data*

The net changes in the percentage of patients with a chemotherapy recommendation or decision (pre-test to post-test) among the UK studies were reductions of 28% to 75% across five Oncotype DX studies.[34-38] The net changes across European (non-UK) studies[39-45] were reductions of 12%[40] to 73% for Oncotype DX. Within studies reporting data by Oncotype DX risk group, there were greater reductions in chemotherapy recommendation in the low-risk and intermediate-risk groups than in the high-risk groups.

**Table 16: Decision impact: Oncotype DX (results by test risk group)**

| Ref Country | Study, setting Years | HR, HER2 | Nodal status Clinical risk | Recom/ decision | Meno status | Test group | RxP | N pts | No CT | No CT to CT | CT | CT to no CT | Pre-test CT | Post-test CT | Net change CT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Oncotype DX: Cut-offs of RS 18 and 30** | | | | | | | | | | | | | | | |
| Holt 2023 (abst)[35] Holt 2023 [unpub] **UK** | 14 centres 2017-2022 | ER+ HER2- | LN1-3 | R-D | All meno | RS 0-17 | - | 400 | 95 (24%) | 3 (1%) | 28 (7%) | 274 (69%) | 302 (76%) | 31 (8%) | -271 (-68%) |
| | | | | | | RS 18-30 | - | 204 | 20 (10%) | 12 (6%) | 88 (43%) | 84 (41%) | 172 (84%) | 100 (49%) | -72 (-35%) |
| | | | | | | RS >30 | - | 58 | 0 (0%) | 2 (3%) | 55 (95%) | 1 (1.7%) | 56 (97%) | 57 (98%) | +1 (+1.7%) |
| Loncaster 2017[36] **UK** | Greater Manchester (NR centres) 2012-2015 | ER+ HER2- | LN+ CT indicated. Post-test decision based on RS | R-D | Post-meno | RS 0-17 | - | 40 | 0 (0%) | 0 (0%) | 40 (100%) | 37 (93%) | 40 (100%) | 3 (8%) | -37 (-93%) |
| | | | | | | RS 18-30 | - | 19 | 0 (0%) | 0 (0%) | 19 (100%) | 7 (37%) | 19 (100%) | 12 (63%) | -7 (-37%) |
| | | | | | | RS >30 | - | 6 | 0 (0%) | 0 (0%) | 6 (100%) | 1 (17%) | 6 (100%) | 5 (83%) | -1 (-17%) |
| Eiermann 2013[45] **Germany** | 15 centres 2010-2011 | ER+ HER2- | LN1-3 | R-R | All meno | RS 0-17 | - | 67 | - | 2 (3%) | - | 30 (45%) | - | - | - |
| | | | | | | RS 18-30 | - | 44 | - | 8 (18%) | - | 4 (9%) | - | - | - |
| | | | | | | RS 31+ | - | 11 | - | 2 (18%) | - | 0 (0%) | - | - | - |
| Llombart-Cussac 2023[43] **Spain** | KARMA Dx (8 centres) 2016-2017 | ER+ HER2- | LN1-3 High clin risk CT indicated | R-R | All meno | RS 0-17 | - | 86 | 0 (0%) | 0 (0%) | 8 (9%) | 78 (91%) | 86 (100%) | 8 (9%) | -78 (-91%) |
| | | | | | | RS 18-30 | - | 57 | 0 (0%) | 0 (0%) | 26 (46%) | 31 (54%) | 57 (100%) | 26 (46%) | -31 (-54%) |
| | | | | | | RS 31-100 | - | 7 | 0 (0%) | 0 (0%) | 7 (100%) | 0 (0%) | 7 (100%) | 7 (100%) | No change |
| Zambelli 2020[44] **Italy** | BONDX (4 centres, Lombardy) 2017-2018 | ER+ HER2- | LN1-3 Int clin risk | R-R | All meno | RS 0-17 | - | 71 | 56 (79%) | 0 (0%) | 1 (1%) | 14 (20%) | 15 (21%) | 1 (1%) | -14 (-20%) |
| | | | | | | RS 18-30 | - | 48 | 23 (48%) | 0 (0%) | 16 (33%) | 9 (19%) | 25 (52%) | 16 (33%) | -9 (-19%) |
| | | | | | | RS 31-100 | - | 8 | 0 (0%) | 0 (0%) | 8 (100%) | 0 (0%) | 8 (100%) | 8 (100%) | No change |
| **Oncotype DX: Cut-offs of RS 11 and 25** | | | | | | | | | | | | | | | |
| Holt 2023 (abst)[35] Holt 2023 [unpub] **UK** | 14 centres 2017-2022 | ER+ HER2- | LN1-3 | R-D | All meno | RS 0-13 | - | 261 | 68 (26%) | 2 (1%) | 13 (5%) | 178 (68%) | 191 (73%) | 15 (6%) | -176 (-67%) |
| | | | | | | RS 14-25 | - | 305 | 48 (16%) | 7 (2%) | 72 (24%) | 178 (58%) | 250 (82%) | 79 (26%) | -171 (-56%) |
| | | | | | | RS 26-100 | - | 98 | 1 (1%) | 8 (8%) | 86 (88%) | 3 (3%) | 89 (91%) | 94 (96%) | +5 (+5%) |
| | | | | | Pre-meno | RS 0-25 | - | 127 | 23 (18%) | 4 (3%) | 43 (34%) | 57 (45%) | 100 (79%) | 47 (37%) | -53 (-42%) |
| | | | | | | RS 26-100 | - | 25 | 0 (0%) | 2 (8%) | 22 (88%) | 1 (4%) | 23 (92%) | 24 (96%) | +1 (+4%) |
| | | | | | Post-meno | RS 0-25 | - | 439 | 93 (21%) | 5 (1%) | 42 (10%) | 299 (68%) | 341 (78%) | 47 (11%) | -294 (-67%) |

| Ref Country | Study, setting Years | HR, HER2 | Nodal status Clinical risk | Recom/ decision | Meno status | Test group | RxP | N pts | No CT | No CT to CT | CT | CT to no CT | Pre-test CT | Post-test CT | Net change CT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Pre-RxP | 292 | 57 (20%) | 1 (0.3%) | 40 (14%) | 194 (66%) | 234 (80%) | 41 (14%) | -193 (-66%) |
| | | | | | | | Post-RxP | 147 | 36 (24%) | 4 (3%) | 2 (1%) | 105 (71%) | 107 (73%) | 6 (4%) | -101 (-69%) |
| | | | | | | RS 26-100 | - | 73 | 1 (1%) | 6 (8%) | 64 (88%) | 2 (3%) | 66 (90%) | 70 (96%) | +4 (+5%) |
| | | | | | | | Pre-RxP | 44 | 1 (2%) | 5 (11%) | 36 (82%) | 2 (5%) | 38 (86%) | 41 (93%) | +3 (+7%) |
| | | | | | | | Post-RxP | 29 | 0 (0%) | 1 (3%) | 28 (97%) | 0 (0%) | 28 (97%) | 29 (100%) | +1 (+3%) |
| Dieci 2019[41] **Italy** | ROXANE; 9 centres 2017-2018 | HR+ HER2- | LN1-3 Most high clin risk (mAOL) | R-R | All (55% post) | RS<11 | - | 31 | - | - | - | - | 19 (61%) | 3 (10%) | -16 (-52%) |
| | | | | | | RS 11-25 | - | 61 | - | - | - | - | 28 (46%) | 17 (28%) | -11 (-18%) |
| | | | | | | RS 11-17 | - | NR | - | - | - | - | - (49%) | - (19.5%) | NR (-29.5%) |
| | | | | | | RS 18-25 | - | NR | - | - | - | - | - (40%) | - (45%) | NR (+5%) |
| | | | | | | RS ≥26 | - | 7 | - | - | - | - | 7 (100%) | 7 (100%) | No change |

*CT - chemotherapy; D - decision; ER - oestrogen receptor; HER2 - human epidermal growth factor receptor 2; HR - hormone receptor positive; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; Pre/post-RxP - Pre/post publication of RxPONDER; R - recommendation; RS - Recurrence Score (Oncotype DX)*

## 3.7 Results: HRQoL and anxiety

### *3.7.1 Overview of data on HRQoL and anxiety*

No studies (or subgroups) were identified which assessed HRQoL or anxiety associated with use of tumour profiling tests in a LN+ population. Therefore, a brief summary of such studies in a LN0 or mixed nodal status population is provided below (these were all included in the previous review by Harnan *et al.*[10] for NICE DG34 and no subsequent studies were identified). The studies described below are not counted as included studies for the purposes of the PRISMA flow chart; see Appendix 2.

### *3.7.2 Overview of data on HRQoL and anxiety in a LN0 or mixed population*

*Oncotype:* Of two studies of Oncotype DX in LN0/LN+ patients in the USA, one (Evans *et al.*, 2016)[84] reported no difference between pre- and post-test values on the Impact of Events Scale ($p$=0.09) and no statistically significant interaction with RS risk group. The other (Lo *et al.*, 2010)[85] reported a statistically significant improvement in overall State-Trait Anxiety Inventory (STAI) score between pre- and post-test values ($p$=0.007), but no statistically significant change in HRQoL measured via Functional Assessment of Cancer Therapy - Breast cancer (FACT-B) or Functional Assessment of Cancer Therapy - General (FACT-G) ($p$=0.55 and $p$=0.49, respectively).

*MammaPrint:* A study of MammaPrint (Retel *et al.*, 2013)[86] which included LN0/LN+ patients screened for MINDACT in the Netherlands used Lynch's distress scale and Lerman's Cancer Worry Scale, and reported statistically significantly higher distress when the genomic test failed; when the patient was categorised as high-risk by both clinical scoring and MammaPrint; and in patients with discordant results when the treatment matched the MammaPrint risk but not the clinical risk. Only patients with high clinical risk and no genomic test result, or high clinical risk and high genomic risk, had a statistically significant decrease in HRQoL via FACT-B.

*EndoPredict:* One study of EndoPredict[87] in LN0/LN+ patients in England reported a statistically significant decrease in STAI for those whose treatment decision changed from chemotherapy to no chemotherapy on the basis of EndoPredict ($p$=0.045), and an increase in STAI for those whose treatment decision changed from no chemotherapy to chemotherapy ($p$=0.001).

*Prosigna:* Two studies assessed Prosigna in LN0 patients in Spain (Martin *et al.*, 2015)[88] and Germany (Wuerstlein *et al.*, 2016).[89] In both studies, state anxiety reduced significantly in low-risk patients ($p$<0.001 and $p$=0.008) but not in the intermediate- or high-risk groups. Both studies reported FACT-G; one[88] reported no change in overall scores, whereas the other[89] reported a statistically significant change in emotional and physical well-being ($p$=0.030, $p$=0.005, respectively).

### *3.7.3 Conclusions for HRQoL and anxiety data*

No studies of HRQoL or anxiety were identified in a LN+ population. Across studies undertaken in a LN0 or mixed population, some reported a significant improvement in anxiety before and after testing, whilst others reported no significant change in anxiety or HRQoL. Patients reported a decrease in anxiety after a low-risk test result or when their treatment was downgraded to no chemotherapy post-test, but an increase in anxiety when treatment was upgraded to chemotherapy, or after scoring high-risk both on the test and clinical measures.

# 4. COST-EFFECTIVENESS

This chapter presents a systematic review of published economic evaluations of tumour profiling tests to guide treatment decisions in people with ER+, HER2-, LN+ early breast cancer (Section 4.1), a summary and critique of the economic models submitted to NICE by the test manufacturers (Section 4.2) and the methods and results of an independent economic analysis undertaken by the EAG (Section 4.3). A discussion of the key issues around the cost-effectiveness of the tumour profiling tests is presented in Section 4.4.

## 4.1 Review of existing economic analyses

### *4.1.1 Cost-effectiveness review - methods*

Systematic searches were undertaken to identify existing economic evaluations of tumour profiling tests to guide treatment decisions in people with ER+, HER2-, LN+ early breast cancer. The review includes studies identified within the previous review undertaken to inform NICE DG34 (Harnan *et al.*[10]) as well as more recent studies published since February 2017 (the cut-off date for the search applied in Harnan *et al.*[10]). The review was undertaken with the purpose of exploring methodological choices and their potential relevance to the current decision problem, rather than to assess the results of the published economic evaluations or the potential sources of bias which might affect these.

A systematic search was undertaken to identify all economic evaluations of the four tumour profiling tests listed in the NICE scope[11] (Oncotype DX, EndoPredict, MammaPrint and Prosigna) for breast cancer.

Literature searching for economic evaluation studies was undertaken in May 2023 in the following electronic databases:

- MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations: Ovid, 1946 to present
- EMBASE: Ovid, 1974 to present
- Science Citation Index Expanded (SCI-E): Web of Science, 1900 to present
- Conference Proceedings Citation Index – Science (CPCI): Web of Science, 1990 to present.

The search strategies comprised Medical Subject Heading (MeSH) or Emtree Thesauri terms and free-text synonyms for: (i) ‘tumour profiling tests’ and ‘breast cancer’ and (ii) ‘breast cancer’ only. Searches for all four tests were limited by publication date from 2017. Searches were translated across databases and were not limited by language. The search strategies are presented in Appendix 1. Search filters designed to identify economic evaluations and reviews were used in MEDLINE and other databases, where appropriate. Reference and citation searching of included papers was also undertaken. In addition, economic studies listed in the Cytel cost-effectiveness analysis (CEA) report[83] (provided as

part of the Agendia company submission [CS]), the Exact Sciences CS[22] and RFI documents provided to NICE by Veracyte[25] and Myriad[24] were checked to ensure that no relevant studies had been missed by the electronic searches.

In order to be considered potentially relevant for inclusion in the review, studies were required to meet all of the following criteria:

- Full economic evaluations comparing tumour profiling tests for breast cancer against other tools and/or current practice
- Published in English
- Available in full text format (studies which were available in abstract form only were excluded from the review)
- Relevant to the population included within the final NICE scope.[11] Studies were only considered includable if they related to patients with ER+, HER2-, LN+ early breast cancer. Studies which reflect a mixed population were included only if the majority of the population used to inform clinical outcomes in the model had LN+ disease (≥80% patients) or if subgroup analyses for LN+ women were presented separately.

#### *4.1.2 Cost-effectiveness review results - summary of included studies*

Following de-duplication, the electronic searches identified a total of 404 studies. Of these, 65 studies were deemed to be potentially includable and full texts were obtained for further scrutiny. Five of these studies met the inclusion criteria for the review.[10, 90-93] Seven further studies[94-100] which were included in the previous systematic review[10] reflected a LN+ population and were also included in this review. No additional studies were identified from handsearching the reference lists for the SLRs reported in the submission from Exact Sciences,[22] the Cytel CEA report[83] or from the information provided to NICE by Veracyte and Myriad.[24, 25] The scope addressed within the 12 included studies and the key aspects of the modelling approaches are summarised Table 17 and Table 18, respectively.

The 12 included economic studies were undertaken to reflect a range of settings, including the UK, the US, Canada and Germany. Most of the included studies adopted a direct health care perspective. Where reported, the time horizon ranged from 25 years to the patient's remaining lifetime. Ten of the 12 included studies reflected an exclusively LN+ population or reported separate subgroup analyses for women with LN+ disease; the remaining two studies[91, 93] included mixed cohorts in which the majority of patients were reported to have LN+ disease. Where reported, the modelled populations range between the ages of 56 and 62 years for most studies. All but one of the included studies[91] evaluated Oncotype DX. EndoPredict, Prosigna and MammaPrint were each included in less than half of the included studies. Across all studies, the comparator was consistently either current decision-making (i.e., no

tumour profile testing) or chemotherapy for all patients. None of the studies reported incremental cost-effectiveness ratios (ICERs) comparing tumour profiling tests against each other.

With the exception of one study,[99] all of the included economic analyses adopted a cohort-level hybrid modelling approach comprising a decision tree to determine genomic risk classification and a state transition (Markov) component to estimate long-term outcomes. The cycle lengths applied in the Markov models ranged from 1 month to 1 year. The Markov models typically included three key health states: (i) relapse-free; (ii) DM and (iii) dead. Several models also included further health states describing the impact of short- and/or long-term complications associated with chemotherapy, including: nausea/vomiting or other toxicity; febrile neutropenia (FN); acute myeloid leukaemia (AML); heart failure (HF) and chronic myeloid leukaemia (CML). One model included a separate health state for local recurrence (LR).[95] The majority of models which evaluated Oncotype DX assumed that this test is predictive of chemotherapy benefit, whereby the relative treatment effect of CET versus ET alone is assumed to differ according to Oncotype DX RS. Only one study[93] reported an analysis which included an assumption of predictive benefit for MammaPrint and Prosigna. None of the studies included an assumption of predictive benefit for EndoPredict. There was variation amongst the included models regarding assumptions about the extent of chemotherapy use with and without tumour profiling tests – some studies compared tumour profiling tests against a strategy of chemotherapy for all, whilst others applied estimates of the proportion of patients receiving chemotherapy with/without testing from published literature and/or from routine data.

Only one of the included studies included an analysis of all four tumour profiling tests listed in the final NICE scope[11] for this appraisal (Harnan *et al.*[10]). As newer relevant clinical evidence has been published since this economic model was developed - in particular, RxPONDER[28] and longer-term follow-up data from MINDACT[29] - and because treatment pathways for breast cancer have changed since the publication of NICE DG34,[13] none of the existing published studies identified by the review provide a sufficient basis for informing the current appraisal.

**Table 17: Existing economic evaluations - analytic scope**

| Author | Population | Percent LN+ | Age | Intervention | Comparator | Country | Perspective | Time horizon | Discount rate |
|---|---|---|---|---|---|---|---|---|---|
| Berdunov *et al.* (2021)[90] | Patients with ER+/HER2- EBC and one to three positive axillary lymph nodes, unrestricted by clinical or genomic risk | 100% | Starting age unclear | Oncotype DX | Clinical risk tools alone | UK | NHS & PSS | Lifetime | 3.5% |
| Hinde *et al.* (2019)[91] | Women with ER+, HER2- EBC | 95%* | Mean age 56.5 years | EndoPredict (EPclin) | Standard risk tools only | UK | NHS | Lifetime | 3.5% |
| Masucci *et al.* (2019)[92] | Patients with ER+, HER2-, LN+ EBC | 100% | Mean age 60 years | Oncotype DX, MammaPrint, Prosigna, MammaTyper, IHC4-AQUA, IHC4 | Current practice | Canada | Health care payer | Lifetime | 1.5% |
| Harnan *et al.* (2019)[10] | Patients with ER+, HER2-, LN+ EBC | 100% in LN+ subgroup | Mean age 58 years | Oncotype DX, EPclin, Prosigna, IHC4+C, MammaPrint | Current practice | UK | NHS & PSS | Lifetime | 3.5% |
| Hall *et al.* (2017)[93] | Women aged 40 years or older with ER+, HER2-, clinically high risk (1-9 axillary lymph nodes, or LN0 with a tumour size ≥30mm) surgically treated early invasive breast cancer | 81% | Starting age unclear | Oncotype DX, MammaPrint, Prosigna, MammaTyper, IHC4-AQUA, IHC4 | Chemotherapy for all | UK | NHS | Lifetime | 3.5% |
| Stein *et al.* (2016)[94] | ER+, HER2- ESBC patients | 100% | Median age 58 years | Oncotype DX; MammaPrint/ Bluetest; Prosigna | Chemotherapy for all | UK | NHS | Lifetime (up to age 100 years) | 3.5% |
| Hannouf *et al.* (2014)[95] | Post-menopausal women with ER+/PR+ axillary LN+ ESBC | 100% | Mean age 61 years | Oncotype DX | Current practice | Canada | Canadian public health care system | Lifetime | 5.0% |
| Blohmer *et al.* (2013)[96] | Patients with ER+, HER2-, LN0 or LN+ (up to 3 nodes) ESBC. | 100% in LN+ subgroup | Mean age 56.3 years | Oncotype DX | Conventional diagnostic procedures | Germany | Health care payer | 30 years | 3.0% |

| Author | Population | Percent LN+ | Age | Intervention | Comparator | Country | Perspective | Time horizon | Discount rate |
|---|---|---|---|---|---|---|---|---|---|
| Lamond *et al.* (2012)[97] | ER-sensitive, LN0 and LN+ BC | 100% in LN+ subgroup | Median age 50 years | Oncotype DX | Current practice (population-based study) | Canada | Canadian health care system perspective | 25 years | 3.0% |
| Hall *et al.* (2012)[98] | LN+, ER+ ESBC | 100% | Baseline age 60 years | Oncotype DX | Standard care (chemotherapy for all) | UK | NHS | Lifetime (up to maximum age 100 years) | 3.5% |
| Wong *et al.* (2012)[99] | Women with LN+ HR+ breast cancer (1-3 nodes) | 100% | Reflective of RxPONDER | Oncotype DX | Current care (US NCCN guidelines) | US | Payer | Lifetime (40 years) | 3.0% |
| Vanderlaan *et al.* (2011)[100] | Minimally LN+, ESBC | 100% | Mean age 62 years | Oncotype DX | Current care (US NCCN guidelines) | US | US payer (managed care) perspective | 30-years | 3.0% |

** The supplementary appendices to Hinde et al. include a histogram which indicates that most patients included in the decision impact study used to inform the model (Bloomfield et al.) had LN+ breast cancer. Further communication with the authors of the decision impact study indicates that most of these patients actually had LN0 disease. Hinde et al. has been retained in this review for completeness.*

*ER - oestrogen receptor; HR - hormone receptor; PR - progesterone receptor; HER2 - human epidermal growth factor receptor 2; LN - lymph node; EBC - early breast cancer; ESBC - early stage breast cancer; UK - United Kingdom; US - United States; NCCN - National Comprehensive Cancer Network; NHS - National Health Service; PSS - Personal Social Services; mm - millimetre*

**Table 18: Existing economic evaluations - modelling approach and assumptions regarding predictive benefit and chemotherapy use**

| Author | Model approach | Cycle length | Model type | Does model claim predictive benefit for test? | Assumptions on chemotherapy use | Long-term health states |
|---|---|---|---|---|---|---|
| Berdunov *et al.* (2021)[90] | Decision tree and Markov model | 6 months | Classification to low-, intermediate- and high-risk | Yes - predictive benefit included in base case analysis. Scenarios assuming no predictive benefit also presented | Chemotherapy use following RS based on the Clalit registry. Chemotherapy use in current practice based on NCRAS data used in DG34 | 4 states: (1) recurrence-free; (2) distant recurrence; (3) AML; (4) dead |
| Hinde *et al.* (2019)[91] | Decision tree and Markov model | 1 year | Classification to low-, intermediate- and high-risk | No - single HR applied for chemotherapy benefit | Chemotherapy with and without EPclin drawn directly from trial (Bloomfield *et al.*) | 3 states: (1) disease-free; (2) metastases; (3) death |
| Masucci *et al.* (2019)[92] | Markov model | 1 year | Classification to low-, intermediate- and high-risk | Yes - HRs based on SWOG-8814 and clinical expert opinion | Based on literature and clinical opinion | 9 states: (1) chemotherapy; (2) chemotherapy nausea/vomiting; (3) chemotherapy febrile neutropenia; (4) chemotherapy; (5) disease-free; (6) distant recurrence; (7) CHF; (8) leukaemia; (9) death |
| Harnan *et al.* (2019)[10] | Decision tree and Markov model | 6 months | Classification to low-, intermediate- and high-risk | Base case analyses assume no predictive benefit for any test. Scenario analysis presented for predictive benefit for Oncotype DX only | Chemotherapy use following test based on literature. Chemotherapy use in current practice based on NCRAS data | 4 states: (1) recurrence-free; (2) distant recurrence; (3) AML; (4) dead |
| Hall *et al.* (2017)[93] | Decision tree and modified Markov model | 1 year | Classification to low- and high-risk | Yes - predictive benefit incorporated by modelling log HR for 10-year RFS as linear function of Oncotype DX RS. | All high-risk patients receive chemotherapy | 6 health states: (1) disease-free; (2) distant recurrence; (3) local recurrence; (4) disease-free after local recurrence; (5) heart failure; (6) dead. |
| Stein *et al.* (2016)[94] | Decision tree and modified Markov model | 1 year | Classification to low- and high-risk | Separate analyses undertaken including predictive benefit and assuming constant benefit across risk groups | All high-risk patients receive chemotherapy | 7 states: (1) disease-free; (2) distant recurrence; (3) local recurrence; (4) disease-free after local recurrence; (5) CHF; (6) CML; (7) dead. |
| Hannouf *et al.* (2014)[95] | Markov | 1 month | Classification to low-, intermediate- and high-risk with separate Markov nodes for | Unclear – appears to assume predictive benefit | Model assumes 50% IR patients receive chemotherapy | ET only model - 5 states: (1) remission; (2) local recurrence; (3) distant recurrence; (4) dead. CT+ET model - 5 states: (1) |

| Author | Model approach | Cycle length | Model type | Does model claim predictive benefit for test? | Assumptions on chemotherapy use | Long-term health states |
|---|---|---|---|---|---|---|
| | | | CT+ET versus ET alone (accounting for chemotherapy-related AEs) | | | remission with chemotherapy SAEs; (2) remission without chemotherapy SAEs; (3) local recurrence; (4) distant recurrence; (5) dead. |
| Blohmer *et al.* (2013)[96] | Decision tree and Markov model | 1 year | Classification to low-, intermediate- and high-risk | Yes - relative risk reductions of 0% applied to LR and IR, relative risk reduction of 41% applied to HR | Based on data reported by Eiermann *et al.* | 3 states: (1) recurrence-free; (2) distant recurrence; (3) dead |
| Lamond *et al.* (2012)[97] | Markov | 1 month | Classification to low-, intermediate- and high-risk | Yes - only in low risk and high risk | For no test, based on Canadian population-based study; for test, based on RS score. Usage in intermediate group assumed to be the same in both groups | 10 states: (1) chemotherapy; (2) CINV; (3) FN; (4) disease-free; (5) local relapse; (6) distant relapse; (7) treated local relapse; (8) AML/MDS; (9) CHF; (10) dead. |
| Hall *et al.* (2012)[98] | Decision tree and modified Markov model | Not reported | Classification to low- and high-risk | Unclear - data contained within the appendices appear to suggest predictive benefit is modelled | All high-risk patients receive chemotherapy | 6 states: (1) disease-free; (2) distant recurrence; (3) local recurrence; (4) disease-free after local recurrence; (5) CHF; (6) dead. |
| Wong *et al.* (2012)[99] | Decision tree with partitioned survival approach to determine sojourn time | Not reported | For patients whose treatment decision was based on US NCCN criteria classification to low-risk or high-risk. For patients whose treatment was based on the Oncotype DX test results classification to low-, intermediate- or high-risk | Yes – different treatment effects applied for each risk category | ~55% women assumed to receive chemotherapy | Not clearly reported - appears to be 3 states: (1) disease-free; (2) relapsed; (3) dead. |
| Vanderlaan *et al.** (2011)[100] | Appears to be Markov | Not reported | Classification to low- and high-risk. | No - same recurrence rates for all high-risk patients | 71% of women in usual care assumed to receive chemotherapy treatment | 3 states: (1) non-progressed disease; (2) progressed disease; (3) death. |

*NCRAS - National Cancer Registration and Analysis Service; DG - Diagnostics Guidance; AML - acute myeloid leukaemia; HR - hazard ratio; SWOG - Southwest Oncology Group; RS - recurrence score; CHF - congestive heart failure; ET - endocrine therapy; CT - chemotherapy; CML - chronic myeloid leukaemia; AE - adverse event; SAE - serious adverse event; CINV - chemotherapy-induced nausea and vomiting; FN - febrile neutropenia; MDS - myelodysplastic syndromes; US - United States; NCCN - National Comprehensive Cancer Network*

## 4.2 Review and critique of economic analyses of tumour profiling tests submitted by the test manufacturers

This section provides a summary and critique of the economic analyses submitted by the test manufacturers. Executable economic models were submitted to NICE by Exact Sciences (Oncotype DX) and Agendia (MammaPrint). No submissions were received from Myriad (EPclin) or Veracyte (Prosigna).

### *4.2.1 Exact Sciences model summary and critique (Oncotype DX)*

4.2.1.1 Summary of economic analysis submitted by Exact Sciences

In May 2023, Exact Sciences submitted an executable economic model and an accompanying written submission which details the methods and results of the model (hereafter referred to as the Exact Sciences CS.[22]). The company also provided responses to clarification questions from the EAG in June 2023,[101] which included an updated version of the economic model. The executable model is an adaptation of the earlier economic analysis reported by Berdunov *et al.*[90] (see Section 4.1.2) which in turn, was based largely on the EAG's model developed to inform NICE DG34 (Harnan *et al.*[10]). The Exact Sciences model differs from the model developed to inform DG34 in that it includes evidence on test risk classifications and DRFI from the RxPONDER trial,[76] as well as other updated parameter estimates which are intended to reflect changes in the downstream breast cancer pathway since DG34 was published in 2018.

The Exact Sciences CS[22] presents cost-effectiveness estimates for Oncotype DX versus clinical-pathological tools alone in terms of the incremental cost per QALY gained from the perspective of the NHS and PSS in England over a 45 year (lifetime) time horizon. The model applies a 6-month cycle length and includes half-cycle correction to account for the timing of events. Health outcomes and costs are discounted at a rate of 3.5% per annum. Costs are valued at 2020 prices.

The Exact Sciences base case analysis is presented across three populations: (i) the overall ER+, HER2-, LN+ (1-3 nodes) early breast cancer population; (ii) pre-menopausal women with ER+, HER2-, LN+ early breast cancer and (iii) post-menopausal women with ER+, HER2-, LN+ early breast cancer. Men with early breast cancer are not considered in the model. Comparisons of Oncotype DX versus other tumour profiling tests (MammaPrint, EPclin and Prosigna) are not included in the company's base case analyses but are included in additional exploratory analyses presented in the CS.

The general structure of the Exact Sciences model is similar to the model used to inform DG34.[13] The model structure adopts a hybrid approach comprising an initial decision tree component which stratifies patients according to their genomic risk based on the tumour profiling test result, followed by a Markov component which estimates long-term health outcomes and costs conditional on genomic risk and

whether the patient receives adjuvant CET or ET alone. The decision tree component includes three levels (low-, intermediate- and high-risk), although DRFI is assumed to be the same for patients with low risk (RS <13) and intermediate risk (RS 13-25). The long-term Markov model includes four health states: (i) recurrence-free; (ii) DM; (iii) AML and (iv) dead. LR is captured as a transient event in a proportion of patients who develop DM. Within the base case analysis and all scenario analyses, the model assumes that Oncotype DX is predictive of chemotherapy benefit, with an HR for distant recurrence for CET versus ET alone of 0.89 assumed in the Oncotype DX RS 0-25 category and an HR of 0.59 assumed in the RS >25 category for the overall LN+ population.[31, 76] The equivalent HRs for CET versus ET in the RS 0-25 category for the pre-menopausal and post-menopausal subgroups are 0.64 and 1.12, respectively; the HR of 0.59 is also applied to the RS >25 category in these subgroup analyses.[31, 76]

QALYs are modelled as a function of whether patients receive CET or ET alone, which subsequently determines the risk of DM, AML and death. The model includes a short-term disutility associated with chemotherapy-related toxicity in the first model cycle which corresponds to once-only QALY loss of 0.038 for all patients receiving chemotherapy. A once-only QALY loss is also applied for patients experiencing LR in any model cycle. QALYs are adjusted for increasing age using utility multipliers based on Ara and Brazier[102] which are aggregated into age bands (one band for patients aged <30 years, one band for those aged >85 years, and 5-year bands for those aged 30-85 years).

The model includes resource costs associated with:

- The Oncotype DX test (the costs of other tumour profiling tests are included in exploratory analyses only).
- Adjuvant therapy, including chemotherapy, ET and supportive medications
- Management of AEs
- Health state management costs whilst patients are recurrence-free (mammograms and outpatient visits)
- Treatment of LR
- Treatments for DM
- Treatments for AML
- End of life care.

The scenarios presented in Exact Sciences CS are summarised in Table 19.

**Table 19: Summary of economic comparisons presented in the Exact Sciences CS**

| Analysis | Population | Intervention and comparators | Key sources of DRFI risk and chemotherapy benefit | Chemotherapy benefit assumptions | Additional EAG comments |
|---|---|---|---|---|---|
| Base case analysis, overall LN+ population | ER+, HER2-, LN+ (1-3 nodes), pre- and post-menopausal | • Oncotype DX<br>• Clinical-pathological tools alone | RxPONDER,[76] TransATAC,[19] SWOG-8814[31] | Predictive benefit assumed for Oncotype DX | Analysis uses weighted DRFI risk and HRs for the pre-menopausal and post-menopausal subgroups. |
| Base case analysis, pre-menopausal LN+ subgroup | ER+, HER2-, LN+ (1-3 nodes), pre-menopausal | Same as overall LN+ population | RxPONDER (pre-menopausal subgroup),[76] TransATAC,[19] SWOG-8814[31] | Predictive benefit assumed for Oncotype DX | Relevant to decision problem set out in final NICE scope.[11] |
| Base case analysis, post-menopausal LN+ subgroup | ER+, HER2-, LN+ (1-3 nodes), post-menopausal | Same as overall LN+ population | RxPONDER (post-menopausal subgroup),[76] TransATAC,[19] SWOG-8814[31] | Predictive benefit assumed for Oncotype DX | Relevant to decision problem set out in final NICE scope.[11] |
| Exploratory analyses | ER+, HER2-, LN+ (1-3 nodes)<br><br>Data for comparators reflect exclusively or mostly post-menopausal patients | • Oncotype DX<br>• MammaPrint<br>• EndoPredict EPclin<br>• Prosigna<br>• Clinical-pathological tools alone<br><br>All analyses presented as pairwise comparisons of test versus clinical-pathological tools alone | RxPONDER,[76] TransATAC,[19] SWOG-8814,[31] MINDACT,[29] EBCTCG[15] | Predictive benefit assumed for Oncotype DX | Relevant to decision problem set out in final NICE scope.[11] |

*CS - company's submission; EAG - External Assessment Group; DRFI - distant recurrence-free interval; HR - hazard ratio; ER - oestrogen receptor; HER2 - human epidermal growth factor receptor 2; LN - lymph node; NICE - National Institute for Health and Care Excellence; SWOG - Southwest Oncology Group; EBCTCG - Early Breast Cancer Trialists' Collaborative Group*

*Key assumptions applied in the Exact Sciences base case analyses*

The Exact Sciences base case analyses employ the following key assumptions:

- Oncotype DX is predictive of chemotherapy benefit. This benefit is captured implicitly through the use of observed 5-year DRFI estimates for CET versus ET alone in the RS 0-25 groups from RxPONDER,[76] together with the use of external data to estimate the risks of DM with ET and CET in women with an RS of >25 (based on TransATAC[19] and SWOG-8814[31]). EPclin and Prosigna are not predictive of chemotherapy benefit (HR=0.76 in all states), although slightly different HRs are applied to MammaPrint low- and high-risk patients (low-risk HR=0.85 versus high-risk HR=0.79).
- In the absence of tumour profiling testing, most patients (~80%) will receive CET.
- The baseline risk of developing DM due to breast cancer with ET alone is reduced by 50% at 10 years; this reduction in baseline risk is retained indefinitely for the remainder of the modelled time horizon.
- Patients who develop DM receive a cyclin-dependent kinase 4/6 inhibitor (CDK4/6i), CET and/or ET. Patients may receive up to three lines of therapy for DM.
- Once patients develop AML, this diagnosis determines the patient's subsequent prognosis regardless of their prior history of distant recurrence of breast cancer.
- Negative effects of chemotherapy on HRQoL are applied for one year.
- Health outcomes and costs differ between pre-menopausal and post-menopausal women, as reflected in the subgroup analyses of RxPONDER.[76]

*Evidence sources used to inform the Exact Sciences model*

The evidence sources used to inform the parameters of the Exact Sciences model are summarised in Table 20, together with brief comments from the EAG.

**Table 20: Key evidence sources used to inform the Exact Sciences base case analyses (overall LN+ population)**

| Parameter group | Source | EAG comments |
|---|---|---|
| ***Clinical parameters*** | | |
| Start age | Unclear | The start age of 55 years is not cited in Exact Sciences CS.[22] |
| Test risk classification probabilities | RS 0-13 and RS 14-25: RxPONDER[28]<br>RS >25: Number of women who were excluded from RxPONDER[28] due to RS >25 divided by number of women registered for screening in the trial | Assumes that all other women who were screened and excluded from RxPONDER[28] had an RS of 0-25. |

| Parameter group | Source | EAG comments |
|---|---|---|
| DRFI probabilities for ET alone | RS 0-25: RxPONDER, ET arm[28]<br>RS >25: TransATAC[19] | Use of external data from TransATAC[19] is necessary because patients with an RS of >25 were excluded from RxPONDER.[28] |
| Risk tapering | Ward *et al*[103] and expert opinion | Same as DG34 model up to 15 years. |
| Chemotherapy probability under current decision-making | Holt *et al.*[17] | Unpublished decision impact study of Oncotype DX in women with LN+ early breast cancer. These data were submitted to NICE as part of the Exact Sciences CS and the submission from the Peony Breast Cancer Unit. |
| Chemotherapy probability conditional on Oncotype DX result | Holt *et al.* [17] | Unpublished decision impact study of Oncotype DX in women with LN+ early breast cancer. These data were submitted to NICE as part of the Exact Sciences CS and the submission from the Peony Breast Cancer Unit. |
| Chemotherapy benefit | RS 0-25: RxPONDER, CET arm[28]<br>RS >25: HR from SWOG-8814[31] applied to ET risk from TransATAC[19] | The inclusion of the HR from SWOG-8814[31] for the RS >25 group indirectly introduces an assumption of predictive benefit for Oncotype DX. |
| Death risk with DM | Abemaciclib plus fulvestrant arm of MONARCH2 trial[104] | Not fully consistent with model assumptions about treatments for DM, whereby only 65% of patients receive a CDK4/6 inhibitor as first-line therapy. |
| Probability of AML | Petrelli *et al.*[105] | The Exact Sciences CS[22] justifies the use of Petrelli *et al.*[105] on the basis that it is more recent than Wolff *et al.*[106] However, Wolff *et al.* is more recent than Petrelli *et al.* |
| Death risk with AML | NICE TA552[107] (liposomal cytarabine–daunorubicin for untreated AML) | Reflects a more recent source than that used in the DG34 model. The use of the median OS underestimates mean OS. |
| Probability of LR | De Bock *et al.*[108] | Same as DG34 model. |
| ***HRQoL parameters*** | | |
| Utility, recurrence-free | Lidgren *et al.*[109] | Same as DG34 model. |
| Utility, DM | Lidgren *et al.*[109] | Same as DG34 model. |
| Utility, AML | NICE TA552[107] | Consistent with source of modelled AML mortality risk. |
| QALY loss, chemotherapy | Campbell *et al.*[110] | Same as DG34 model. |
| QALY loss, LR | Campbell *et al.*[110] | Same as DG34 model. |
| Utility age adjustment | Ara and Brazier[102] (banded estimates) | Values reported by Hernández-Alava *et al.*[111] are more up-to-date. The use of age bands is unnecessary. |
| ***Resource use and cost parameters*** | | |
| Tumour profiling test costs | NICE DG34[13] (list prices) | Price discounts for other tumour profiling tests are not known to the company. |
| Adjuvant chemotherapy regimens used and associated resource use | Clinical opinion. Costing approach includes acquisition, administration and supportive medications. | The EAG's clinical advisors agree that the assumed distribution of chemotherapy regimens generally reflects current practice but noted that there is an increasing shift away from anthracycline-based regimens in certain patient groups. |
| ET usage | Ward *et al.*[103] | Based on DG34 model. |

| Parameter group | Source | EAG comments |
|---|---|---|
| Treatments for DM | Clinical opinion, Kurosky *et al.*[112] and MONARCH2 trial[104] | The model assumes that 65% of women with DM will receive a CDK4/6i as first-line treatment. The EAG's clinical advisors commented that CDK4/6i treatment would now be offered as first-line therapy for the vast majority of women with distant recurrence. |
| Cost AML (initial one-off cost and ongoing cyclical cost) | Zeidan *et al.*[113] | The model applies a once-only cost of intensive therapy in the first 6 months after diagnosis and an ongoing 6-monthly cost to reflect the cost of BSC for patients surviving beyond the initial 6 months. |
| AE frequency | TACT trial[114] | The model applies the frequency of AEs associated with FEC-D to anthracycline-taxane combinations and the costs of FEC60 to all other regimens. |
| Unit costs | NHS Reference Costs,[115, 116] eMIT,[117] BNF,[118] PSSRU[119] | Appropriate sources applied |
| Cost of death | Hinde *et al.*[91] | - |

*EAG - External Assessment Group; CS - company's submission; RS - recurrence score; DRFI - distant recurrence-free interval; OS - overall survival; DM - distant metastases; LR - local recurrence; ET - endocrine therapy; DG - Diagnostics Guidance, HRQoL - health-related quality of life; LN - lymph node; CET - chemotherapy and endocrine therapy; BSC - best supportive care; HR - hazard ratio; SWOG - Southwest Oncology Group; AML - acute myeloid leukaemia; CDK4/6i - cyclin dependent kinase 4/6 inhibitor; NICE - National Institute for Health and Care Excellence; TA - Technology Appraisal; AE - adverse event; eMIT - electronic Market Information Tool; NHS - National Health Service; BNF - British National Formulary; PSSRU - Personal Social Services Research Unit; FEC - fluorouracil, epirubicin, and cyclophosphamide; FEC-D - fluorouracil, epirubicin, and cyclophosphamide followed by docetaxel*

*Model evaluation methods*

The Exact Sciences CS[22] presents base case results for the overall LN+ population and for the pre-menopausal and post-menopausal LN+ subgroups using both the probabilistic and deterministic versions of the model. The CS also presents the results of the probabilistic sensitivity analysis (PSA) using cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs). The results of deterministic sensitivity analyses (DSAs) are presented using tornado diagrams and in tabular form. In addition, the CS presents the results of a number of scenario analyses exploring the use of alternative evidence sources and assumptions; these are presented separately for the overall LN+ population and for the pre- and post-menopausal LN+ subgroups. The CS also presents the results of exploratory analyses in the overall LN+ population in terms of deterministic pairwise ICERs for Oncotype DX, MammaPrint, EPclin and Prosigna versus clinical-pathological tools alone; key sources used to inform these exploratory analyses are shown in Table 19.

*Results of the Exact Sciences model*

The results presented in the Exact Sciences CS[22] are summarised in Table 21. Overall, the model suggests that Oncotype DX dominates clinical-pathological tools alone in the overall LN+ population and in the post-menopausal subgroup, but is dominated by clinical-pathological tools alone in the pre-menopausal subgroup. The exploratory comparisons of other tests suggest that the ICER for EPclin versus clinical-pathological tools alone is £9,355 per QALY gained, whereas the ICERs for Prosigna and MammaPrint versus clinical-pathological tools alone are substantially higher, at £41,773 and £50,626 per QALY gained, respectively.

**Table 21: Summary of cost-effectiveness results presented in the Exact Sciences CS (based on the company's revised model provided as part of their clarification response)**

| Analysis type | Base case, overall LN+ population | Base case, LN+ pre-menopausal subgroup | Base case, LN+ post-menopausal subgroup |
|---|---|---|---|
| Deterministic ICERs | Oncotype DX dominates clinical-pathological tools alone | Oncotype DX is dominated by clinical-pathological tools alone | Oncotype DX dominates clinical-pathological tools alone |
| Probabilistic ICERs | Oncotype DX dominates clinical-pathological tools alone | Oncotype DX is dominated by clinical-pathological tools alone | Oncotype DX dominates clinical-pathological tools alone |
| Probability test is cost-effective at WTP= £20,000/QALY gained | Probability = 0.93 | Probability = 0.07 | Probability = 1.00 |
| DSAs | Assuming WTP=£20,000/QALY, NMB is positive for all DSAs except lower value of HR from RxPONDER | Assuming WTP=£20,000/QALY, NMB is negative for all DSAs except upper value of HR from RxPONDER | Assuming WTP=£20,000/QALY, NMB is positive for all DSAs |
| Additional scenario analyses | Oncotype DX dominates clinical-pathological tools alone in all scenarios tested | Oncotype DX is dominated by clinical-pathological tools alone in all scenarios tested | Oncotype DX dominates clinical-pathological tools alone in all scenarios tested |
| Exploratory analyses of other tests versus clinical-pathological tools alone | Oncotype DX dominates clinical-pathological tools alone.<br><br>MammaPrint: ICER versus clinical-pathological tools alone = £50,626 per QALY gained<br><br>EPclin: ICER versus clinical-pathological tools alone = £9,355 per QALY gained<br><br>Prosigna: ICER versus clinical-pathological tools alone = £41,773 per QALY gained | Not presented | Not presented |

*CS - company's submission; LN - lymph node; ICER - incremental cost-effectiveness ratio; WTP - willingness-to-pay; NMB - net monetary benefit; DSA - deterministic sensitivity analysis; QALY - quality-adjusted life year; HR - hazard ratio*

4.2.1.2 EAG critique of the Exact Sciences model

The EAG's main concerns regarding the Exact Sciences model are summarised in Box 2. These concerns are discussed in detail in the subsequent sections.

**Box 1: Summary of the EAG's main concerns regarding the Exact Sciences model**

(1) Uncertainty surrounding the predictive benefit of Oncotype DX
(2) Analyses presented for the overall LN+ population mask the cost-ineffectiveness of Oncotype DX in the pre-menopausal subgroup
(3) Uncertainty around the probability of being in the Oncotype DX RS >25 group
(4) Uncertainty around relevant cut-offs for NICE decision-making
(5) Model errors and other minor implementation issues

*(1) Uncertainty surrounding the predictive benefit of Oncotype DX*

All of the economic analyses of Oncotype DX presented in the Exact Sciences CS[22] are informed by RxPONDER[76] for patients with an RS of 0-25 and by TransATAC[19] and SWOG-8814[31] for patients with an RS of >25. The use of HRs for the effect of CET versus ET alone which are drawn from separate studies for different genomic risk groups indirectly introduces an assumption that Oncotype DX is predictive of chemotherapy benefit. This assumption applies to all three of the company's base case analyses. As noted in Section 3.5.8, there remains some uncertainty around the predictive benefit of Oncotype DX: RxPONDER provides no information about the benefit of chemotherapy in women with an RS of >25 and did not demonstrate a predictive effect in women with an RS below this cut-off, whereas the interaction tests for chemotherapy effect and risk group in SWOG-8814 were statistically significant in some analyses, but not others. The EAG considers that it would have been useful to explore whether the model results are sensitive to this assumption, for example, through consideration of risk classification probabilities and DRFI estimates from TransATAC study.[19] This type of analysis is presented by Berdunov *et al.*[90] but is not included in the Exact Sciences CS;[22] Berdunov *et al.* reported that Oncotype DX was dominated by clinical-pathological tools alone when the assumption of predictive effect was removed from the model.

*(2) Analyses presented for the overall LN+ population mask the cost-ineffectiveness of Oncotype DX in the pre-menopausal subgroup*

The base case results for the overall LN+ population suggest that Oncotype DX dominates current decision-making using clinical-pathological tools alone (see Table 21). However, within the pre-menopausal subgroup the company's model suggests the opposite conclusion, as Oncotype DX is dominated by clinical-pathological tools alone. This is largely a consequence of the favourable HR for CET versus ET alone applied to patients with an RS of 0-25 in the pre-menopausal subgroup and the unfavourable HR for CET versus ET applied to patients with an RS of 0-25 in the post-menopausal

subgroup. The cost-ineffectiveness of Oncotype DX in the pre-menopausal LN+ subgroup is masked within the company's analysis of the overall LN+ population. As such, the EAG believes that it is appropriate to focus on the pre- and post-menopausal subgroup analyses separately.

The EAG notes that the Exact Sciences CS[22] (page 26) states that *"For selected premenopausal patients with N1 breast cancer, a low RS result (defined based on the clinical judgement of a multi-disciplinary team) may be valuable to guide the decision for hormonal treatment including potentially ovarian function suppression in place of adjuvant chemotherapy. The RS result may also help some premenopausal women with comorbidities which affect their suitability for chemotherapy treatment to decide between CET or ET (potentially with ovarian function suppression), based on their individual risk estimate."* In response to a request for clarification from the EAG, the company stated that their analysis may not have captured the full value of the Oncotype DX test for the subset of younger women who may prefer to avoid the harmful effects of chemotherapy, including permanent effects on reproductive health. The company also stated that clinicians recognise the value of the information provided by the Oncotype DX test in order to make better decisions for adjuvant treatment in this subgroup. For these reasons, the company advised caution in interpreting the results of their economic analysis in the pre-menopausal LN+ subgroup.

The EAG's clinical advisors commented that based on the findings of RxPONDER,[28] they considered that the use of Oncotype DX in pre-menopausal women may provide additional clinical information on the individual patient's risk of breast cancer recurrence, but commented that this would not influence their decision-making on whether to recommend chemotherapy. Overall, the EAG notes that based on the findings of RxPONDER,[28] the clinical value of Oncotype DX appears to be considerably stronger in post-menopausal women, and based on the Exact Sciences model, Oncotype DX appears to represent an inefficient use of NHS resources in pre-menopausal women.

*(3) Uncertainty around the probability of being in the Oncotype DX RS >25 group*

The Exact Sciences model assumes that 11% of women in the target population will have an Oncotype RS of >25. This calculation is based on the number of women who were screened for eligibility for entry into RxPONDER who had an RS >25 as the numerator (N=1,035) and the overall number of women who were registered for screening in RxPONDER as the denominator (N=9,383).[28] Kalinsky *et al.*[28] report that a total of 4,300 women were excluded from RxPONDER and that these women were excluded for various reasons: ineligible (N=164); no RS (N=84); had RS >26 (N=1,035); declined to participate (N=2,372); had recurrence (N=23) or had other or unknown reason (N=622). It is likely that some of the 3,265 women who were excluded for other reasons would actually have had an RS of >25. The EAG believes that the denominator for this calculation should reflect those women who were eligible for the trial and for whom a known Oncotype DX test result was available. As such, the company's model likely underestimates the probability that a woman will have an RS >25. The EAG

believes that within the RxPONDER trial, the proportion of women who had an RS of >25 lies somewhere between a minimum value of 0.11 (assuming that all other excluded patients have an RS of <25 [1035/9383]) and a maximum value of 0.17 (including only patients with a known RS in the calculation [1035/6118]). This range may vary across study populations.

In response to a request for clarification from the EAG,[101] the company highlighted that RxPONDER is an independently conducted study and that the company only had access to the information provided in the study protocol and the trial publication.[28] The company's response also highlights that the proportion of women with an RS >25 in RxPONDER is broadly similar to that in Stemmer *et al.*[120] and SEER.[78]

*(4) Uncertainty around relevant cut-offs for NICE decision-making*

The Exact Sciences CS[22] does not clearly state how the company intends the Oncotype DX test results to be used in clinical practice. Page 19 of the CS states that *"The RS result is typically defined as low (0 to 25) or high (26-100) however clinicians may apply different thresholds based on their interpretation of the evidence."* This suggests that Oncotype DX would be used as a two-level test (giving results as either low- or high-risk) based on an RS cut-off of 25. However, Table 2B of the Instructions For Use document for Oncotype DX[121] refers to chemotherapy benefit by RS in LN+ patients based on three levels: low - RS 0-17; intermediate - RS 18-30, and high - RS 31-100. The company's economic model is based specifically on the cut-offs applied in RxPONDER[28] (RS 0-25 and RS >25). This is the only scenario in which the cost-effectiveness of Oncotype DX has been evaluated within the CS.

In response to a request for clarification from the EAG,[101] the company stated that *"the validated RS result cut-offs used in the RxPONDER study should be used to categorise patients according to their risk of distant recurrence"* and that the Instructions For Use document will be updated to reflect this. At the time of writing, this document had not been updated and so there remains some uncertainty around the most relevant cut-offs for NICE decision-making.

*(5) Model errors and other minor implementation issues*

The EAG double-programmed the deterministic version of the Exact Sciences model to verify its implementation. The EAG was able to generate almost identical estimates of life years gained (LYGs), QALYs and costs for both the Oncotype DX and the comparator group in the LN+ population. Overall, the EAG considers the Exact Sciences model to be well programmed and free from major errors. During this double-programming exercise, the EAG identified the following minor issues:

(a) The Exact Sciences CS[22] states that the model assumes that women receive ET for 5 years. However, the executable model assumes that women receive ET for 5.5 years. In practice, many women will receive extended ET for a longer time period.
(b) The model includes a QALY loss due to LR. However, in the model, the disutility value for LR is applied to all women with DM, rather those women with DM who also develop LR.
(c) The model applies a cost of LR of £23,099. This is substantially higher than the cost of LR reported in the original source (Karnon *et al.*[122]) even when uplifted to current values. It is unclear which cost estimate from Karnon *et al.* has been used in the company's model prior to uplifting.
(d) The 6-month probability of death with AML is based on a median overall survival estimate of 9.6 months (based on the liposomal cytarabine–daunorubicin arm of Study 301[123]). However, the survival distribution is skewed and the mean survival estimate will be higher than the median value. The EAG believes it would be more appropriate to use the mean survival estimate to estimate mortality risk in each model cycle.
(e) The cost estimates for adjuvant TAC (docetaxel, doxorubicin and cyclophosphamide) assume that doxorubicin is given at a dose of 500mg/m$^2$. The EAG's clinical advisors commented that doxorubicin should have been assumed to be given at a dose of 50mg/m$^2$.
(f) The cost estimates for adjuvant EC90 (epirubicin plus cyclophosphamide) apply the cost of a first intravenous (IV) administration in every chemotherapy cycle, rather than applying the subsequent IV administration costs after the first treatment cycle. This results in the overestimation of administration costs for this regimen.
(g) The model includes age-adjusted by age band using Ara and Brazier.[102] The EAG believes that it would be preferable to adjust utility values for each individual age using more recent Euroqol 5-Dimensions (EQ-5D) values reported by Hernández-Alava *et al.*[111]
(h) The exploratory analyses which compare EPclin and Prosigna against clinical-pathological tools alone apply an HR of 0.76, taken from Harnan *et al.,*[10] which in turn, was estimated from the EBCTCG meta-analysis.[15] This value is a 10-year relative risk (RR); the estimated HR based on the same annual event rate data used to estimate the RR is approximately 0.71.
(i) In the pre-menopausal LN+ subgroup analysis, the model assumes a start age of 43 years and a time horizon of 45 years. By the final model cycle (at age 88 years), around 30% of the modelled population is still alive. The EAG believes that a lifetime horizon should have been applied.

#### 4.2.1.3 Additional analyses undertaken using the Exact Sciences model

As part of their response to clarification questions from the EAG,[101] the company provided an updated version of the model which addresses errors (a) and (b) listed in Section 5.2.1.2. The company's written response[101] noted that the presence of these errors had a negligible impact on the model results. The EAG further amended the company's revised model to also address issues (c), (d), (e), (g) and (i). Issue

(f) could not be easily resolved in the company's existing model structure and issue (h) was not resolved as it applies only to the company's exploratory analyses of other tumour profiling tests. The inclusion of these model amendments by the EAG had only a small impact on the results and did not affect the company's original base case economic conclusions, with Oncotype DX dominating clinical-pathological tools alone in the overall LN+ population and in the post-menopausal LN+ subgroup, and clinical-pathological tools alone dominating Oncotype DX in the pre-menopausal LN+ subgroup.

### *4.2.2 Agendia model summary and critique (MammaPrint)*

#### 4.2.2.1 Summary of economic analysis submitted by Agendia

In May 2023, Agendia submitted an executable model and an accompanying written document prepared by Cytel which details the methods and results of the model (hereafter referred to as the Cytel CEA report[83]). The company also provided responses to clarification questions from the EAG in May and June 2023,[124] which included an updated version of the economic model. The Cytel CEA report (page 5) states that the objective of the report is *"To evaluate the cost-effectiveness of the MammaPrint test compared to other tumour profiling tests (Oncotype DX, Prosigna and EPclin) as well as clinical risk tools (NPI and mAOL) to guide the use of adjuvant chemotherapy in ER+/HER2- early breast cancer patients."* The Cytel CEA report (page 30) also states *"Agendia is seeking reimbursement consideration to include LN+ [1-3, nodes] patients."*

The Cytel CEA report[83] presents cost-effectiveness estimates for MammaPrint in terms of the incremental cost per QALY gained compared with three other tumour profiling tests and usual care (current decision-making using clinical-pathological tools alone) from the perspective of the UK NHS and PSS over a 45.5 year (lifetime) horizon. The model includes a cycle length of 6 months and includes half-cycle correction. Health outcomes and costs are discounted at 3.5% per annum. Costs are valued at 2021/22 prices, including uplifting of unit cost estimates using Hospital and Community Health Services (HCHS) indices and NHS Cost Inflation Indices (NHSCII) where necessary.

The Cytel CEA report[83] presents a base case analysis and five additional scenario analyses. The modelled population and the comparators under consideration differ between these analyses (see Table 22). The base case analysis compares the use of MammaPrint versus Oncotype DX, EPclin, Prosigna and usual care in a population of women who are ER+ and HER2-, who have either LN0 or LN+ (1-3 nodes), and who are considered to be clinical high-risk according to NPI or mAOL. In the base case analysis, at model entry, the population is assumed to be aged 58.9 years based on the NHS England (NHSE) Access Scheme Dataset, which reflects LN0 patients who received Oncotype DX testing following DG10.[125] The modelled population is intended to reflect both pre- and post-menopausal women, although post-menopausal women are considered as a specific subgroup in Scenario 4 (see Table 22). Men with early breast cancer are not reflected in the modelled population.

The general structure of the Agendia model is similar to the model used to inform DG34.[13] The model uses a hybrid approach and is comprised of an initial decision tree component which stratifies patients according to their genomic risk based on the tumour profiling test, followed by a Markov component which estimates long-term outcomes and costs conditional on genomic risk and whether the patient receives adjuvant CET or ET alone. The decision tree component of the model stratifies patients into either high- or low-risk for 2-level tests, or high-, intermediate- or low-risk for 3-level tests, and determines whether the patient receives adjuvant chemotherapy. The long-term Markov model includes four health states: (i) recurrence free; (ii) DM; (iii) AML and (iv) dead. The model also includes separate tunnel states to reflect the impact of LR (prior to DM), which is assumed to impact on QALYs and costs, but does not affect the patient's underlying health state or mortality risk. The benefit of chemotherapy is modelled using RRs applied to the risk of DM with ET alone. Within the base case analysis and Scenarios 1-4, the company's model assumes that MammaPrint is predictive of chemotherapy benefit, with RRs for DM for CET versus ET alone of 1.0 and 0.38 applied to the genomic low-risk and genomic high-risk groups, respectively. In Scenario 5, which reflects a pure LN+ population, RRs of 0.97 and 0.28 are applied in the genomic low-risk and high-risk groups. All other tests are assumed to be prognostic only, except for Oncotype DX in Scenario 4 (see Table 22). This is a key assumption which favours MammaPrint over all of the other comparator tests and usual care.

QALYs are modelled as a function of whether patients receive adjuvant chemotherapy and the long-term trajectory of patients through the Markov model health states. Lower utility values are applied to patients receiving CET versus ET alone for 3 years, which are intended to represent the disutility resulting from toxicity associated with adjuvant chemotherapy (net loss per patient treated with CET versus no ET alone = 0.29 QALYs). The model applies comparatively lower utility values to the DM and AML states than the recurrence-free state. The model also includes a disutility value associated with LR of -0.11.[110] The model includes age-adjustment of utility values based on Ara and Brazier.[102]

The model includes resource costs associated with:

- The tumour profiling tests
- Adjuvant chemotherapy and supportive medications (applied in the first 6 months only)
- Management of chemotherapy-related AEs
- ET (acquisition and administration costs, for up to 8.5 years)
- Bisphosphonates (zoledronic acid, acquisition and administration costs, for up to 4 years)
- Resource use whilst patients are receiving chemotherapy (applied in the first 6 months only)
- Additional resource use whilst patients remain recurrence-free (up to 3 years)
- Treatments for locoregional recurrence (costed per local / contralateral recurrence event)
- Treatments for DM
- Treatments for AML
- End of life care.

**Table 22: Summary of economic analyses presented in the Cytel CEA report on MammaPrint**

| Analysis | Population | Intervention and comparators | Key sources of DRFI risk and chemotherapy benefit | Chemotherapy benefit assumptions | Additional EAG comments |
|---|---|---|---|---|---|
| Base case. Clinical high-risk patients (ER+, HER2-). | Clinical high-risk. LN0 NPI>3.4 and LN+ patients weighted in blended analysis. | • MammaPrint<br>• Oncotype DX<br>• Prosigna<br>• EPclin<br>• Usual care | TransATAC,[19] MINDACT,[29] EBCTCG[15] | Predictive benefit included for MammaPrint. All other tests assumed to be prognostic only. | Assumes mAOL high-risk is equivalent to NPI>3.4.<br>Analysis includes a minority of LN+ patients. |
| Scenario 1. Full ER+, HER2- population stratified by 2-level clinical test. | Clinical low-risk and clinical high-risk patients. LN0 NPI>3.4 and LN+ patients weighted in blended analysis. | Same as Agendia base case | | | Clinical low-risk patients included in model but only clinical high-risk patients get the genomic test.<br>Analysis includes a minority of LN+ patients. |
| Scenario 2. Full ER+, HER2- population – stratified by 3-level clinical test. | Clinical low-risk and clinical high-risk patients. LN0 NPI>3.4 and LN+ patients weighted in blended analysis. | Same as Agendia base case | | | Analysis partitions population into clinical low-risk, clinical high-risk and LN+, but only clinical high-risk patients receive the genomic test.<br>Analysis includes a minority of LN+ patients. |
| Scenario 3. ER+, HER2- post-menopausal women stratified by 2-level clinical test. | Post-menopausal clinical low-risk and clinical high-risk patients. LN0 NPI>3.4 and LN+ patients weighted in blended analysis. | Same as Agendia base case | | | Clinical low-risk patients included, but only clinical high-risk patients receive the genomic test.<br>Analysis includes a minority of LN+ patients. |
| Scenario 4. TAILORx clinical study stratified by 2-level clinical test. | Clinical low-risk and clinical high-risk patients. LN0 patients only. | • MammaPrint<br>• Oncotype DX<br>• Usual care. | TAILORx[126] MINDACT[29] | Predictive benefit included for MammaPrint and Oncotype DX. | Clinical low-risk patients included, but only clinical high-risk patients receive the genomic test.<br>Analysis excludes LN+ patients. |
| Scenario 5. ER+, HER2-, LN+ subgroup. | Clinical high risk. LN+ patients only. | • Same as base case scenario, but restricted to LN+ subgroups from MINDACT[29] and TransATAC[19] | | | This is the only analysis which directly addresses the decision problem set out in the NICE scope.[11] |

*DRFI - distant recurrence-free interval; EAG - External Assessment Group; ER - oestrogen receptor; HER2 - human epidermal growth receptor 2; mAOL - modified Adjuvant! Online; EBCTCG - Early Breast Cancer Trialists' Collaborative Group; LN - lymph node; NPI - Nottingham Prognostic Index; NICE - National Institute for Health and Care Excellence*

*Key assumptions applied in the Agendia base case analysis*

The Agendia model makes the following key assumptions:

- MammaPrint is predictive of chemotherapy benefit. All other comparator tests have prognostic benefit only (a predictive effect of Oncotype DX is included in Scenario 4).
- MammaPrint would be used only in patients who are clinical high-risk.
- Other tests (EPclin, Prosigna and Oncotype DX) are included in the analysis based on the assumption that NPI>3.4 in TransATAC[19] is equivalent to mAOL high-risk in MINDACT.[29]
- In the absence of tumour profile testing, most patients (~79%) will receive adjuvant chemotherapy.
- Chemotherapy-related AEs impact on HRQoL for 3 years.
- The baseline risk of distant recurrence of breast cancer is reduced by 50% at 10 years and by 100% at 15 years.
- Once patients develop AML, their AML determines their prognosis regardless of prior history of distant recurrence of breast cancer.
- Tests with the same number of levels are not interpreted in the same way (e.g., the probability that a patient with a low-risk result from MammaPrint receives adjuvant chemotherapy differs from that for a patient with a low-risk result from EPclin).
- The modelled population is intended to reflect both pre-menopausal and post-menopausal women.

*Evidence sources used to inform the Agendia model*

The evidence sources used to inform the parameters of the Agendia model are summarised in Table 23, together with brief comments from the EAG.

**Table 23: Key evidence sources used to inform Agendia model - base case analysis and Scenario 5 (pure LN+ subgroup)**

| Parameter group | Source | EAG comments |
|---|---|---|
| ***Clinical parameters*** | | |
| Start age | NHSE Access Dataset[127] | Reflects patients with LN0 disease. |
| Test risk classification probabilities | MINDACT,[29] TransATAC[19] | Analyses between MammaPrint and other tests assume equivalence of populations enrolled in TransATAC[19] and MINDACT.[29]<br><br>MammaPrint data are based on the HR+/HER2- population of MINDACT (LN0 and/or LN+)[29] |
| DRFI probabilities | MINDACT,[29] TransATAC[19] | DRFI For MammaPrint low-risk ET and MammaPrint high-risk chemotherapy plus ET groups are based on re-analyses of MINDACT IPD.[83] |
| Risk tapering | Ward *et al.*[103] | Same as DG34 model. |

| Parameter group | Source | EAG comments |
|---|---|---|
| Chemotherapy probability under usual care | NCRAS[16] and expert opinion[83] | The Cytel CEA report[83] states that *"there are no empirical evidence sources which provide estimates of baseline chemotherapy use for patients who are mAOL high-risk or mAOL low risk"* hence the need to rely on expert opinion. |
| Chemotherapy probability conditional on genomic tests | MammaPrint – Kuijer *et al*.[128]<br>Oncotype DX – Crolley *et al.*[129]<br>Prosigna – UKBCG survey (3-level)[10]<br>EPclin – UKBCG survey (2-level)[10] | Use of different sources for each test implicitly assumes that 2-level tests (MammaPrint and EPclin) are interpreted differently and that 3-level tests (Oncotype DX and Prosigna) are interpreted differently.<br><br>Crolley *et al.,*[129] was undertaken in a purely LN0 population and Kuijer *et al.*[128] was undertaken in women without axillary lymph node involvement (pN0 or pN1mi). |
| Chemotherapy benefit for MammaPrint | Assumptions of predictive benefit based on interpretation of MINDACT data[29] | Assumes predictive benefit – RRs of 1.00 and 0.38 are applied to patients receiving CET in the MammaPrint low- and high-risk patients, respectively. RRs of 0.97 and 0.28 applied in Scenario 5 (LN+ subgroup). |
| Chemotherapy benefit for other tests | EBCTCG meta-analysis[15] | An RR of 0.76 is applied to all patients receiving CET in the Oncotype DX, Prosigna, EPclin and usual care comparator groups. |
| Death risk with DM | Wang *et al.*[130] | This dataset reflects patients diagnosed between 2010 and 2015 and therefore is unlikely to reflect survival for patients receiving current first-line treatments for DM (e.g., CDK4/6i therapies). |
| Probability of AML | Wolff *et al.*[106] | Same as DG34 model. |
| Death risk with AML | Edlin *et al.*[131] | Same as DG34 model. |
| LR transition probabilities | Geurts *et al.*[132] | This aspect of the model was not included in the DG34 model. LR is assumed to impact only on QALYs and costs without affecting the patient's underlying health state or survival. |
| ***HRQoL parameters*** | | |
| Utility values, recurrence-free with adjuvant CET or ET alone | Lidgren *et al*.[109] | Disutility value applied for patients receiving CET for 3 years. The source of assumption about duration of disutility is unclear. |
| Utility value, DM | Lidgren *et al*.[109] | Same as DG34 model. |
| Utility value, AML | Younis *et al*.[133] | Same as DG34 model. |
| Disutility, LR | Campbell *et al*.[110] | Same as DG34 model. |
| Utility age adjustment | Ara and Brazier[102] | Values reported by Hernández-Alava *et al.*[111] are more up-to-date. |
| ***Resource use and cost parameters*** | | |
| Tumour profiling test costs | NICE DG34[11] (list prices) | Price discounts for other tests are not known to the company. Test prices are uplifted using inflation indices. |
| Adjuvant chemotherapy regimens used and associated resource use | Clinical opinion. Structure of costing approach based on Hall *et al.*[93] | Updated from DG34 model. |
| ET usage | Ward *et al.*[103] | Same as DG34 model. |
| Bisphosphonates | Ward *et al.*[103] | Similar to DG34 model. |

| Parameter group | Source | EAG comments |
|---|---|---|
| Cost DM | Thomas *et al.*[134] | Same as DG34 model. |
| Cost AML | Russell-Smith *et al.*[135] | The population of the model reported by Russell-Smith *et al.* relates to patients with *de novo* AML which is not therapy-related. |
| AE frequency | Based on various trials[136-140] | Values used in the company's original model are unclear and appear highly inflated. These values were amended in the company's revised model. |
| Unit costs | NHS Reference Costs,[141] eMIT,[142] MIMS,[143] PSSRU[119] | Appropriate sources applied. |
| Cost of death | Georgiou and Bardsley[144] | - |

*EAG - External Assessment Group; NHSE - National Health Service England; LN - lymph node; HR – hormone receptor; HER2 - human epidermal growth factor receptor; CDK4/6i - cyclin dependent kinase 4/6 inhibitor; DRFI - distant recurrence-free interval; HRQoL - health-related quality of life; NCRAS - National Cancer Registration and Analysis Service; DG - Diagnostics Guidance; ET - endocrine therapy; CET - chemotherapy plus endocrine therapy; IPD - individual patient data; LR - local recurrence; UKBCG - UK Breast Cancer Group; RR - relative risk: AML - acute myeloid leukaemia; DM - distant metastases; AE - adverse event; NHS - National Health Service; BNF - British National Formulary; eMIT - electronic Market Information Tool; MIMS - Monthly Index of Medical Specialities; PSSRU - Personal Social Services Research Unit; pN0 - lymph node negative (pathological); pN1(mi) - lymph node negative with micrometastases (pathological); NPI - Nottingham Prognostic Index; mAOL - modified Adjuvant! Online; CEA - cost-effectiveness analysis*

*Model evaluation methods*

The headline results of the company's model are presented in terms of ICERs based on the deterministic version of the model. The Cytel CEA report[83] also presents PSA results for the base case scenario; the results of these analyses are reported in terms of probabilistic ICERs, cost-effectiveness planes and CEACs. The Cytel CEA report also presents the results of DSAs in the form of tornado diagrams as well as a number of deterministic scenario analyses which explore the impact of alternative assumptions around: the time horizon; discount rates; alternative clinical model parameters; chemotherapy probabilities; utility values and costs. The report presents the results of these uncertainty analyses only for the base case scenario.

*Model results*

The Cytel CEA report[83] presents the results of a large number of analyses. For brevity, these are summarised in Table 24. Across the base case and scenario analyses, the Agendia model suggests that MammaPrint dominates all other tumour profiling tests and usual care.

**Table 24: Summary of cost-effectiveness results presented in the Cytel CEA report (includes company's correction of errors at the clarification stage)**

| Analysis type | Base case - Clinical high-risk patients | Scenario 1 - Full population stratified by 2-level clinical test | Scenario 2 - Full population – stratified by 3-level clinical test | Scenario 3 - Post-menopausal women stratified by 2-level clinical test | Scenario 4 - TAILORx clinical study stratified by 2-level clinical test | Scenario 5 – LN+ subgroup |
|---|---|---|---|---|---|---|
| Deterministic ICERs | MammaPrint dominates all comparators | MammaPrint dominates all comparators | MammaPrint dominates all comparators | MammaPrint dominates all comparators | MammaPrint dominates Oncotype DX and usual care | MammaPrint dominates all comparators |
| Probabilistic ICERs | MammaPrint dominates all comparators | Not presented | Not presented | Not presented | Not presented | Not presented |
| Probability test is cost-effective at λ= £20,000 | Probability = 0.91 | Not presented | Not presented | Not presented | Not presented | Not presented |
| DSAs | Highest ICER for MammaPrint vs comparators across all analyses = £392 per QALY gained | Not presented | Not presented | Not presented | Not presented | Not presented |
| Additional scenario analyses | Highest ICER for MammaPrint vs comparators across all analyses = £3,647 per QALY gained | Not presented | Not presented | Not presented | Not presented | Not presented |

*CEA - cost-effectiveness analysis; ICER - incremental cost-effectiveness ratio; LN - lymph node; QALY - quality-adjusted life year; DSA - deterministic sensitivity analysis*

4.2.2.2 EAG critique of the Agendia model

The EAG's main concerns regarding the Agendia model are summarised in Box 2. These concerns are discussed in detail in the subsequent sections.

**Box 2: Summary of the EAG's main concerns regarding the Agendia model**

1. Relevance of base case analyses to the decision problem set out in the final NICE scope
2. Reliability of comparisons against other tumour profiling tests
3. Inappropriate assumption that some women in the MammaPrint group will not receive the test
4. Questionable assumptions around post-test chemotherapy probabilities
5. Questionable assumption of predictive benefit of chemotherapy for MammaPrint group
6. Concerns regarding HRQoL assumptions
7. Concerns regarding costs
8. Model errors

*(1) Relevance of the model population to the decision problem set out in the final NICE scope*

The final NICE scope[11] for this appraisal describes the target population as *"People with ER positive and/or PR positive, HER2 negative, early breast cancer with 1 to 3 positive lymph nodes, who are deciding whether to have adjuvant chemotherapy."* The Agendia base case analysis reflects a mixed population of women with either LN0 or LN+ early breast cancer. Clinical outcomes for the comparison of MammaPrint versus usual care are informed by data from the MINDACT trial.[29] Within this trial, 79.0% of patients were LN0 and 21.0% were LN+; the precise proportion of women who were LN+ in the HR+, HER2- population used in the analysis is unclear but is likely to be similar. With the exception of Scenario Analysis 5, which uses unpublished individual patient data (IPD) for the LN+ subgroup of women in MINDACT, the EAG considers the economic analysis presented by Agendia to be of limited relevance to the population under consideration within this appraisal.

*(2) Reliability of the comparisons against other tumour profiling tests*

The NICE scope[11] defines the comparator as *"Current decision making, which may include any tool, or clinical and pathological features, used to assess risk."* NICE DG34[13] did not make any specific recommendations on the use of tumour profiling tests within the LN+ population; hence, tumour profiling tests do not reflect current decision-making in the relevant population for this appraisal. The EAG believes that the comparisons against the other tumour profiling tests included in the Agendia model are problematic as they assume that the characteristics of the patient populations enrolled in TransATAC[19] and MINDACT[29] are identical with respect to prognostic factors and treatment effect modifiers. The EAG believes that it may be more appropriate to focus on the comparisons of MammaPrint versus usual decision-making, which do not necessitate the use of naïve indirect comparisons between the tests.

*(3) Inappropriate assumption that some women in the MammaPrint group will not receive the test*

Within the Scenario 5 (the pure LN+ subgroup), the Agendia model assumes that 16% of women in the MammaPrint group will not receive the MammaPrint test. These patients are instead assumed to accrue the outcomes and costs associated with the usual care (no testing) group. The Cytel CEA report[83] states that these women are clinically low-risk and that the company only intends MammaPrint to be used in women who are clinically high-risk. The EAG considers that the economic model should reflect the target clinical high-risk population only and that clinical low-risk patients who are not eligible for MammaPrint should be excluded from the model. The same issue also applies to Scenario 1-4, albeit with higher proportions of women (54% to 68%) not receiving the tumour profiling test.

*(4) Questionable assumptions around post-test chemotherapy probabilities*

The post-test probabilities of receiving chemotherapy in the Agendia model are summarised in Table 25. The base case analysis uses separate studies to estimate the probability of receiving chemotherapy conditional on genomic risk classification. This implies that different tests with the same number of levels will be interpreted differently by clinicians and will lead to different probabilities of patients receiving chemotherapy. For example, the probability that a patient who is low-risk according to the 2-level MammaPrint test goes on to receive adjuvant chemotherapy is assumed to be 0.05 (based on Kuijer *et al.*[128]), whereas the probability that a patient who is low-risk according to the 2-level EPclin test goes on to receive adjuvant chemotherapy is assumed to be 0.15 (based on the UK Breast Cancer Group [UKBCG] survey reported by Harnan *et al*.[10]). It is unclear whether clinicians would interpret the results of tumour profiling tests differently, but it may be the case that the differences included in the Agendia model reflect heterogeneity between the patient populations enrolled in the studies or differences in preferences for the use of adjuvant chemotherapy between the countries in which those studies were undertaken. The EAG also notes that Crolley *et al.,*[129] was undertaken in an LN0 population and Kuijer *et al.*[128] was undertaken in women without axillary lymph node involvement (pN0 or pN1mi). These studies are therefore not aligned with the population defined in the final NICE scope for this appraisal.[11]

**Table 25: Probability of receiving chemotherapy conditional on genomic risk classification applied in the Agendia model**

| Test | Low-risk | High-risk | Source |
|---|---|---|---|
| **2-level tests** | | | |
| MammaPrint | 0.05 | 0.97 | Kuijer *et al.*[128] |
| EPclin | 0.15 | 0.92 | UKBCG survey[10] |
| **3-level tests** | | | |
| Oncotype DX | 0.03 | 0.91 | Crolley *et al.*[129] |
| Prosigna | 0.04 | 0.92 | UKBCG survey[10] |

*UKBCG - UK Breast Cancer Group*

*(5) Highly questionable assumption of predictive benefit of chemotherapy for MammaPrint group*

The Agendia model assumes that MammaPrint is predictive of chemotherapy benefit, and that the other tumour profiling tests are prognostic only (with the exception of Oncotype DX in Scenario 4, see Table 22). In the base case scenario, the model applies an RR for distant recurrence for CET versus ET of 1.00 (i.e., no chemotherapy benefit) for patients who are MammaPrint low-risk and an RR of 0.38 (i.e., substantial chemotherapy benefit) for patients who are MammaPrint high-risk. In the pure LN+ subgroup (Scenario 5), RRs of 0.97 and 0.28 are applied. The company's justification for this assumption of predictive benefit is based on the non-significant HR obtained from an adjusted Cox model fitted to MINDACT IPD by the company to estimate the effect of CET versus ET alone in HR+, HER2- women (including both LN0 and LN+) who were clinical high-risk and genomic low-risk (HR 0.74, 95% CI 0.43 to 1.28).[124] The company's model assumes that because this HR is not statistically significant, it should be interpreted as being equivalent to chemotherapy having no effect (i.e., HR=1.00), and because chemotherapy is known to be clinically effective overall,[15] a considerably greater treatment effect must therefore apply to the clinical high-risk genomic high-risk group who were not randomised in MINDACT. The logic underpinning the company's calculations of chemotherapy treatment effects by genomic risk group in the base case analysis is as follows:

(a) Based on the EBCTCG meta-analysis,[15] an overall RR of 0.76 is expected across the full spectrum of clinical high-risk patients.
(b) Within the overall ER+, HER2- population, 61% of the clinical high-risk population is MammaPrint low-risk. These patients will obtain no benefit from adjuvant chemotherapy because the HR from the Cox model is not statistically significant. An RR of 1.00 is applied to these patients.
(c) Given points (a) and (b), the necessary RR in the remaining 39% of patients who are MammaPrint high-risk must therefore be 0.38 (i.e., 0.61 x 1.00 + 0.39 x 0.38 = 0.76).

The EAG notes that the updated MINDACT publication by Piccart *et al.*[29] reports an HR for DMFS for the overall clinical-high genomic-low risk group of 0.66 (95% CI 0.48 to 0.92) and an HR for DRFI in this same population of 0.66 (95% CI 0.46 to 0.95). After excluding the 21% of participants in this group that are not relevant to the decision problem, the HR for the HR+, HER2- population (regardless of nodal status) provided by the company was 0.74 (95% CI 0.43 to 1.28). The wider CI is expected due to the reduction in sample size and although no longer statistically significant, the point estimate of 0.74 is still the best estimate of the treatment effect for this group. Based on the Z-score (calculated on a log scale), this implies that 86% of individuals in this group are likely to have a beneficial response to chemotherapy (HR <1.0). If the estimated HR of 0.74 is applied directly (rather than assuming that it is 1.0), then the HR for the 39% of patients who are MammaPrint high-risk would be 0.79 (i.e., 0.61 x 0.74 + 0.39 x 0.79 = 0.76). This is similar to the HR for the MammaPrint low-risk group.

Piccart *et al.*[29] also reports an estimate of the relative treatment effect on DMFS specifically for the clinical high genomic low LN1-3 population, including those with HER2+ and ER- or PR- disease (HR 0.84, 95% CI 0.51 to 1.37; see Table 12). This analysis indicates a comparatively lesser effect of chemotherapy in women with LN1-3 which is again not statistically significant and the sample size is small (N=658). It is unclear why the point estimate of the effect of chemotherapy is less pronounced in this subgroup.

Overall, the EAG considers that there is insufficient evidence from the MINDACT trial to support the argument that MammaPrint is predictive of chemotherapy benefit, and there is no external evidence to inform the relative treatment effect of chemotherapy in the clinical high genomic high group. As such, the EAG believes that the company's interpretation of the results of their IPD analysis are flawed and that it is more reasonable to apply the same treatment effect estimate for chemotherapy to both the MammaPrint low-risk and MammaPrint high-risk groups.

*(6) Concerns regarding HRQoL assumptions*

The Cytel CEA report[83] (page 30) states that *"A disutility associated with short-term AEs related to adjuvant chemotherapy is applied once during the first model cycle only (whilst the patient is receiving treatment)."* This is not an accurate description of the assumptions employed in the Agendia model. Instead, the model applies treatment-specific utility values for patients receiving adjuvant CET and for patients receiving ET alone, based mostly on values reported by Lidgren *et al.*,[109] with higher utility scores applied to the ET group. These differences are assumed to persist for 3 years, although the source of the assumption of a 3-year disutility is unclear. The utility values applied for patients receiving CET or ET alone in the Agendia model are summarised in Table 26. The model suggests that a patient who is treated with CET who survives for three years without experiencing relapse will lose 0.29 QALYs compared with an equivalent patient who receives ET alone. This modelled QALY loss is substantially larger than the QALY losses applied in the majority of other economic models of tumour profiling tests in LN+ women included in the EAG's review and the Exact Sciences Model (see Section 4.1, Table 27 and Section 5.2.1.1). The only studies in which a similar or higher chemotherapy-related QALY loss is applied are those reported by Vanderlaan *et al.*[100] and Wong *et al.*[99] In both of these studies, QALY losses associated with adjuvant chemotherapy appear to be based on assumptions rather than empirical evidence. Overall, the EAG has concerns that the QALY loss associated with chemotherapy in the Agendia model has likely been overestimated. The EAG also notes that the Agendia model includes age-adjustment of utility values based on Ara and Brazier;[102] the EAG believes that it would be more appropriate to use more recent estimates by Hernández-Alava *et al.*[111]

**Table 26: Utility values associated with chemotherapy and endocrine therapy applied in the Agendia model**

| Model health state | Utility value | Source |
|---|---|---|
| Recurrence-free, CET, year 1 | 0.620 | Lidgren *et al.*[109] State P *"First year after primary breast cancer"*, receiving adjuvant chemotherapy |
| Recurrence-free, CET, year 2-3 | 0.743 | Value used in model not reported in Lidgren *et al.*[109] |
| Recurrence-free, CET, year 4+ | 0.824 | Lidgren *et al.*[109] State S *"Second and following years after primary breast cancer / recurrence"*, receiving ET |
| Recurrence-free, ET alone, year 1 | 0.744 | Lidgren *et al.*[109] State P *"First year after primary breast cancer"*, receiving ET |
| Recurrence-free, ET alone, year 2+ | 0.824 | Lidgren *et al.*[109] State S *"Second and following years after primary breast cancer / recurrence"*, receiving ET |

*ET - endocrine therapy; CET - chemotherapy plus endocrine therapy*

**Table 27: Adjuvant chemotherapy disutility values and QALY losses applied in the Agendia model and other models included in the EAG's systematic review**

| Model | Disutility / QALY loss associated with adjuvant chemotherapy toxicity per patient treated | Source of disutility value / QALY loss |
|---|---|---|
| Agendia model[83] | QALY loss of 0.29 | Lidgren *et al.*[109] |
| Harnan *et al.* (2019)[10] | QALY loss of 0.04 | Campbell *et al.*[110] |
| Berdunov *et al.* (2021)[90] | Disutility of -0.04, 1-year duration | Campbell *et al.*[110] |
| Blohmer *et al.* (2013)[96] | QALY loss of 0.07 | Peasgood *et al.*[145] |
| Hall *et al.* (2012)[98] | QALY loss of 0.08 | Lidgren *et al.*[109] |
| Hall *et al.* (2017)[93] | Disutility of -0.096, 1-year duration | Campbell *et al.*[110] |
| Hannouf *et al.* (2014)[95] | QALY loss in first year of approximately 0.025. Treatment-specific utility values favouring ET over chemotherapy are applied thereafter. | Assumptions |
| Hinde *et al.* (2019)[91] | QALY loss of 0.12 | Lidgren *et al.*[109] |
| Lamond *et al.* (2012)[97] | QALY loss appears to be around 0.06 (excluding the impact of CINV and FN) | Tufts,[146] Tengs,[147] Ward[103] |
| Masucci *et al.* (2019)[92] | QALY loss of 0.06 | Lidgren *et al.*[109] |
| Stein *et al.*(2016)[94] | Disutility of -0.096, 1-year duration | Lidgren *et al.*[109] |
| Vanderlaan *et al.* (2011)[100] | QALY loss of 0.50 over lifetime | Assumption |
| Wong *et al.* (2012)[99] | Disutility of -0.30, duration unclear | Assumption |

*QALY - quality-adjusted life year; CINV - chemotherapy-induced nausea and vomiting; FN - febrile neutropoenia; ET - endocrine therapy*

*(7) Concerns regarding costs*

The Agendia model includes the same estimates of mortality risk and HRQoL for AML as those used in the EAG's model developed to inform DG34.[10] However, the Agendia model applies a substantially higher lifetime cost estimate for AML compared with the DG34 model (Harnan *et al.*[10] cost = £10,600 versus Russell-Smith *et al.*[135] cost = £132,039). This cost estimate has been taken from an economic modelling study of gemtuzumab ozogamicin (GO) plus standard care (SC) chemotherapy in people with *de novo* CD33-positive AML. The population included in the analysis reported by Russell-Smith *et al.*

is not consistent with the patient population reflected in the Agendia model, as the latter has secondary (therapy-related) AML. In addition, as the AML lifetime cost estimate has been updated in the Agendia model but the mortality risk and health utility value have not, this implies that the treatment of AML has become substantially more expensive without any improvement in health outcomes. The company's approach favours all tumour profiling tests which reduce the incidence of AML as a consequence of fewer patients receiving chemotherapy. The EAG considers that it would be preferable for the model to reflect mortality risks, QALYs and costs associated with current therapies estimated in patients with secondary AML.

In addition, the company has applied inflation indices to uplift the prices of all four tumour profiling tests considered in the model. However, since DG34[13] was published in 2018, the list prices of EPclin and Oncotype DX have not changed and the marginal cost per Prosigna test has decreased (see Section 4.3, Table 38).

*(8) Model errors*

The EAG identified three sets of programming errors in the Agendia model:

(i) *AE frequencies.* The original version of the Agendia model included programming errors which led to implausibly high AE frequencies. This issue was resolved in the revised version of the model provided in the company's response to clarification questions from the EAG.[124]

(ii) *Half-cycle correction.* The half-cycle correction calculations count the costs and outcomes for the first cycle 1.5 times. This is an unequivocal error.

(iii) *Supportive care administration cost calculations.* The model formulae used to calculate supportive care administration costs (Model worksheet "Model Parameters", cells P519:P526) erroneously exclude dollar signs to anchor the cell references. Consequently, incorrect cell references are used in the calculations. This is also an unequivocal error.

4.2.2.3 Additional exploratory analysis undertaken by the EAG

The EAG undertook an additional analysis which attempts to address six of the issues identified in the EAG's critique:

(i) The analysis was restricted to the LN+ subgroup (Scenario 5) as this is consistent with the population listed in the scope of the appraisal.[11]

(ii) The assumption of predictive benefit for MammaPrint was removed from the model. An RR for the effect of chemotherapy on distant recurrence of 0.76 was applied to all patients regardless of their genomic risk, based on the EBCTCG meta-analysis.[15]

(iii) The utility value for women who remain recurrence-free was assumed to be equal to 0.824 after 1 year, regardless of whether they receive ET or CET. This means that the disutility value for chemotherapy-related toxicity is applied for a duration of 1 year only.

(iv) The errors in the formulae used to apply the half-cycle correction and the supportive care administration costs were rectified.

(v) The assumption that 16% of women in the MammaPrint group do not receive the MammaPrint test was removed.

(vi) Other comparator tests were excluded from the analysis.

The results of this analysis are shown in Table 28 together with the results of the company's deterministic base case analysis. The EAG's additional analysis suggests that MammaPrint leads to a small reduction in survival, a small increase in QALYs and a small decrease in costs; hence, MammaPrint remains dominant. However, the EAG remains concerned that there are still some minor errors in this model – the EAG ran a scenario in which no patients receive chemotherapy in either the MammaPrint or usual care groups and the model still suggests that MammaPrint generates additional QALYs over usual care – this clearly reflects an error. The EAG also notes that this model does not reflect all of the EAG's preferred assumptions and evidence sources (see Section 4.3). As such, the EAG believes that the results of this re-analysis should be interpreted with caution.

**Table 28: Results of additional analysis undertaken by the EAG**

| Option | LYGs* | QALYs | Costs | Inc. LYGs | Inc. QALYs | Inc. Costs | ICER |
|---|---|---|---|---|---|---|---|
| **Company's Scenario Analysis 5 (LN+), other tests excluded** | | | | | | | |
| MammaPrint | 23.45 | 11.57 | £23,327 | 1.43 | 0.70 | -£4,676 | **MammaPrint** |
| Usual care | 22.02 | 10.87 | £28,003 | - | - | - | **dominating** |
| **Company's Scenario Analysis 5 (LN+), other tests excluded, including EAG amendments** | | | | | | | |
| MammaPrint | 23.86 | 11.60 | £21,570 | -0.08 | 0.02 | -£26 | **MammaPrint** |
| Usual care | 23.93 | 11.59 | £21,596 | - | - | - | **dominating** |

*EAG – External Assessment Group; LYG – life year gained; QALY – quality-adjusted life year gained; ICER – incremental cost-effectiveness ratio; inc. – incremental; LN+ - lymph node positive*
** Undiscounted*

## 4.3 Independent EAG economic analysis

### *4.3.1 Scope of the EAG economic analysis*

The EAG developed a health economic model to assess the cost-effectiveness of Oncotype DX, Prosigna, EPclin and MammaPrint versus current decision-making. The scope of the EAG's model is summarised in Table 29. The model assesses health outcomes and costs associated with each tumour profiling test and current decision-making over a lifetime horizon (up to age 100 years) from the perspective of the NHS and PSS. All costs and health outcomes are discounted at a rate of 3.5% per annum. The analysis adopts a formal price year of 2022/2023, including uplifting of older cost estimates using inflation indices,[148] where necessary.

**Table 29: Scope of the EAG economic analysis**

| | |
|---|---|
| Population | Women with ER+/PR+, HER2-, LN+ early breast cancer (1-3 nodes).<br>For the evaluation of Oncotype DX using the newer cut-offs (RS 0-25 and RS >25), pre-menopausal and post-menopausal subgroups are considered separately.<br>For the evaluation of MammaPrint, the modelled population reflects the mAOL clinical high-risk ER+, HER2-, LN+ subgroup within the MINDACT trial.[29] |
| Interventions | (1) Oncotype DX (two sets of cut-offs assessed: (a) new cut-offs – low RS 0-25, RS high >25; (b) old cut-offs – low RS <18; intermediate RS 18-30; high RS >30)<br>(2) Prosigna (cut-offs LN+: low 0-15, intermediate 16-40, high 41-100)<br>(3) EPclin (cut-off: 3.3: low <3.3; high ≥3.3)<br>(4) MammaPrint (cut-off: low >0, high ≤0). |
| Comparator | Current decision-making, which may include any tool, or clinical and pathological features, used to assess risk.<br>For MammaPrint, the clinical high-risk subgroup is based on mAOL, as per the design of the MINDACT trial.[29, 149]<br>Due to evidence limitations, the tumour profiling tests are not compared incrementally against each other.† |
| Main economic outcome | Incremental cost per QALY gained |
| Additional model outcomes | • Incremental LYGs<br>• Incremental QALYs gained<br>• Incremental costs<br>• Impact on chemotherapy use |
| Perspective | NHS and PSS |
| Time horizon | Lifetime |
| Discount rate | 3.5% per annum |
| Price year | 2022/2023 |

*ER - oestrogen receptor; HER2 - human epidermal growth factor receptor; LN - lymph node; LYG - life year gained; QALY - quality-adjusted life year; NHS - National Health Service; PSS - Personal Social Services; RS - recurrence score; mAOL - modified Adjuvant! Online*

*† Risk classification probabilities and DRFI probabilities for MammaPrint, Oncotype DX (using the newer cut-offs) and other tests are derived from different sources. Risk classifications and DMFI probabilities from the TransATAC trial are based on datasets which feature different sample sizes between the tests.*

*Population*

Overall, the population reflected in the economic model relates to women with ER+/PR+, HER2-, LN+ (1-3 nodes) early breast cancer. This is consistent with the final NICE scope.[11] The following issues should be noted with respect to the modelled population:

- In line with the Cytel CEA report,[83] the analysis of MammaPrint is focussed on a subgroup of patients who are defined as clinical high-risk based on mAOL. Women who are at low clinical risk of distant recurrence are not included in the EAG's model for MammaPrint.
- All patients included in the model are women. Owing to a lack of evidence, no economic analysis has been conducted for men with breast cancer.
- The studies used to inform baseline DRFI rates with ET alone are TransATAC,[19] RxPONDER[28, 76] and MINDACT.[29, 149] TransATAC included post-menopausal women only.

RxPONDER recruited both pre- and post-menopausal women and separate outcomes data are available for each of these groups. MINDACT recruited pre- and post-menopausal women, but separate results by menopausal status are not available within the LN+ subgroup. The extent to which menopausal status can be reflected in the economic model therefore differs between the tumour profiling tests.

- Oncotype DX, EPclin and MammaPrint are indicated both for pre-menopausal and post-menopausal women. Prosigna is indicated for post-menopausal women only (see Table 4).

*Interventions*

The EAG's economic analysis includes all four tumour profiling tests included in the final NICE scope:[11] Oncotype DX, Prosigna, EPclin and MammaPrint. The cut-offs assumed for each of these tests are described in Table 29. These cut-offs are in line with the way in which each test is currently used in clinical practice, or how they are expected to be used in the future. For Oncotype DX, two different sets of cut-offs are applied: (a) the newer cut-offs of RS 0-25 and RS >25, as assessed in RxPONDER,[28] and (b) the older cut-offs of RS <18, RS 18-30 and RS >30, as applied in DG34.[13] For EPclin, Prosigna and MammaPrint, only a single set of cut-offs is assumed.

Each of the tests are assumed to be applied together with clinical-pathological factors and patient choice. As such, a high-risk test result does not necessarily lead to a decision to receive chemotherapy and a low-risk result does not necessarily lead to a decision to forgo chemotherapy.

*Comparator*

The comparator reflected in the model is current decision-making. Advice received from the EAG's clinical advisors suggested that current decisions on the use of adjuvant chemotherapy may be informed by risk prediction tools such as PREDICT or NPI, or through consideration of specific clinical and/or pathological factors without the use of a statistical risk prediction tool. A specific decision-making tool is not reflected in the model. Instead, current decision-making is characterised as the pre-test probability of receiving chemotherapy in the absence of tumour profiling testing.

Within the MINDACT trial,[29, 149] clinical high-risk was defined using mAOL. During the appraisal consultation process, the company highlighted that within the HR+, HER2-, LN1-3 population, mAOL high-risk is equivalent to NPI>3.4.

Owing to the use of different evidence sources on clinical outcomes (test risk classification probabilities and DRFI estimates) between the tumour profiling tests, the overlapping but non-identical samples used between alternative tests in TransATAC,[19] and the availability of evidence by menopausal status for some tests but not for others, each test is compared only against current decision-making; tests are not compared incrementally against each other.

*Base case scenarios presented by the EAG*

The EAG's economic analyses are comprised of seven base case scenarios; hereafter, these are denoted "BC" followed by the scenario number. These scenarios have been designed to reflect: (i) the analyses presented by the EAG to inform DG34;[10] (ii) more recent evidence on the tests published since DG34, and (iii) key scenarios presented in the Cytel CEA report[83] and the Exact Sciences CS.[22] The EAG scenarios presented in this report are summarised in Box 3.

**Box 3: Summary of EAG base case scenarios**

- **BC1 –** Oncotype DX versus current decision-making, RxPONDER pre-menopausal LN+ subgroup,[76] supplemented using external data on women with an RS of >25 (thereby assuming predictive benefit).[19, 31] This scenario is similar to the pre-menopausal LN+ subgroup analysis presented in the Exact Sciences CS.[22]
- **BC2 –** Oncotype DX versus current decision-making, RxPONDER post-menopausal LN+ subgroup,[76] supplemented using external data on women with an RS of >25 (thereby assuming predictive benefit).[19, 31] This scenario is similar to the post-menopausal LN+ subgroup analysis presented in the Exact Sciences CS.[22]
- **BC3 –** Oncotype DX versus current decision-making, TransATAC post-menopausal LN+ population,[19] assuming predictive benefit based on SWOG-8814.[31] This scenario is similar to the EAG analysis which included predictive benefit for Oncotype DX in the LN+ population in Harnan *et al.*[10]
- **BC4 –** Oncotype DX versus current decision-making, TransATAC post-menopausal LN+ population,[19] assuming prognostic benefit only.[15] This scenario is similar to the EAG base case scenario for Oncotype DX in the LN+ population in Harnan *et al.*[10]
- **BC5** – Prosigna versus current decision-making, TransATAC post-menopausal LN+ population,[19] assuming prognostic benefit only.[15] This scenario is similar to the EAG non-predictive base case scenario for EPclin in the LN+ population in Harnan *et al.*[10]
- **BC6 –** EPclin versus current decision-making, TransATAC post-menopausal LN+ population,[19] assuming prognostic benefit only.[15] This scenario is similar to the EAG non-predictive base case scenario for Prosigna in the LN+ population in Harnan *et al.*[10]
- **BC7 –** MammaPrint versus current decision-making, MINDACT pre-/post-menopausal, LN+ clinical high-risk subgroup, assuming prognostic benefit only. This scenario is similar to the LN+ subgroup analysis presented in the Cytel CEA report,[83] but excludes the company's assumption of a predictive benefit for MammaPrint (see Section 4.2.2). One-third of patients are assumed to be pre-menopausal; insufficient data were available to allow for subgroup analyses by menopausal status.

Alongside the additional clinical evidence incorporated into Scenarios BC1, BC2 and BC7, all of the EAG's analyses also include assumptions and parameter values which have been updated since DG34. These model amendments are described in Section 4.3.3.

The EAG notes that Exact Sciences have indicated that the validated RS result cut-offs used in RxPONDER[76] should be used to categorise patients according to their risk of distant recurrence and that the Instructions For Use for the Oncotype DX test will be updated to reflect this.[101] The EAG's clinical advisors also commented that they would use Oncotype DX based on these newer cut-offs. As such, BC3 and BC4 are consistent with the previous analyses used to inform DG34,[13] but may be less relevant for NICE decision-making if the older Oncotype DX RS cut-offs are no longer used in practice.

### *4.3.2 Model structure*

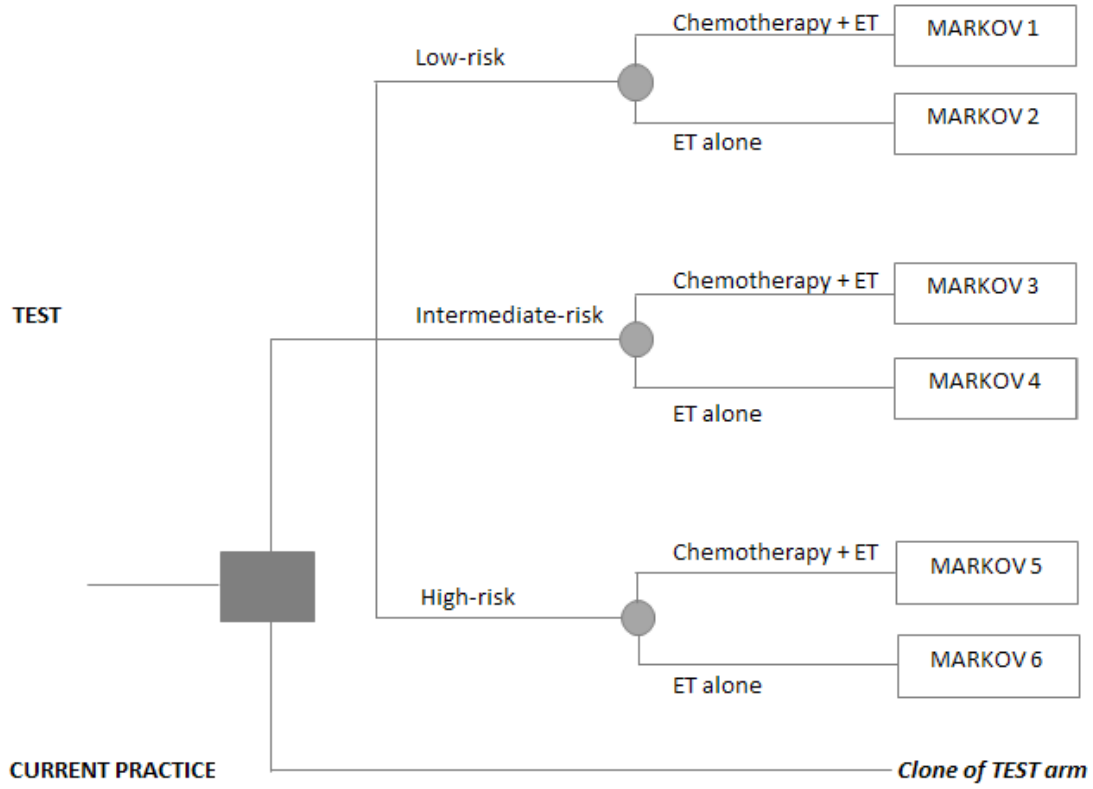
The EAG's model is based on the economic analysis used to inform DG34,[13] together with updated evidence both on the tumour profiling tests and updated evidence and assumptions regarding downstream events, health outcomes and costs. The general model structure is consistent with the majority of studies identified in economic review (see Section 4.1) as well as the models submitted by Agendia[83] and Exact Sciences[22] (see Section 4.2). The EAG's model is intended to capture the key trade-offs in the use of tumour profiling tests in guiding the decision to receive or forgo chemotherapy. Specifically, the model reflects the benefits associated with adjuvant chemotherapy in terms of the reduction in the risk of developing DM and the avoidance of adverse impacts of relapse on HRQoL, survival and costs, as well as its negative effects, which include short-term toxicities and late effects (AML) and the costs of the adjuvant chemotherapy itself. Within the model, the benefits of the tumour profiling tests are modelled by changing the probability that patients receive chemotherapy. In scenarios in which the test is assumed to be predictive of chemotherapy benefit, the relative treatment effect for chemotherapy versus ET alone differs between genomic risk classification groups.

The model takes the form of a hybrid decision tree and long-term Markov model. The decision tree component stratifies patients according to their genomic risk (low-, intermediate- or high-risk for 3-level tests [Prosigna and Oncotype DX using the older RS cut-offs] or low- or high-risk for 2-level tests [Oncotype DX using the newer RS cut-offs, EPclin and MammaPrint]) and according to whether the patient receives chemotherapy conditional on their genomic risk classification. As such, the decision tree determines the distribution of patients across up to six categories:

(i) Genomic low-risk, chemotherapy plus ET
(ii) Genomic low-risk, ET alone
(iii) Genomic intermediate-risk, chemotherapy plus ET (used for 3-level tests only)
(iv) Genomic intermediate-risk, ET alone (used for 3-level tests only)
(v) Genomic high-risk, chemotherapy plus ET
(vi) Genomic high-risk, ET alone.

Each of these branches is linked to a long-term Markov model which predicts lifetime QALY gains and costs conditional on the patient's risk of DM, whether or not they receive chemotherapy, and the magnitude of the treatment effect of chemotherapy on DM.

**Figure 5: EAG's model – decision tree component**

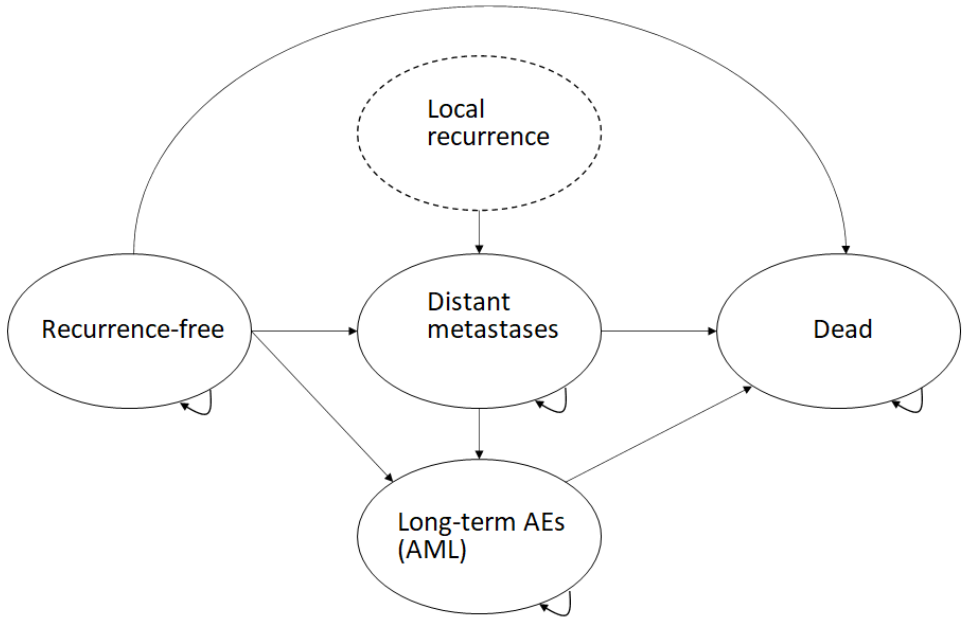



*ET – endocrine therapy*
*Notes: For Oncotype DX under the new cut-offs, EPclin, MammaPrint, four branches are used due to the absence of an intermediate-risk category for these tests. All patients are also assumed to receive ET.*

The structure of the long-term Markov sub-models is illustrated in Figure 6. Each Markov sub-model is evaluated using six-monthly cycles until the patient cohort has reached age 100. Patients enter the model aged 62 years if post-menopausal, or aged 44 years if pre-menopausal, and the evaluation is continued until the cohort has reached age 100 years. Each Markov sub-model includes 4 mutually exclusive and jointly exhaustive health states: (1) recurrence-free; (2) DM; (3) AML and (4) dead. Each sub-model differs with respect to the patient's risk of developing DM, as determined by their genomic risk classification and whether or not they receive chemotherapy. For all Markov sub-models, patients enter the model in the distant recurrence-free state. During each 6-month cycle, patients who are recurrence-free can remain recurrence-free, develop DM, develop AML (if they have previously received adjuvant chemotherapy) or die. CHF was not explicitly included as a late effect in the model because the EAG's clinical experts stated that oncologists are generally able to select out those patients who are at risk of this event based on clinical risk factors, baseline cardiac function and biochemical tests.[150] Patients who are alive with DM can remain in their current health state, develop AML or die.

For patients who have developed AML, the only remaining transition is to the dead state. Patients may die from breast cancer, AML or other causes. Adverse effects of chemotherapy are captured through the inclusion of a short-term (1-year) toxicity-related QALY loss and additional AE management costs applied in the first model cycle, and through the inclusion of the AML state which impacts on survival, HRQoL and costs. The benefit of chemotherapy is modelled through the application of an HR which is applied to the probability of developing DM with ET alone within each genomic risk category. In all scenarios, the use of the tumour profiling test impacts on health outcomes and costs by influencing the probability that a patient receives chemotherapy. In BC1-3, a predictive benefit is assumed for Oncotype DX; hence, the HR for distant recurrence for CET versus ET alone is assumed to differ between the risk classification groups.

**Figure 6: EAG's model – long-term Markov model component**




*AE - adverse event; AML - acute myeloid leukaemia*
*Note: In line with the model developed to inform DG34, once-only costs and QALY losses associated with local recurrence are modelled for women who develop DM (see Section 4.3.3)*

QALYs gained are estimated by assigning health utility values to each of the Markov sub-model health states. The model also includes a short-term QALY loss associated with AEs resulting from the use of adjuvant chemotherapy and a QALY loss associated with the incidence of LR. Health utility values are adjusted for increasing age.

The model includes costs associated with: the tumour profiling tests; acquisition and administration of adjuvant chemotherapy and supportive medications, ET, other drug treatments (bisphosphonates and/or ovarian suppression), routine follow-up visits and tests, treatments for LR, treatments for DM, treatments for AML and end of life care.

The cost-effectiveness of the tumour profiling tests is evaluated using pairwise comparisons for each test versus current decision-making.

*Key EAG model assumptions*

The model employs the following structural assumptions:

- The pre-menopausal model population (BC1) enters the model aged 44 years. The post-menopausal model population (BC2-6) enters the model aged 62 years. For the MammaPrint evaluation (BC7), one-third of the population is assumed to be pre-menopausal; hence, they enter the model aged 56 years.
- Oncotype DX is assumed to be predictive of chemotherapy benefit in some of the scenarios evaluated (BC1-3, see Box 3). Prosigna, EPclin and MammaPrint are assumed to have prognostic benefit only.
- Three-level test results (low-, intermediate- and high-risk) would be interpreted in the same way across all 3-level tests. Two-level test results (low- and high-risk) would be interpreted in the same way across all 2-level tests.
- The risk of DM with ET alone remains constant over time.
- The risk of death in women who remain recurrence-free is assumed to be equivalent to that of the age-matched female general population.
- The model includes a structural constraint which ensures that the risk of death in women with DM or AML is at least as high as the risk of death in the age-matched female general population.
- All women are assumed to receive a CDK4/6 inhibitor as first-line treatment for DM.
- Chemotherapy-related AEs impact on patient HRQoL for 1 year.
- LR impacts on patient HRQoL for 1 year.
- ET is assumed to be given 5 years for all women, with extended therapy given for 10 years in 80% of women.
- Based on clinical input, ovarian suppression is assumed to be offered to 60% of pre-menopausal women for 5 years.
- Based on clinical input, bisphosphonates are assumed to be offered to 60% of post-menopausal women for 3 years.
- Follow-up visits and imaging are assumed to continue for up to 5 years.

### *4.3.3 Evidence sources used to inform the model parameters*

Table 30 summarises the key evidence sources used to inform the model. Further details on the individual model parameters are provided in the subsequent sections.

**Table 30: Evidence sources used in the EAG's base case model**

| Parameter group | Source |
|---|---|
| **Patient characteristics** | |
| Patient age | Holt *et al.*[17] |
| Mean BSA | Assumption |
| **Clinical parameters** | |
| Risk classification probabilities | TransATAC,[19] RxPONDER,[28] MINDACT[29] |
| 6-month DRFI on ET alone | TransATAC,[19] RxPONDER,[76] MINDACT IPD[83] |
| Pre-test probability of receiving adjuvant chemotherapy | Holt *et al.*[17, 35] |
| Post-test probability of receiving chemotherapy (3-level tests) | Holt *et al.*[17, 35] |
| Post-test probability of receiving chemotherapy (2-level tests) | Holt *et al.*[17, 35] |
| HRs for DM, CET vs. ET | EBCTCG,[15] RxPONDER,[76] SWOG-8814[31] |
| 6-month probability of death due to DM | Rebuilt model based on Suri *et al.*[151] |
| Probability of developing LR | De Bock *et al.*[108] |
| 6-month probability of developing AML | Wolff *et al.*[106] |
| 6-month probability of death due to AML | Rebuilt model based on Bewersdorf *et al.*[152] |
| Other-cause mortality | ONS life tables for England[153] |
| **Health-related quality of life parameters** | |
| Utility, recurrence-free | Lidgren *et al.*[109] |
| Utility, distant recurrence | Lidgren *et al.*[109] |
| Utility AML | Rebuilt model based on Bewersdorf *et al.*[152] |
| QALY loss due to chemotherapy-related AEs | Campbell *et al.*[110] |
| QALY loss due to LR | Campbell *et al.*[110] |
| **Cost parameters** | |
| Tumour profiling test costs | Test manufacturers[22-25] |
| Adjuvant chemotherapy and supportive medications | Proportions based on expert opinion.[90] Unit costs taken from eMIT[154] and BNF.[155] |
| ET | Distribution and duration of treatments based on expert opinion. Drug costs taken from eMIT[154] and BNF |
| Bisphosphonates | Proportion based on expert opinion. Unit costs taken from eMIT[154] and BNF.[155] |
| AEs | Frequency based on Ellis *et al.*[114] Costs taken from NHS Reference Costs 2021/22.[141] |
| Routine follow-up | Frequency based on expert opinion. Unit costs taken from Ward *et al.*[103] and NHS Reference Costs 2021/22.[141] |
| LR (once-only cost) | Karnon *et al.*[122] |
| DM (lifetime cost) | Mean cost reported by Suri *et al.*[151] |
| AML (lifetime cost) | Costs of intensive therapy, HSCT and subsequent BSC from Zeidan *et al.*[113] applied to rebuilt model based on Bewersdorf *et al.*[152] |

*BSA - body surface area; DRFI - distant recurrence-free interval; ET - endocrine therapy; HR - hazard ratio; LR - local recurrence; DM - distant metastases; AML - acute myeloid leukaemia; CET - chemotherapy and endocrine therapy; BSC - best supportive care; eMIT - electronic Market Information Tool; - BNF - British National Formulary; ONS - Office for National Statistics; QALY - quality-adjusted life year; AE - adverse event; NHS - National Health Service; HSCT - haematopoietic stem cell transplantation*

*Patient characteristics*

The model assumes that pre-menopausal women who are eligible for tumour profiling testing have a mean age of 44 years (applied in BC1), whereas post-menopausal women who are eligible for tumour profiling testing have a mean age of 62 years (applied BC2-6), based on the age distribution of patients included in Holt *et al.*[17] The analysis of MammaPrint (BC7) applies a mean age of 56 years, assuming that one-third of women are pre-menopausal.[149] Patients are assumed to have a mean body surface area (BSA) of 1.75$m^{2}$; this assumption influences the costs of adjuvant chemotherapy regimens.

*Clinical parameters*

*Risk classification probabilities*

The test risk classification probabilities applied in the EAG's economic model are summarised in Table 31. Across the seven base case comparisons presented, risk classification probabilities for the tests were drawn from three different sources:

- Within BC1 and BC2, risk classification probabilities for Oncotype DX were based RxPONDER.[28] The probability of being in the RS >25 group was estimated based on the number of women who were excluded from RxPONDER due to an RS >25, divided by the number of women registered for screening in the trial and who were eligible for trial entry (N=9,112; see Figure 1 in Kalinsky *et al.*[28]).
- Within BC3-6, risk classification probabilities for Oncotype DX, EPclin and Prosigna were taken from the published analysis of TransATAC.[19]
- Within BC7, risk classification probabilities for MammaPrint were based on the HR+, HER2-, LN+ clinical high-risk group in MINDACT (taken from Piccart *et al.*,[29] Supplementary Appendix, page 29, Table S10).

**Table 31: Risk classification probabilities used in the EAG's model**

| Scenario | Scenario description | Test | Test risk classification probability | | | Source |
|---|---|---|---|---|---|---|
| | | | Low | Int. | High | |
| BC1 | RxPONDER pre-menopausal | Oncotype DX | 0.89 | N/a | 0.11* | Kalinsky *et al.*[28] |
| BC2 | RxPONDER post-menopausal | Oncotype DX | 0.89 | N/a | 0.11* | Kalinsky *et al.*[28] |
| BC3 | TransATAC, predictive | Oncotype DX | 0.57 | 0.32 | 0.11 | Sestak *et al.*[19] |
| BC4 | TransATAC, non-predictive | Oncotype DX | 0.57 | 0.32 | 0.11 | Sestak *et al.*[19] |
| BC5 | TransATAC, non-predictive | Prosigna | 0.08 | 0.32 | 0.60 | Sestak *et al.*[19] |
| BC6 | TransATAC, non-predictive | EPclin | 0.23 | N/a | 0.77 | Sestak *et al.*[19] |
| BC7 | MINDACT clinical high-risk | MammaPrint | 0.69 | N/a | 0.31 | Piccart *et al.*[29] |

*BC - base case; Int. - intermediate; N/a - not applicable*

** Calculated as the number of women excluded from RxPONDER because they had an RS >25 divided by the number of women screened and eligible for entry into the trial*

*DRFI on ET alone*

DRFI estimates for women receiving ET alone were taken from RxPONDER, TransATAC and MINDACT.[19, 76, 83] Across all seven base case scenarios, the source used to inform DRFI was consistent with the source used to inform test risk classification probabilities described in the previous section. The EAG notes the following:

- For the analyses of Oncotype DX using RxPONDER (BC1 and BC2), DRFI probabilities were taken from slides presented by Kalinsky *et al.*[76] This source was used because it reports DRFI estimates by menopausal status.
- The timepoints for reporting DRFI differ between the three trials, with RxPONDER,[76] MINDACT[29] and TransATAC[19] reporting estimates at 5-, 8- and 10-years, respectively. Within the economic model, the cumulative probability of DRFI for the reported time period in each trial was converted to a 6-month probability, assuming a constant event rate.
- For the comparison of MammaPrint versus current decision-making (BC7), the DRFI estimate for women with HR+, HER2-, LN+ breast cancer who are clinical high-risk and genomic low-risk was estimated by the company using IPD from MINDACT.[83] The DRFI estimate for women with HR+, HER2-, LN+ breast cancer who are both clinical high-risk and genomic high-risk was taken from Piccart *et al.*,[29] (Supplementary Appendix, page 29, Table S10). The vast majority of these women received chemotherapy, and no DRFI estimate is reported for women who did not receive chemotherapy. The DRFI for women who are clinical high-risk and genomic high-risk who receive ET alone was estimated by applying the inverse of the HR from the EBCTCG meta-analysis[15] (1/0.71) to the DRFI estimate for the clinical high-risk genomic high-risk group. This approach assumes no predictive benefit for MammaPrint.

**Table 32: Cumulative DRFI probabilities for ET alone used in the EAG's model**

| Scenario | Scenario description | Test | DRFI time point reported | Cumulative DRFI | | | Source |
|---|---|---|---|---|---|---|---|
| | | | | Low | Int. | High | |
| BC1 | RxPONDER pre-menopausal | Oncotype DX | RS0-25: 5 years RS>25: 10 years | 0.06 | N/a | 0.38 | RS0-25: Kalinsky *et al.*[76] RS >25: Sestak *et al.*[19] |
| BC2 | RxPONDER post-menopausal | Oncotype DX | RS0-25: 5 years RS>25: 10 years | 0.03 | N/a | 0.38 | RS0-25: Kalinsky *et al.*[76] RS >25: Sestak *et al.*[19] |
| BC3 | TransATAC, predictive | Oncotype DX | 10 years | 0.19 | 0.29 | 0.38 | Sestak *et al.*[19] |
| BC4 | TransATAC, non-predictive | Oncotype DX | 10 years | 0.19 | 0.29 | 0.38 | Sestak *et al.*[19] |
| BC5 | TransATAC, non-predictive | Prosigna | 10 years | 0.00 | 0.21 | 0.31 | Sestak *et al.*[19] |
| BC6 | TransATAC, non-predictive | EPclin | 10 years | 0.06 | N/a | 0.30 | Sestak *et al.*[19] |
| BC7 | MINDACT clinical high-risk | MammaPrint | 8 years | 0.09 | N/a | 0.26* | MINDACT IPD[83] |

**Based on the cumulative DRFI for clinical high-risk genomic high-risk women raised to the power of the inverse HR from the EBCTCG meta-analysis, thereby assuming no predictive effect*

*BC - base case; DRFI - distant metastasis-free interval; RS - recurrence score; int. - intermediate; IPD - individual patient data; N/a - not applicable*

*HR for distant recurrence, chemotherapy versus no chemotherapy*

Estimates of relative treatment effects for CET versus ET alone were taken from several sources and are assumed to differ between the base case scenarios, depending primarily on whether the test is assumed to be predictive of chemotherapy benefit.

- For the analysis of Oncotype DX using the newer cut-offs (BC1 and BC2), the model applies the competing risks adjusted HRs by menopausal subgroup, as reported in the additional analysis of RxPONDER by Kalinsky *et al.*[76] As women with an RS of >25 were excluded from RxPONDER, the HR for chemotherapy in the RS >25 group was based on the HR for women with an RS of ≥31 in SWOG-8814 (Albain *et al*.[31]). This indirectly assumes that Oncotype DX is predictive of chemotherapy benefit.
- For the analysis of Oncotype DX using the older cut-offs and including an assumption of predictive benefit (BC3), the model applies different HRs by Oncotype DX RS category (low RS 0-18; intermediate RS 18-30; high RS ≥31) based on SWOG-8814.[31]
- For the analyses of all tests without predictive benefit (BC4-7), the model applies an HR for DRFI based on the EBCTCG meta-analysis.[15] The model used to inform DG34 applied an RR of 0.76, based on the annual event rates for DM for anthracycline-based regimens versus no chemotherapy (EBCTCG meta-analysis,[15] Web Extra Material, Analysis P11, page 12). For simplicity, the EAG's model for this appraisal instead applies an estimated HR of 0.71, based on the same event rate data used in Harnan *et al*.

**Table 33: HRs for DM for chemotherapy versus no chemotherapy applied in the EAG's model**

| Scenario | Scenario description | Test | Test risk | | | Source of HRs |
|---|---|---|---|---|---|---|
| | | | Low | Int. | High | |
| BC1 | RxPONDER pre-menopausal | Oncotype DX | 0.64 | N/a | 0.59 | RS 0-25: Kalinsky *et al.*[76]<br>RS >25: Albain *et al.*[31] |
| BC2 | RxPONDER post-menopausal | Oncotype DX | 1.12 | N/a | 0.59 | RS 0-25: Kalinsky *et al.*[76]<br>RS >25: Albain *et al.*[31] |
| BC3 | TransATAC, predictive | Oncotype DX | 1.02 | 0.72 | 0.59 | Albain *et al.*[31] |
| BC4 | TransATAC, non-predictive | Oncotype DX | 0.71 | 0.71 | 0.71 | EBCTCG[15] |
| BC5 | TransATAC, non-predictive | Prosigna | 0.71 | 0.71 | 0.71 | EBCTCG[15] |
| BC6 | TransATAC, non-predictive | EPclin | 0.71 | N/a | 0.71 | EBCTCG[15] |
| BC7 | MINDACT clinical high-risk | MammaPrint | 0.71 | N/a | 0.71 | EBCTCG[15] |

*BC - base case; Int. - intermediate; RS - recurrence score; EBCTCG - Early Breast Cancer Trialists' Collaborative Group*
*Note – values shown in the table are median estimates. These are converted to mean values within the model.*

*Pre-test probability of receiving adjuvant chemotherapy*

The probability of receiving adjuvant chemotherapy without the test was taken from the unpublished decision impact study reported by Holt *et al.*[17] This study was provided as part of the Peony Breast Cancer Unit submission to NICE. This decision impact study was conducted in a cross section of UK NHS hospitals and was designed to measure the decision impact of using Oncotype DX test in women with HR+, HER2-, LN+ breast cancer. Within this study, 530 of 664 (79.82%) women had an initial recommendation to receive chemotherapy. This value is used for the current decision-making group across all key base case scenarios.

**Table 34: Pre-test probability of receiving chemotherapy**

| Scenario | Scenario description | Test | Pre-test probability | Source |
|---|---|---|---|---|
| BC1-7 | All scenarios | All tests | 0.80 | Holt *et al.*[17] |

*BC - base case*

*Post-test probability of receiving chemotherapy*

The post-test chemotherapy probabilities applied in the model are shown in Table 35. Post-test chemotherapy probabilities were selected based on consideration of the studies included in the systematic review of decision impact studies (see Section 3.6):

- For the analyses of Oncotype DX at the newer cut-offs of RS 0-25 and >25 (BC1 and BC2), the model uses estimates of post-test chemotherapy probabilities reported by Holt *et al.*.[17] This study was selected because it is a recent UK-based study undertaken in a LN+ population, because it reports chemotherapy use according to the RxPONDER cut-offs and because separate data are available by menopausal status. This study was also used as the source of the pre-test chemotherapy probability, which provides consistency between data sources used in the model.
- A further abstract of the same decision impact study reported by Holt *et al.*[35] provides estimates of post-test chemotherapy probabilities according to the older cut-offs of RS <18, 18-30 and >30 for the same patient population. These estimates were applied for Oncotype DX in BC3 and BC4.
- The systematic review did not identify any relevant decision impact studies of Prosigna in people with LN+ breast cancer. In BC5, the model assumes that Prosigna test results would be interpreted in the same way as other 3-level tests. The model uses the post-test chemotherapy probabilities derived from the re-analysis of the Holt *et al.* data[35] based on the older Oncotype DX cut-offs (the same estimates applied in BC3 and BC4).
- The systematic review did not identify any relevant decision impact studies for either EPclin or MammaPrint. For consistency with the analyses of Oncotype DX using the newer cut-offs, the model uses post-test chemotherapy probabilities for the post-menopausal subgroup of Holt *et*

*al*.[17] for BC6 (EPclin) and weighted pre- and post-menopausal estimates from Holt *et al*. for BC7 (MammaPrint).

Additional sensitivity analyses were undertaken to explore the impact of applying post-test chemotherapy probabilities derived from other studies identified within the systematic review of decision impact studies (see Section 3.6). These alternative sources include Llombart Cussac *et al.*,[43] Loncaster *et al.*,[36] Zambelli *et al.*,[44] Dieci *et al.*[41] and the UKBCG survey reported in Harnan *et al.*[10]

**Table 35: Post-test probability of receiving chemotherapy**

| Scenario | Scenario description | Test | Test risk | | | Source |
|---|---|---|---|---|---|---|
| | | | Low | Int. | High | |
| BC1 | RxPONDER pre-menopausal | Oncotype DX | 0.37 | N/a | 0.96 | Holt *et al.,*[17] pre-menopausal subgroup |
| BC2 | RxPONDER post-menopausal | Oncotype DX | 0.11 | N/a | 0.96 | Holt *et al.,*[17] post-menopausal subgroup |
| BC3 | TransATAC, predictive | Oncotype DX | 0.08 | 0.49 | 0.98 | Holt *et al.*[35] |
| BC4 | TransATAC, non-predictive | Oncotype DX | 0.08 | 0.49 | 0.98 | Holt *et al.*[35] |
| BC5 | TransATAC, non-predictive | Prosigna | 0.08 | 0.49 | 0.98 | Holt *et al.*[35] |
| BC6 | TransATAC, non-predictive | EPclin | 0.11 | N/a | 0.96 | Holt *et al.,*[17] post-menopausal subgroup |
| BC7 | MINDACT clinical high-risk | MammaPrint | 0.19 | N/a | 0.96 | Holt *et al.,*[17] post-menopausal subgroup |

*BC - base case; Int. - intermediate-risk; N/a - not applicable*

*Long-term risk of DM on ET alone*

The previous model developed to inform DG34[10] assumed that the risk of DM decreases by 50% at 10 years and drops to zero at 15 years and subsequent timepoints. However, a meta-analysis of 88 trials involving 62,923 women with ER+ breast cancer reported by Pan *et al.*[156] suggests that the risk of DM in women with breast cancer with 1-3 involved nodes remains generally flat out to 20 years. Based on the findings of this study, the EAG's model does not assume any risk tapering for patients receiving ET alone.

*Six-month probability of death due to DM*

The previous model developed to inform DG34[10] applied a 6-month probability of death due to DM based on a study of hospital records of 77 UK women who had relapsed breast cancer between 2000 and 2005 (Thomas *et al.*[134]). The EAG's clinical advisors commented that the vast majority of women with ER+ breast cancer who develop DM in England would now receive a CDK4/6i (abemaciclib, palbociclib or ribociclib) as first-line treatment. This aspect of the EAG's model was therefore updated to account for the impact of CDK4/6 inhibitors on OS. The EAG identified a published model-based

economic evaluation of ribociclib plus letrozole versus palbociclib plus letrozole for the treatment of post-menopausal women with HR+, HER2- advanced breast cancer.[151] The published model by Suri *et al.* reports on the incremental cost-effectiveness of ribociclib plus letrozole versus palbociclib plus letrozole over a lifetime horizon from the perspective of the NHS and PSS. Suri *et al.* report the parameters of the baseline Weibull model for OS and HRs obtained from a matching-adjusted indirect comparison (MAIC) of ribociclib plus letrozole versus placebo plus letrozole based on the MONALEESA-2 and PALOMA-1 studies.[157, 158] The EAG replicated the published OS model for the ribociclib plus letrozole group; this model suggests a mean OS of 4.63 years for patients receiving this treatment in the first-line setting. The EAG's model applies a 6-month probability of death due to DM of 0.102, assuming a constant event rate.

*Probability of LR conditional on DM*

The probability of LR was based on a multistate modelling study reported by de Bock *et al.*[108] Within this study, the authors analysed 3,601 women enrolled in three European Organisation for Research and Treatment of Cancer (EORTC) RCTs. The study included both LN0 and LN+ women who had been treated for early-stage breast cancer. Of the 1,224 women who developed DM, 129 women experienced a previous locoregional recurrence. The EAG's model therefore assumes that 10.54% of women who develop DM have a prior LR. The EAG's model does not separately take into account the time spent alive with LR; instead, the impact of LR is applied in the model as a once-only cost and QALY loss. This parameter has not been updated since DG34 and is also used in the models submitted by Agendia and Exact Sciences.[22, 83]

*Six-month probability of developing AML*

The probability of developing AML was derived from a study of the frequency of marrow neoplasms in 20,063 patients with Stage I to III breast cancer treated at US academic centres between 1998 and 2007 (Wolff *et al.*[106]). Within this study, the 10-year cumulative incidence of developing marrow neoplasms was reported to be 0.49% (95% CI 0.11% to 0.87%). The EAG's model applies a 6-monthly probability of developing AML of 0.00025. This probability is applied only to those women who receive chemotherapy. This parameter has not been updated since DG34 and this same probability is used in the model submitted by Agendia.[83]

*Six-month probability of death due to AML*

Within the previous model used to inform DG34, the risk of death due to AML was taken from the EAG report produced to inform the NICE appraisal of azacitidine for the treatment of myelodysplastic syndromes (MDS), chronic myelomonocytic leukaemia (CMML) and AML (TA218).[131] This parameter was updated using more recent evidence. Within the current model, the 6-month probability of death due to therapy-related AML was estimated by reconstructing the intervention group outcomes from a

published model-based economic analysis of liposomal cytarabine/daunorubicin compared with conventional cytarabine/daunorubicin (Bewersdorf *et al.*[152]). The published model presents plots of cumulative survival probabilities based on log-logistic models fitted to data on event-free survival (EFS) and overall survival (OS) from Study 301[123] over a time horizon of 10 years. The EAG digitised the modelled cumulative survival probabilities for EFS and OS in the liposomal cytarabine/daunorubicin group. As around 7% of patients were estimated to still be alive at 10 years in the study by Bewersdorf *et al.*, the EAG extrapolated outcomes to a lifetime horizon using mortality risks from English life tables, together with a standardised mortality ratio (SMR) of 2.3 based on Martin *et al*.[159] The EAG then estimated mean OS using the trapezium rule. This replicated model suggests a mean undiscounted survival duration for the liposomal cytarabine/daunorubicin group of 2.27 years. Mean OS was then converted to a 6-monthly probability of death due to AML of 0.20, assuming a constant event rate.

*All-cause mortality*

Age-specific probabilities of all-cause death were estimated using Office for National Statistics (ONS) life tables for England (years 2018-2020).[153] These mortality risks are applied to all women who remain in the recurrence-free state. These probabilities are also used as constraints in the DM and AML states to ensure that the risk of death with DM and AML remain at least as high as the risk of death in the general population in every model cycle.

*Health-related quality of life*

The utility values and QALY losses applied in the EAG's model are summarised in Table 36. The derivation of each individual utility value / QALY loss is described in further detail in the subsequent sections. Within the economic model, all utility values were adjusted for increasing age using EQ-5D-3L estimates for the general population of the UK reported by Hernández-Alava *et al.*[111]

**Table 36: Utility values and QALY losses applied in the EAG's model**

| Parameter | Mean value | Source |
|---|---|---|
| Utility, recurrence-free | 0.824 | Lidgren *et al.*[109] |
| Utility, DM | 0.685 | Lidgren *et al.*[109] |
| Utility AML | 0.59 | Estimated by dividing the mean QALYs by mean the LYGs in the rebuilt model based on Bewersdorf *et al.*[152] |
| QALY loss chemotherapy | -0.038 | Campbell *et al.*[110] |
| QALY loss LR | -0.108 | Campbell *et al.*[110] |

*AML - acute myeloid leukaemia; QALY - quality-adjusted life year; LYG - life year gained; LR - local recurrence; DM - distant metastases*

*Utility values associated with recurrence-free and DM states and QALY loss associated with chemotherapy-related toxicity*

The model developed to inform DG34[83] applied utility values to the recurrence-free and DM health states based a cross-sectional observational study of 361 patients with previous diagnosis of breast cancer who attended the outpatient clinic at the Karolinska University Hospital in Sweden between April and May 2005 (Lidgren *et al.*[109]). Within this study, patients were asked to complete both the EQ-5D-3L and a direct time trade-off (TTO) question. Patients were then divided into mutually exclusive groups based on their breast cancer disease state: State "P" – first year after primary breast cancer; State "R" first year after recurrence; State "S" – second and following years after primary breast cancer or recurrence and State "M" – metastatic disease. Lidgren *et al.* report a utility value of 0.824 for patients in State S who were receiving ET and a utility value of 0.685 for patients in State M. These values were applied to the recurrence-free and DM states in the model. The disutility associated with chemotherapy was derived from a previous economic model reported by Campbell *et al.*[110]

The EAG undertook a further review to identify other potentially relevant studies which have been published since 2017 (the cut-off date in Harnan *et al.*[10]). Systematic searches were undertaken to identify studies reporting on HRQoL associated with different health states for women with breast cancer. Searches were undertaken in May 2023 in the following electronic databases:

- MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations: Ovid, 1946 to present
- EMBASE: Ovid, 1974 to 2017 July 07
- Science Citation Index Expanded (SCI-E): Web of Science, 1900 to present
- Conference Proceedings Citation Index – Science (CPCI): Web of Science, 1990 to present.

The searches focussed specifically on studies which report HRQoL estimates for health states measured and valued using the EQ-5D. The search strategy comprised sensitive MeSH or Emtree Thesauri terms and free-text synonyms for 'breast cancer' combined with free-text synonyms for 'EQ-5D'. The search strategies are presented in Appendix 1. Studies were considered potentially relevant if they reported EQ-5D valuations for both non-metastatic/early breast cancer and DM states, thereby reflecting key health states in the model. Studies which reported disutilities associated with AEs resulting from the use of chemotherapy were also retained for separate consideration. Studies were sifted by title and abstract according to the inclusion criteria. Full texts were retrieved for all potentially relevant studies identified at the title/abstract stage. In order to be considered for inclusion in the review, studies had to meet the following criteria:

- Must be published in the English language
- Study population or subgroup must reflect early breast cancer population receiving ET (i.e., patients must not be receiving adjuvant or neoadjuvant chemotherapy)

- Must report EQ-5D-3L values for patients who are recurrence-free on ET and for patients who have DM, or must report a disutility associated with receiving CET versus ET alone
- Must reflect a similar patient group to the target population (either European or UK).

The searches identified a total of 404 studies. The full texts of 23 studies were retrieved for more detailed review. None of these studies reported EQ-5D-3L estimates for patients with non-metastatic breast cancer receiving ET and for patients with DM. One "near miss" was identified in which the authors reported EQ-5D-3L utility values for patients with early and metastatic breast cancer (Verrill *et al.*[160]). This study was a UK cross-sectional study of 299 adult patients with HER2+ early or metastatic breast cancer. The authors report mean EQ-5D-3L values of 0.73 for early breast cancer on treatment post-surgery, 0.73 for early breast cancer after completion of adjuvant treatment, and 0.60 for metastatic breast cancer. Given that the population in Verrill *et al.*[160] reflects a HER2+ population, whereas the target population for this appraisal relates to a HER2- population, this study was considered only in sensitivity analyses. As such, the EAG's model retains the use of Lidgren *et al.*[109] as the primary source of utility values.

No other studies were identified which report on the disutility of chemotherapy. The model therefore retains the estimated QALY loss of -0.038 from the model-based economic analysis of chemotherapy for breast cancer reported by Campbell *et al.*[110] This disutility value is also used in the Exact Sciences model[22] and is the same as the value used in the EAG's model used to inform NICE DG34.[10]

*Utility value associated with AML*

The utility value for the AML state was estimated based on the same model used to estimate survival with AML (Bewersdorf *et al.*[152]). The mean utility value over the patient's lifetime for patients was estimated as the mean undiscounted QALYs divided by the mean undiscounted LYGs (1.34/2.27=0.59).

*QALY loss due to LR*

The model applies a QALY loss of -0.108 for patients experiencing LR. This estimate was also taken from Campbell *et al.*[110] This value is also used in the Exact Sciences model[22] and the Agendia model[83] and is the same as the value used in the EAG's model used to inform NICE DG34.[10]

*Resource use and cost parameters*

*Summary of resource use and cost parameters applied in the EAG's model*

The EAG's model includes the costs associated with: the tumour profiling tests; drug treatments (ET, chemotherapy and supportive medications, bisphosphonates and ovarian suppression treatments), routine follow-up visits and tests, treatments for LR, treatments for DM, treatments for AML and end of life care. Table 37 provides a summary of the costs applied in the economic model. All costs were uplifted to current prices using the NHSCII and the HCHS index for published cost estimates valued at 2009 prices or earlier.

**Table 37: Summary of costs applied in the EAG's model**

| Resource use component | | Mean cost |
|---|---|---|
| Tumour profiling tests (list prices) | Oncotype DX | £2,580 |
| | Prosigna | £1,896 |
| | EPclin | £1,500 |
| | MammaPrint | £2,616 |
| Adjuvant chemotherapy (once-only) | | £7,410.48 |
| ET years 1-2 (per cycle) | | £66.95 |
| ET years 3-5 (per cycle) | | £66.44 |
| ET years 6-10 (per cycle) | | £53.16 |
| Bisphosphonates (per cycle)* | | £320.84 |
| Ovarian suppression (per cycle)† | | £496.73 |
| AEs (once-only) | | £1,249.58 |
| Follow-up, year 1 (per year) | | £360.48 |
| Follow-up, years 2-5 (per year) | | £139.00 |
| LR (once-only) | | £16,494.23 |
| DM (once-only) | | £117,482.09 |
| AML (once-only) | | £132,185.91 |
| End of life care (once-only) | | £4,898.17 |

** Applied to post-menopausal women only*
*† Applied to pre-menopausal women only*
*ET - endocrine therapy; AE - adverse event; LR - local recurrence; DM - distant metastases; AML - acute myeloid leukaemia*

*Tumour profiling tests*

The list prices of the tests applied in the EAG's model are summarised in Table 38. Confidential price discounts apply to Oncotype DX, Prosigna and EPclin. The results of the economic analyses including these discounts are provided in a confidential appendix to this report.

**Table 38: Costs of tumour profiling tests**

| Test | List price excluding VAT | Source |
|---|---|---|
| Oncotype DX | £2,580.00 | Exact Sciences CS,[22] May 2023. Price includes costs of all activities required to conduct the testing service, including shipping, materials, customer support, online customer portal for accessing orders and results information. |
| Prosigna | £1,896.00 | Veracyte RFI document,[25] February 2023. Price reflects in-house NHS testing, including costs of gene signature assay, nCounter DX analysis, nCounter servicing, RNA isolation kit and laboratory staff costs. |
| EPclin | £1,500.00 | Myriad RFI document,[24] March 2023. Price includes all reaction agents and consumables. Price reflects locally run testing service. |
| MammaPrint | £2,616.00 | Agendia value dossier,[23] February 2023. Price includes transport, specimen processing and all other costs associated with reporting the result. |

*VAT - Value Added Tax; CS - company's submission; RFI - request for information; RNA - ribonucleic acid; NHS – National Health Service*

*Adjuvant chemotherapy and supportive medications*

The costs associated with adjuvant chemotherapy and supportive medications are summarised in Table 39. The proportionate use of each chemotherapy regimen was taken from Berdunov *et al,*[90] which in turn, was based on the costing approach used by Hall *et al.*[93] The EAG's clinical advisors agreed that these proportions reflect the current use of anthracycline- and taxane-based chemotherapy regimens, but noted that there is an increasing shift away from the use of anthracyclines, particularly for certain patient groups (e.g., those without nodal involvement, those with cardiac co-morbidities and younger patients). In line with Berdunov *et al.*, the model assumes that an anti-emetic (aprepitant) is given in 20% of all chemotherapy cycles. G-CSF (filgrastim) is assumed to be given in 20% of cycles of anthracycline-based regimens and in 100% of cycles of docetaxel regimens and accelerated regimens. The costs also include the costs of pharmacy preparation, outpatient monitoring visits and tests (full blood counts [FBCs], liver function tests [LFTs] and urea and electrolytes [U&Es] and electrocardiograms [ECGs] in 25% of patients). Drug acquisition costs were taken from eMIT.[154] The costs of delivering chemotherapy and tests were taken from NHS Reference Costs 2021/22.[141] The cost of pharmacy preparation was taken from Ward *et al.*[103]

Within the economic model, a weighted mean cost of £7,410.48 is applied to all patients who receive chemotherapy. This cost is applied in the first model cycle as a once-only cost.

**Table 39: Per-cycle adjuvant chemotherapy costs applied in the EAG's model**

| Regimen | Proportion | Drug acquisition cost per model cycle | Administration, pharmacy, visits and monitoring per model cycle | Total cost per model cycle |
|---|---|---|---|---|
| FEC75 (6 cycles) | 0.00% | £821.13 | £4,110.70 | £4,931.83 |
| FEC100-T (3+3 cycles) | 23.75% | £1,581.22 | £4,110.70 | £5,691.92 |
| TC (4 cycles) | 10.00% | £1,640.50 | £2,735.53 | £4,376.03 |
| EC90/T75 (4+4 cycles) | 28.75% | £2,101.75 | £5,485.87 | £7,587.62 |
| EC90 (4 cycles) | 0.00% | £547.90 | £2,735.53 | £3,283.43 |
| C-D (6 cycles) | 2.50% | £2,458.68 | £4,110.70 | £6,569.38 |
| TAC (6 cycles) | 1.25% | £2,509.74 | £4,110.70 | £6,620.44 |
| Accelerated EC90/P (4+4 cycles) | 23.75% | £3,381.97 | £5,485.87 | £8,867.84 |
| Weekly P (12 weeks) | 2.50% | £155.13 | £8,236.21 | £8,391.34 |
| EC/weekly P (4 cycles, 12 weeks) | 7.50% | £703.03 | £10,986.55 | £11,689.58 |
| Weighted cost | - | £2,096.50 | £5,313.98 | **£7,410.48** |

*FEC75 - fluorouracil, epirubicin and cyclophosphamide; FEC100-T - fluorouracil, epirubicin, cyclophosphamide and docetaxel; TC - docetaxel and cyclophosphamide; EC90 - epirubicin and cyclophosphamide; EC90/T75 - epirubicin and cyclophosphamide followed by docetaxel; C-D - carboplatin plus docetaxel; TAC - docetaxel, doxorubicin and cyclophosphamide; Accelerated EC90/P - epirubicin, cyclophosphamide followed by paclitaxel; Weekly P - weekly paclitaxel; EC/weekly P - epirubicin and cyclophosphamide followed by paclitaxel*

*Adverse events associated with adjuvant chemotherapy*

The frequency of Grade 3/4 AEs was informed by the TACT trial[114] (see Table 40). Unit costs were taken from NHS Reference Costs 2021/22[141] based on the same service codes as those used in the Exact Sciences model.[22] The model applies the expected costs associated with AEs in the FEC-D group to all

docetaxel-containing regimens and the costs associated with AEs in the control group to the other regimens included in the model, based on the distribution of regimen usage shown in Table 39. The model applies a weighted cost of £1,249.58 to all patients receiving adjuvant chemotherapy. This cost is applied as a once-only cost in the first 6-month model cycle.

**Table 40: Frequency of AEs and unit costs applied in the EAG's model**

| AE | FEC-D - frequency | Control - frequency | Unit cost | Cost source |
|---|---|---|---|---|
| Anaemia | 0.006 | 0.007 | £1,439.66 | NHS Reference Costs 2021/22:[141] SA04G-L, non-elective short and long stay |
| Febrile neutropoenia | 0.071 | 0.029 | £3,676.55 | NHS Reference Costs 2021/22:[141] SA35A-E, non-elective long stay |
| Leucopoenia | 0.246 | 0.175 | £501.80 | NHS Reference Costs 2021/22:[141] weighted average of JA12D-L, non-elective short and long stay and NHS Reference Costs 2020/21: consultant-led outpatient visit, WF01A medical oncology |
| Neutropoenia | 0.455 | 0.384 | £501.80 | NHS Reference Costs 2021/22:[141] weighted average of JA12D-L, non-elective short and long stay and NHS Reference Costs 2020/21: consultant-led outpatient visit, WF01A medical oncology |
| Thrombocytopenia | 0.006 | 0.013 | £2,163.16 | NHS Reference Costs 2021/22:[141] SA12G-K, non-elective short and long stay |
| Alopecia | 0.102 | 0.103 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Diarrhoea | 0.037 | 0.028 | £1,446.84 | NHS Reference Costs 2021/22:[141] FD10J-M, non-elective short and long stay |
| Infection | 0.142 | 0.088 | £1,628.07 | NHS Reference Costs 2021/22:[141] DZ22K-Q, non-elective short and long stay |
| Lethargy | 0.221 | 0.131 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Musculoskeletal (other) | 0.070 | 0.015 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Myalgia/ arthralgia | 0.050 | 0.001 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Nausea/ vomiting | 0.097 | 0.099 | £1,579.14 | NHS Reference Costs 2021/22:[141] FD11K, non-elective short and long stay |
| Neuropathy | 0.048 | 0.005 | £1,886.35 | NHS Reference Costs 2021/22:[141] AA26C-H, non-elective short and long stay |
| Oedema | 0.008 | 0.003 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Pain | 0.028 | 0.001 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Skin disorder (including nail changes) | 0.033 | 0.012 | £221.48 | NHS Reference Costs 2021/22:[141] consultant-led outpatient visit, WF01A medical oncology |
| Stomatitis | 0.076 | 0.036 | £1,978.14 | NHS Reference Costs 2021/22:[141] CB01F, non-elective short and long stay |

*AE - adverse event; FEC-D – fluorouracil, epirubicin and cyclophosphamide followed by docetaxel; FEC - fluorouracil, epirubicin and cyclophosphamide; NHS - National Health Service*

*Endocrine therapy*

The model assumes that, whilst recurrence-free, all women will receive ET for 5 years, and that 80% of women will receive extended ET for a further 5 years. During the first 5 years following surgery, the model assumes that 15% of women will receive tamoxifen, 23% receive anastrozole, 24% receive letrozole, 23% receive exemestane and 15% receive tamoxifen for 2 years, then exemestane, anastrozole of letrozole for 3 years. These proportions were based on clinical input. The model applies this same distribution of treatments for years 3-5 to those women who continue to receive extended ET during years 6-10. The prices of anastrozole, letrozole and exemestane were taken from eMIT[154] whereas the price of tamoxifen was taken from the British National Formulary (BNF).[155] Monthly pharmacy preparation and dispensing costs were taken from the Personal Social Services Research Unit (PSSRU).[148] The expected cost of ET (including pharmacy prescribing costs) is estimated to be £66.95 in years 1-2, £66.44 in years 3-5 and £53.16 in years 6-10.

*Cost of routine follow-up*

The model assumes that women undergo routine follow-up for 5 years following surgery for their primary breast cancer. Women are assumed to have three outpatient visits in year 1 followed by one annual outpatient visit during years 2-5. Women are also assumed to undergo one annual mammogram during years 1-5. The cost of outpatient follow-up appointments was taken from NHS Reference Costs 2021/22.[141] The unit cost for mammograms is not listed in NHS Reference Costs 2021/2022; instead, this was taken from Ward *et al.*[103] These costs are applied in each 6-monthly cycle for up to 5 years whilst patients remain recurrence-free.

*Cost of bisphosphonates*

The model assumes that 60% of post-menopausal women who are recurrence-free receive bisphosphonates (4mg zoledronic acid) every six-months for three years. Treatment is assumed to be administered in a chemotherapy day unit and involves an additional blood test and a nurse assessment. The proportion of patients receiving treatment and the duration and frequency of administrations were based on estimates provided by the EAG's clinical advisors. The unit cost of zoledronic acid was taken from eMIT[154] and the cost of administration was taken from NHS Reference Costs 2021/2022.[141] These costs are applied in BC2-6 (for the RxPONDER post-menopausal subgroup and TransATAC post-menopausal analyses) and in BC7 (for the proportion of post-menopausal women in MINDACT).

*Cost of ovarian suppression treatment*

The model assumes that 60% of pre-menopausal women who are recurrence-free receive ovarian suppression treatment for up to 5 years. The model assumes that women receiving ovarian suppression are equally likely to receive goserelin, leuprorelin or triptorelin. Treatment is assumed to be administered in an outpatient setting for 15% of women, with the remaining 85% of women receiving

treatment at a GP surgery. These assumptions were based on input from the EAG's clinical advisors. The unit costs of ovarian suppression drugs were taken from the BNF.[155] Administration costs were taken from NHS Reference Costs 2021/2022[141] and the PSSRU.[148] These costs are applied in BC1 (RxPONDER pre-menopausal subgroup) and in BC7 (pre-menopausal women in MINDACT). These costs were not applied in the model used to inform DG34,[10] or in the base case analyses presented by Exact Sciences or Agendia.[22, 83]

*Cost of treating LR*

The cost of treating LR was taken from a breast cancer costing study reported by Karnon *et al*[122] (uplifted cost = £16,494) This is applied as a once-only cost to 10.5% of patients who experience distant recurrence (based on de Bock *et al.*,[108]).

*Lifetime cost of treating DM*

The lifetime cost of treating DM was based on the discounted cost for the ribociclib plus letrozole group of the model reported by Suri *et al.*[151] The EAG's model applies a once-only cost of £117,482 to patients entering the DM health state.

*Lifetime cost of treating AML*

The lifetime cost of AML was based on the same replicated model used to estimate mortality risk and health utility with AML (based on Bewersdorf *et al.*[152]), together with treatment costs reported by Zeidan *et al.*[113] The EAG applied an initial 6-month cost of intensive induction and consolidation therapy to 65% of patients and an initial cost of haematopoietic stem cell transplantation (HSCT) to 35% of patients in the first cycle of the replicated model, based on the proportion of patients proceeding to HSCT in Study 301[123] (weighted initial cost = £72,869). From month 6 onwards, the model applies a monthly cost of BSC of £704. Based on these costing assumptions, the replicated model suggests a mean undiscounted lifetime cost for standard AML treatments of £88,863. The additional cost of liposomal cytarabine/daunorubicin was not available from the committee papers for NICE TA552; instead, an estimated incremental cost of liposomal cytarabine/daunorubicin versus current therapy was taken from a technical briefing on this drug reported by the National Centre for Pharmacoeconomics (NCPE) which was converted to current UK prices using Purchasing Power Parities (additional cost = £43,322.87). Taken together, this suggests an estimated lifetime cost for this treatment of £132,186. The model applies this mean lifetime cost to all patients upon entry into the AML health state.

*Cost of death*

The cost of death was taken from the economic analysis reported by Hinde *et al.*,[91] which in turn, was based on Karnon *et al.*[122] Within the model, a once-only cost of £4,898 is applied to patients when they enter the dead state.

*Model evaluation methods*

For each of the EAG's base case scenarios, cost-effectiveness results are presented for the tumour profiling test versus current decision-making. Results are presented using both the probabilistic and deterministic versions of the model. All probabilistic ICERs are based on 10,000 Monte Carlo simulations. The results of the PSA are also presented using cost-effectiveness planes and CEACs. The distributions used in the PSA are as follows:

- Risk classification distributions were modelled using Dirichlet distributions
- Probabilities and utility values were modelled using beta distributions
- HRs were modelled using log-normal distributions
- Costs were modelled using gamma distributions.

Where sufficient information was available, distribution parameters were characterised using reported standard errors (SEs) or 95% CIs. Where insufficient information was provided, SEs were assumed to be equal to 10% of the mean.

Alongside the PSA, the EAG also undertook a number of DSAs to explore alternative evidence sources and assumptions. The following analyses were undertaken across each of BCs 1-7 (where relevant):

- *DSA1:* As noted in Section 5.2.1.2, there is some uncertainty around the proportion of women who would obtain an Oncotype DX RS of >25. Within BC1 and BC2 (Oncotype DX using data from RxPONDER), 17% of women were assumed to be in the RS >25 group.
- *DSA2:* The test classification probabilities and DRFI estimates for Prosigna were taken from Gnant *et al.*[54] rather than TransATAC.[19] The test risk classification probabilities for low-, intermediate- and high-risk were 0.04, 0.34 and 0.62, respectively. Across these three risk groups, the 6-month probability of DM was estimated to be 0.000, 0.003, and 0.013, respectively.
- *DSA3:* The test classification and DRFI estimates for EPclin were taken from Filipits *et al.*[161] rather than TransATAC.[19] The test risk classification probabilities for low- and high-risk were 0.35 and 0.65, respectively. Across these two risk groups, the 6-month DRFI was estimated to be 0.01 and 0.01, respectively.
- *DSA4:* The post-test chemotherapy probabilities for 3-level tests were based on estimates reported by Llombart Cussac *et al.*[43] rather than Holt *et al.*[17] (values presented in Table 16).
- *DSA5:* The post-test chemotherapy probabilities for 3-level tests were based on estimates reported by Loncaster *et al.*[36] rather than Holt *et al.*[17] (values presented in Table 16).
- *DSA6:* The post-test chemotherapy probabilities for 3-level tests were based on estimates reported by Zambelli *et al.*[89] rather than Holt *et al.*[17] (values presented in Table 16).
- *DSA7:* The post-test chemotherapy probabilities for 2-level tests were based on estimates reported by Dieci *et al.*,[41] rather than Holt *et al.*[17] (values presented in Table 16).

- *DSA8*: The post-test chemotherapy probabilities for 3-level tests are based on the UKBCG survey reported by Harnan *et al.*,[10] rather than Holt *et al.*[17]
- *DSA9*: The post-test chemotherapy probabilities for 2-level tests are based on the UKBCG survey reported by Harnan *et al.*,[10] rather than Holt *et al.*[17]
- *DSA10:* The model includes risk tapering for women receiving either CET or ET alone, with the risk of DM decreasing by 50% after 10 years and dropping to a risk of 0% after 15 years.
- *DSA11:* The HR for CET versus ET was set equal to 0.60 in all genomic risk groups. This assumes prognostic benefit only for all tests.
- *DSA12:* The HR for CET versus ET was set equal to 0.71 in all genomic risk groups. This assumes prognostic benefit only for all tests.
- *DSA13:* The HR for CET versus ET was set equal to 0.80 in all genomic risk groups. This assumes prognostic benefit only for all tests.
- *DSA14:* The chemotherapy QALY loss was halved (from 0.038 to 0.019 QALYs).
- *DSA15:* The chemotherapy QALY loss was doubled (from 0.038 to 0.076).
- *DSA16:* The chemotherapy QALY loss was tripled (from 0.038 to 0.114).
- *DSA17:* The baseline probability of receiving chemotherapy was increased by 10% (from 0.80 to 0.90).
- *DSA18:* The starting age of the population was increased by 5 years.
- *DSA19:* The starting age of the population was reduced by 5 years.
- *DSA20:* The utility values for the recurrence free and DM health states were based on utility values reported by Verrill *et al.* [160] (recurrence-free utility = 0.73; DM utility = 0.60).
- *DSA21:* The probability of developing AML was removed from the model.
- *DSA22:* The cost of adjuvant chemotherapy was halved (from £7,410 to £3,705).
- *DSA23:* The cost of adjuvant chemotherapy was doubled (from £7,410 to £14,821).
- *DSA24:* The lifetime cost of treating DM was halved (from £117,482 to £58,741).
- *DSA25:* The lifetime cost of treating DM was doubled (from £117,482 to £234,964).
- *DSA26:* The lifetime cost of treating AML was halved (from £132,186 to £66,093).
- *DSA27:* The lifetime cost of treating AML was doubled (from £132,186 to£264,372).

In addition to these sensitivity analyses, the EAG also estimated the impact of changes in chemotherapy use following the use of tumour profiling testing on the number of infusion chair hours, based on infusion times obtained from the EAG's clinical advisors (see Appendix 8).

*Model verification methods*

A number of approaches were used to ensure the credibility of the EAG's model. These included:

- Ensuring that the model is consistent with the NICE Reference Case[162] and published checklists for economic evaluations and models.[163, 164]
- Double-programming the deterministic version of the model by the primary model author.
- Checking model implementation by a third-party modeller who was not involved in developing the model itself.
- Ensuring the accuracy of model input parameters against their original sources.
- Checking the appropriateness of model input parameters and assumptions with clinical experts.
- Checking the face validity of the model predictions with clinical experts.

### *4.3.6 Results of the EAG economic analysis*

This section presents the results of the EAG's economic analysis. The results of the EAG's base case analysis generated using the probabilistic and deterministic versions of the model are presented in Table 41 and Table 42, respectively. The results of the DSAs are presented in Table 43. A summary of the model-predicted impact of tumour profiling testing on chemotherapy use, clinical outcomes, costs and net health benefits (NHBs) per 1,000 women tested is presented in Table 44. CEACs for each comparison can be found in Appendix 9. The results of these analyses are summarised together in the subsequent sections.

*Oncotype DX versus current decision-making (BC1-4)*

The probabilistic version of the model for the pre-menopausal LN+ subgroup suggests that compared with current decision-making, Oncotype DX is expected to result in 0.66 fewer LYGs, 0.18 fewer QALYs and additional costs of £1,810 per patient tested. Consequently, Oncotype DX is dominated by current decision-making in this population. These results are driven by the estimated reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment. Assuming willingness-to-pay (WTP) thresholds of £20,000 and £30,000 per QALY gained, the probability that Oncotype DX generates more net benefit than current decision-making is approximately 0.06. The DSAs indicate that Oncotype DX remains dominated across all analyses, except for DSA23 (cost of chemotherapy doubled).

Within the post-menopausal LN+ subgroup, the probabilistic version of the model suggests that compared with current decision-making, Oncotype DX is expected to generate 0.21 additional LYGs, 0.11 additional QALYs and cost savings of £4,273 per patient tested. Consequently, Oncotype DX dominates current decision-making in this population, provided the assumption of predictive benefit holds. These results are driven by an estimated reduction in the use of adjuvant chemotherapy in women who would not have benefitted from chemotherapy. Assuming WTP thresholds of £20,000 and £30,000

per QALY gained, the probability that Oncotype DX generates more net benefit than current decision-making is approximately 1.00. The DSAs indicate that Oncotype DX remains dominant across all analyses except for those in which the assumption of a predictive benefit of chemotherapy is removed (DSAs 11-13); within these scenarios, the ICER for Oncotype DX is in the South-West quadrant and ranges from £9,772 to £279,599 saved per QALY lost. The analyses of Oncotype DX within the post-menopausal LN+ subgroup based on TransATAC[19] using the older RS cut-offs suggest a similar finding – Oncotype DX dominates current-decision-making when a predictive benefit is assumed, but it is dominated by current decision-making when this assumption is removed. These results remain generally consistent across the range of DSAs tested.

*Prosigna versus current decision-making (BC5)*

The probabilistic version of the model suggests that compared with current decision-making, Prosigna is expected to result in 0.06 additional LYGs, 0.03 additional QALYs and additional costs of £1,084 per patient tested; the ICER for Prosigna versus current decision-making is expected to be £39,357 per QALY gained. The model suggests that the use of Prosigna will result in a small decrease in the use of chemotherapy, a small reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. Assuming WTP thresholds of £20,000 and £30,000 per QALY gained, the probability that Prosigna generates more net benefit than current decision-making is approximately 0.16 and 0.34, respectively. The DSAs resulted in ICERs ranging from £23,859 per QALY gained to dominated. The DSAs indicate that the ICER is sensitive to the source of test risk classification probabilities and associated DRFI estimates, the HR for chemotherapy, and the costs of adjuvant chemotherapy and downstream treatments for DM.

*EPclin versus current decision-making (BC6)*

The probabilistic version of the model suggests that compared with current decision-making, EPclin is expected to result in 0.13 additional LYGs, 0.06 additional QALYs and additional costs of £231 per patient tested; the ICER for EPclin versus current decision-making is expected to be £4,113 per QALY gained. The model suggests that the use of EPclin will result in a small decrease in the use of chemotherapy, a reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. Assuming WTP thresholds of £20,000 and £30,000 per QALY gained, the probability that EPclin generates more net benefit than current decision-making is approximately 0.82 and 0.86, respectively. The DSAs resulted in ICERs ranging from dominating to dominated. The DSAs indicate that the ICER is sensitive to the test risk classification probabilities and associated DRFI estimates, the baseline probability of receiving chemotherapy, the HR for chemotherapy, and the costs of adjuvant chemotherapy and downstream treatments for DM.

*MammaPrint versus current decision-making (BC7)*

The probabilistic version of the model suggests that compared with current decision-making, MammaPrint is expected to result in 0.22 fewer LYGs, 0.07 fewer QALYs and additional costs of £786 per patient tested; hence, MammaPrint is dominated by current decision-making. The model suggests that the use of MammaPrint will result in a large decrease in the use of chemotherapy, an increase in the lifetime probability of developing DM and additional net costs due to the cost of the test. Assuming WTP thresholds of £20,000 and £30,000 per QALY gained, the probability that MammaPrint generates more net benefit than current decision-making is approximately 0.01. The DSAs suggest that MammaPrint is either dominated or results in a South-West quadrant ICER which is less than £30,000 per QALY gained across all scenarios tested.

**Table 41: Central estimates of cost-effectiveness, all EAG base case comparisons, probabilistic**

| Option | LYGs* | QALYs | Costs | Inc. LYGs* | Inc. QALYs | Inc. costs | Incremental cost per QALY gained |
|---|---|---|---|---|---|---|---|
| **BC1 - Oncotype DX, RxPONDER pre-menopausal (predictive benefit)** | | | | | | | |
| Oncotype DX | 32.73 | 14.25 | £41,631 | -0.66 | -0.18 | £1,810 | **Dominated** |
| Current DM | 33.39 | 14.43 | £39,821 | - | - | - | - |
| **BC2 - Oncotype DX, RxPONDER post-menopausal (predictive benefit)** | | | | | | | |
| Oncotype DX | 21.82 | 11.18 | £26,546 | 0.21 | 0.11 | -£4,273 | **Dominating** |
| Current DM | 21.61 | 11.07 | £30,818 | - | - | - | - |
| **BC3 - Oncotype DX, TransATAC, post-menopausal (predictive benefit)** | | | | | | | |
| Oncotype DX | 19.29 | 10.11 | £47,762 | 0.05 | 0.04 | -£1,942 | **Dominating** |
| Current DM | 19.24 | 10.07 | £49,704 | - | - | - | - |
| **BC4 - Oncotype DX, TransATAC, post-menopausal (non-predictive benefit)** | | | | | | | |
| Oncotype DX | 19.28 | 10.11 | £47,806 | -0.44 | -0.17 | £1,811 | **Dominated** |
| Current DM | 19.72 | 10.28 | £45,994 | - | - | - | - |
| **BC5 - Prosigna, TransATAC, post-menopausal (non-predictive benefit)** | | | | | | | |
| Prosigna | 19.73 | 10.28 | £47,427 | 0.06 | 0.03 | £1,084 | **£39,357** |
| Current DM | 19.67 | 10.25 | £46,342 | - | - | - | - |
| **BC6 - EPclin, TransATAC, post-menopausal (non-predictive benefit)** | | | | | | | |
| EPclin | 19.88 | 10.34 | £45,786 | 0.13 | 0.06 | £231 | **£4,113** |
| Current DM | 19.75 | 10.29 | £45,555 | - | - | - | - |
| **BC7 - MammaPrint, MINDACT, LN+ subgroup (non-predictive benefit)** | | | | | | | |
| MammaPrint | 24.50 | 12.04 | £40,614 | -0.22 | -0.07 | £786 | **Dominated** |
| Current DM | 24.72 | 12.10 | £39,828 | - | - | - | - |

** Undiscounted*

*LYG - life year gained; QALY - quality-adjusted life year; Inc. - incremental; DM - decision-making*

**Table 42: Central estimates of cost-effectiveness, all base case comparisons, deterministic**

| Option | LYGs* | QALYs | Costs | Inc. LYGs* | Inc. QALYs | Inc. costs | Incremental cost per QALY gained |
|---|---|---|---|---|---|---|---|
| **BC1 - Oncotype DX, RxPONDER pre-menopausal (predictive benefit)** | | | | | | | |
| Oncotype DX | 32.69 | 14.24 | £41,814 | -0.65 | -0.18 | £1,787 | **Dominated** |
| Current DM | 33.34 | 14.42 | £40,027 | - | - | - | - |
| **BC2 - Oncotype DX, RxPONDER post-menopausal (predictive benefit)** | | | | | | | |
| Oncotype DX | 21.81 | 11.23 | £26,630 | 0.21 | 0.11 | -£4,283 | **Dominating** |
| Current DM | 21.60 | 11.12 | £30,913 | - | - | - | - |
| **BC3 - Oncotype DX, TransATAC, post-menopausal (predictive benefit)** | | | | | | | |
| Oncotype DX | 19.26 | 10.15 | £48,145 | 0.08 | 0.05 | -£2,300 | **Dominating** |
| Current DM | 19.18 | 10.10 | £50,444 | - | - | - | - |
| **BC4 - Oncotype DX, TransATAC, post-menopausal (non-predictive benefit)** | | | | | | | |
| Oncotype DX | 19.27 | 10.16 | £47,986 | -0.45 | -0.17 | £1,862 | **Dominated** |
| Current DM | 19.72 | 10.33 | £46,124 | - | - | - | - |
| **BC5 - Prosigna, TransATAC, post-menopausal (non-predictive benefit)** | | | | | | | |
| Prosigna | 19.71 | 10.32 | £47,650 | 0.06 | 0.03 | £1,108 | **£40,220** |
| Current DM | 19.65 | 10.30 | £46,543 | - | - | - | - |
| **BC6 - EPclin, TransATAC, post-menopausal (non-predictive benefit)** | | | | | | | |
| EPclin | 19.86 | 10.39 | £46,080 | 0.12 | 0.05 | £305 | **£5,580** |
| Current DM | 19.74 | 10.33 | £45,775 | - | - | - | - |
| **BC7 - MammaPrint, MINDACT, LN+ subgroup (non-predictive benefit)** | | | | | | | |
| MammaPrint | 24.50 | 12.06 | £40,621 | -0.22 | -0.07 | £792 | **Dominated** |
| Current DM | 24.71 | 12.13 | £39,830 | - | - | - | - |

** Undiscounted*

*LYG - life year gained; QALY - quality-adjusted life year; Inc. - incremental; DM - decision-making*

**Table 43: Deterministic sensitivity analysis results for all base case comparisons - test versus current decision-making**

| DSA | BC1 – Oncotype DX, RxPONDER pre-menopausal, predictive | BC2 – Oncotype DX, RxPONDER post-menopausal, predictive | BC3 – Oncotype DX, TransATAC post-menopausal, predictive | BC4 – Oncotype DX, TransATAC post-menopausal, non-predictive | BC5 – Prosigna, TransATAC post-menopausal, non-predictive | BC6 – EPclin, TransATAC post-menopausal, non-predictive | BC7 – MammaPrint, MINDACT, non-predictive |
|---|---|---|---|---|---|---|---|
| **Deterministic base case ICER** | **Dominated** | **Dominating** | **Dominating** | **Dominated** | **£40,220** | **£5,580** | **Dominated** |
| DSA1: 17% of women assumed to be in RS >25 group (Oncotype DX only) | Dominated | Dominating | N/a | N/a | N/a | N/a | N/a |
| DSA2: Prosigna test classification probabilities and DRFI from Gnant *et al.*,[54] | N/a | N/a | N/a | N/a | £23,853 | N/a | N/a |
| DSA3: EPclin test classification probabilities and DRFI from Filipits *et al.*[161] | N/a | N/a | N/a | N/a | N/a | Dominated | N/a |
| DSA4: 3-level post-test chemotherapy probabilities - Llombart Cussac *et al*[43] | N/a | N/a | Dominating | Dominated | £37,959 | N/a | N/a |
| DSA5: 3-level post-test chemotherapy probabilities - Loncaster *et al.*[36] | N/a | N/a | Dominating | Dominated | Dominated | N/a | N/a |
| DSA6: 3-level post-test chemotherapy probabilities - Zambelli *et al.*[44] | N/a | N/a | Dominating | Dominated | £62,801 | N/a | N/a |
| DSA7: 2-level post-test chemotherapy probabilities – Dieci *et al.*[41] | Dominated | Dominating | N/a | N/a | N/a | £6,448 | Dominated |
| DSA8: 3-level post-test chemotherapy probabilities - UKBCG survey (3-level tests)[10] | N/a | N/a | Dominating | Dominated | £37,092 | N/a | N/a |
| DSA9: 2-level post-test chemotherapy probabilities - UKBCG survey (2-level tests)[10] | Dominated | Dominating | N/a | N/a | N/a | £12,606 | Dominated |
| DSA10: Risk tapering to 50% at 10 years then 0% at 15 years | Dominated | Dominating | Dominating | Dominated | £40,876 | £7,097 | Dominated |

| DSA | BC1 – Oncotype DX, RxPONDER pre-menopausal, predictive | BC2 – Oncotype DX, RxPONDER post-menopausal, predictive | BC3 – Oncotype DX, TransATAC post-menopausal, predictive | BC4 – Oncotype DX, TransATAC post-menopausal, non-predictive | BC5 – Prosigna, TransATAC post-menopausal, non-predictive | BC6 – EPclin, TransATAC post-menopausal, non-predictive | BC7 – MammaPrint, MINDACT, non-predictive |
|---|---|---|---|---|---|---|---|
| DSA11: CET vs ET HR = 0.60 in all genomic risk groups (non-predictive) | Dominated | £9,772 (SWQ) | Dominated | Dominated | £24,584 | Dominating | Dominated |
| DSA12: CET vs ET HR = 0.71 in all genomic risk groups (non-predictive) | Dominated | £42,518 (SQW) | Dominated | Dominated | £40,220 | £5,580 | Dominated |
| DSA13: CET vs ET HR = 0.80 in all genomic risk groups (non-predictive) | Dominated | £279,599 (SWQ) | Dominated | Dominated | £60,336 | £14,493 | Dominated |
| DSA14: Chemotherapy QALY loss halved | Dominated | Dominating | Dominating | Dominated | £44,427 | £5,820 | Dominated |
| DSA15: Chemotherapy QALY loss doubled | Dominated | Dominating | £757,556 (SWQ) | Dominated | £49,618 | £6,080 | Dominated |
| DSA16: Chemotherapy QALY loss tripled | Dominated | Dominating | £106,021 (SWQ) | Dominated | £53,808 | £6,267 | Dominated |
| DSA17: Baseline probability of chemotherapy = 0.90 | Dominated | Dominating | Dominating | Dominated | Dominated | £13,402 | Dominated |
| DSA18: Start age + 5 years | Dominated | Dominating | Dominating | Dominated | £52,697 | £8,137 | Dominated |
| DSA19: Start age – 5 years | Dominated | Dominating | Dominating | Dominated | £33,567 | £4,379 | Dominated |
| DSA20: Utility values from Verrill *et al.*[160] | Dominated | Dominating | Dominating | Dominated | £44,393 | £6,172 | Dominated |
| DSA21: AML removed from model | Dominated | Dominating | Dominating | Dominated | £47,629 | £7,274 | Dominated |
| DSA22: Chemotherapy cost halved | Dominated | Dominating | Dominating | Dominated | £46,376 | £8,253 | Dominated |
| DSA23: Chemotherapy cost doubled | £5,007 (SWQ) | Dominating | Dominating | £10,361 (SWQ) | £27,908 | £235 | £29,702 (SWQ) |
| DSA24: DM lifetime cost halved | Dominated | Dominating | Dominating | £524 (SWQ) | £46,275 | £12,758 | £1,239 (SWQ) |
| DSA25: DM lifetime cost doubled | Dominated | Dominating | Dominating | Dominated | £28,111 | Dominating | Dominated |
| DSA26: AML costs halved | Dominated | Dominating | Dominating | Dominated | £41,175 | £6,066 | Dominated |
| DSA27: AML costs doubled | Dominated | Dominating | Dominating | Dominated | £38,311 | £4,608 | Dominated |

*BC - base case; ICER - incremental cost-effectiveness ratio; DSA - deterministic sensitivity analysis; N/a – not applicable; RS - recurrence score; DRFI - distant recurrence-free interval; UKBCG - UK Breast Cancer Group; CET - chemotherapy plus endocrine therapy; ET - endocrine therapy; HR - hazard ratio; QALY - quality-adjusted life year; AML - acute myeloid leukaemia; DM - distant metastases; SWQ - South-West quadrant*

**Table 44: Model-predicted incremental clinical and economic outcomes per 1,000 women tested – test versus current decision-making**

| Incremental model outcome (test versus current decision-making) | BC1 – Oncotype DX, RxPONDER pre-menopausal, predictive | BC2 – Oncotype DX, RxPONDER post-menopausal, predictive | BC3 – Oncotype DX, TransATAC post-menopausal, predictive | BC4 – Oncotype DX, TransATAC post-menopausal, non-predictive | BC5 – Prosigna, TransATAC post-menopausal, non-predictive | BC6 – EPclin, TransATAC post-menopausal, non-predictive | BC7 – MammaPrint, MINDACT, non-predictive |
|---|---|---|---|---|---|---|---|
| Number of women receiving chemotherapy | -361 | -594 | -491 | -491 | -46 | -39 | -370 |
| Number of infusion chair hours | -1,854 | -3,051 | -2,520 | -2,520 | -235 | -203 | -1,900 |
| Number of women experiencing DM during their lifetime | 41 | -13 | -2 | 46 | -3 | -8 | 24 |
| LYGs (undiscounted) | -650 | 214 | 81 | -447 | 59 | 122 | -217 |
| QALYs gained (discounted) | -178 | 113 | 53 | -171 | 28 | 55 | -66 |
| Additional costs to NHS/PSS (discounted) | £1,786,628 | -£4,282,569 | -£2,299,836 | £1,862,075 | £1,107,509 | £305,191 | £791,671 |
| Net health benefit (£20,000 per QALY gained) | -267 | 327 | 168 | -265 | -28 | 39 | -105 |
| Net health benefit (£30,000 per QALY gained) | -237 | 255 | 130 | -233 | -9 | 45 | -92 |

*BC - base case; DM - distant metastases; LYG - life year gained; QALY - quality-adjusted life year; NHS - National Health Service; PSS - Personal Social Services*

### 4.4 Discussion

The EAG undertook a systematic review of published economic evaluations of tumour profiling tests to guide adjuvant chemotherapy decisions in women with ER+, HER2-, LN+ early breast cancer. A total of 12 studies were included in the review, including five studies identified from the new searches and seven studies which were included in the previous systematic review by Harnan *et al.*[10] The economic models included in the review adopted similar structures based on a hybrid decision tree and state transition approach, built around three core health states which were defined according to the presence or absence of DM and survival status. Only one of the studies (Harnan *et al.*) included all four tumour profiling tests listed in the final NICE scope for this appraisal.

Two of the test manufacturers, Exact Sciences and Agendia, submitted model-based economic analyses to inform the appraisal. The structures of these models are broadly similar to the approaches used in the published economic analyses identified by the EAG's systematic review. The model of Oncotype DX provided by Exact Sciences presents separate base case analyses for: (i) pre-menopausal women with LN+ early breast cancer; (ii) post-menopausal women with LN+ early breast cancer and (iii) a blended analysis which reflects a mixed pre- and post-menopausal LN+ population. The model is informed by RxPONDER[28, 76] in women with an Oncotype DX RS of 0-25 and by external data (TransATAC[19] and SWOG-8814[31]) for women with an RS of >25. Pre- and post-test chemotherapy probabilities are based on an unpublished UK decision impact study on the use of Oncotype DX undertaken in women with LN+ early breast cancer.[17] All three base case analyses include an assumption that Oncotype DX is predictive of chemotherapy benefit, with different relative treatment effects for adjuvant chemotherapy versus ET applied to women who are low-risk (RS 0-25) and those who are high-risk (RS >25). The company's model suggests that Oncotype DX dominates current decision-making in post-menopausal women with LN+ disease and that Oncotype DX is dominated by current decision-making in pre-menopausal women with LN+ disease. Within the overall LN+ population, the model suggests that Oncotype DX dominates current decision-making; however, this analysis is misleading as it masks the cost-ineffectiveness of the test in the pre-menopausal subgroup.

The model provided by Agendia compares MammaPrint to other tumour profiling tests and usual decision-making across a range of populations, including women with LN0 disease. The company's analysis includes a separate scenario analysis which focuses on a pure LN+ subgroup. The company's analyses include an assumption that MammaPrint is predictive of chemotherapy benefit, based on the finding of a non-significant HR for DRFI for chemotherapy versus no chemotherapy for women who are clinical high-risk and MammaPrint low-risk, which was calculated through a re-analysis of IPD for women with HR+, HER2- disease (LN0 or LN+) from the MINDACT trial.[29] The company's submitted model suggests that MammaPrint dominates current decision-making in the LN+ subgroup. The EAG does not consider the company's assumption of predictive benefit to be a reasonable interpretation of

the results of their re-analysis of the MINDACT IPD. In addition, the EAG believes that the company's model likely overestimates the negative HRQoL impact of chemotherapy toxicity. The EAG also identified some programming errors which affect the model results. The EAG undertook a re-analysis of this model which removes the assumption of predictive benefit, down-weights the chemotherapy-related QALY loss and corrects the programming errors. This re-analysis suggests that MammaPrint leads to a small loss in survival, a small QALY gain and a small cost saving; hence MammaPrint remains dominant. However, the EAG has concerns that this model is still subject to some programming errors and notes that it does not include all of the EAG's preferred assumptions and evidence sources.

The EAG developed a *de novo* health economic model to assess the cost-effectiveness of Oncotype DX, Prosigna, EPclin and MammaPrint, each versus current decision-making. The economic analysis was undertaken from the perspective of the NHS and PSS and was largely based on the model structure used to inform NICE DG34.[10] The EAG's model adopts a hybrid decision tree and state transition structure. Key updates to the previous version of the EAG model include:

- The incorporation of data on test risk classification probabilities and DRFI from RxPONDER for the evaluation of Oncotype DX.[28, 76]
- Separate analyses for Oncotype DX to reflect assumptions that this test is or is not predictive of chemotherapy benefit based on both the older and newer RS cut-offs.
- Re-focusing the target population for MammaPrint to women who are clinically high-risk and who have LN+ early breast cancer.
- The incorporation of more up-to-date DRFI estimates from MINDACT for the evaluation of MammaPrint.[29]
- The incorporation of published analyses of TransATAC.[19]
- The incorporation of estimates of pre- and post-test chemotherapy use, based on Holt *et al.*[17] which are applied to all 2-level and 3-level tests.
- Updated estimates of the costs of adjuvant chemotherapy.
- Updated costing assumptions around the duration of ET, the proportion of post-menopausal women receiving bisphosphonates and the inclusion of ovarian suppression treatments for pre-menopausal women.
- Updated estimates of mortality risk and lifetime costs associated with treatments for DM, assuming first-line treatment with CDK4/6i therapy.
- Updated estimates of mortality risk, HRQoL and lifetime costs for people with secondary (therapy-related) AML.

The EAG's base case analyses suggest the following results:

- *Oncotype DX:* In the pre-menopausal LN+ population, Oncotype DX is dominated by current decision-making. This result is driven by the estimated reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment. In the post-menopausal LN+ population, Oncotype DX dominates current decision-making, providing the assumption of predictive benefit holds. As was the case with the economic analyses in the LN+ subgroup undertaken to inform DG34,[10] removing this assumption of predictive benefit results in a situation whereby Oncotype DX is dominated by current decision-making (based on the older RS cut-offs).
- *Prosigna:* The model suggests that the use of Prosigna will result in a small decrease in the use of chemotherapy, a small reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. The ICER for Prosigna versus current decision-making is expected to be £39,357 per QALY gained.
- *EPclin:* The model suggests that the use of EPclin will result in a small decrease in the use of chemotherapy, a reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. The ICER for EPclin versus current decision-making is expected to be £4,113 per QALY gained.
- *MammaPrint:* The model suggests that the use of MammaPrint will result in a large decrease in the use of chemotherapy in women who would have benefitted from it, an increase in the lifetime probability of developing DM and additional net costs due to the cost of the test. MammaPrint is dominated by current decision-making.

The EAG's model is subject to the following strengths:

- The economic analysis is in line with the NICE Reference Case[162] and relates specifically to the population under consideration within this appraisal.
- The model structure is consistent with the general approach used in most of the economic analyses included in the SLR and the two models submitted by the test manufacturers.
- Where data permit, risk classification probabilities and DRFI estimates for each test have been taken from same source. This approach maintains correlation between these parameters and avoids the potential for spectrum bias.
- For the analyses of Oncotype DX, the assumption of a predictive benefit of chemotherapy has been tested.
- Unlike the analyses presented to inform DG34,[13] the current EAG model applies pre- and post-test chemotherapy probabilities for all tests based on analyses of the same UK decision impact study of Oncotype DX evaluated using both the older 3-level and newer 2-level RS cut-offs (Holt *et al.*[17]).

- A broad range of DSAs have been undertaken to explore uncertainty around all key model inputs.
- The EAG's model and the Exact Sciences model suggest similar economic conclusions for Oncotype DX, Prosigna, EPclin. The Exact Sciences model suggests that MammaPrint has an ICER of more than £50,000 per QALY gained, whereas the EAG's model suggests that this test is dominated by current decision-making.

The EAG's economic analyses are also subject to several weaknesses, many of which stem from uncertainties and gaps in the available evidence:

- There remains some uncertainty around the extent to which Oncotype DX is predictive of chemotherapy benefit. As discussed in Section 3.5, tests for interaction between Oncotype DX RS and chemotherapy benefit on DFS in SWOG-8814[31] were statistically significant for some analyses, but not others. RxPONDER[28] indicates that chemotherapy is not beneficial to post-menopausal women who have an RS of 0-25. The test for interaction between the treatment group and the continuous RS in RxPONDER, when adjusted for the continuous RS, menopausal status, and treatment group, was not statistically significant within the range RS 0-25 ($p$=0.35). The other evidence identified from the EAG's review of predictive benefit does not consistently support or refute the assumption of predictive benefit (see Section 3.5.8). Therefore, the assumption of predictive benefit applied in the Exact Sciences model and the EAG's model is hinged on a clinically plausible assumption about the benefit of chemotherapy benefit in women with an Oncotype DX RS of >25, rather than empirical studies which statistically demonstrate this interaction across the full range of RS scores. The EAG's economic analyses highlight that the conclusions drawn from the model are strongly influenced by the inclusion of this assumption of predictive benefit. The need to draw on external evidence for women with an Oncotype DX RS of >25 from external sources also results in some inconsistency in terms of the cut-off used to characterise the Oncotype DX high-risk group (RxPONDER high-risk = RS >25; TransATAC high-risk = RS >31; SWOG-8814 high-risk = RS ≥31).
- The EAG's review of decision impact studies (see Section 3.6) did not identify any relevant studies for the use of Prosigna, EPclin or MammaPrint in the LN+ early breast cancer population. As such, the EAG's economic analyses use pre- and post-chemotherapy probabilities which are based on a decision impact study of Oncotype DX, defined either as a 2-level or 3-level test (Holt *et al.*[17]). This absence of relevant evidence means that the results of the analyses presented for each of these tests are highly uncertain and should be interpreted with some caution.
- It was only possible to present separate analyses of one test – Oncotype DX – by menopausal status. The analyses of EPclin and Prosigna are based on TransATAC[19] which was undertaken

in a post-menopausal population. EPclin is indicated for both pre-menopausal and post-menopausal women; however, there are insufficient data available to evaluate the use of the test in pre-menopausal women with LN+ disease. Prosigna is not indicated for use in pre-menopausal women. MammaPrint is indicated for both pre- and post-menopausal women; however, it was not possible to undertake separate analysis for these subgroups using the data from MINDACT.

- Owing to the use of different studies across the EAG's base case analyses, and the inclusion of overlapping but non-identical samples used between the tests included in TransATAC,[19] the EAG did not consider it appropriate to undertake indirect comparisons to compare tests incrementally.
- The EAG's model does not explicitly include the effect CHF on HRQoL which is a potential late effect of anthracycline-based chemotherapy. This event was also excluded from the two test manufacturers' models submitted to NICE and the previous EAG model used to inform DG34.[10] The EAG's clinical advisors commented that there is currently a shift away from anthracycline-based regimens in certain patients groups, including those with cardiac comorbidities, and they noted that oncologists are generally able to select out women who are likely to be at risk of CHF.
- Amongst pre-menopausal women, short-term or permanent amenorrhea is a common AE resulting from the use of chemotherapy. The impact of early menopause caused by chemotherapy is not explicitly captured in the EAG's model or the test manufacturers' models. The EAG was unable to identify relevant evidence which provides a quantitative estimate of the disutility associated with temporary or permanent infertility, the duration over which such a disutility might apply, or the proportion of women affected. These factors are complex and may be partly influenced by whether the woman already has children prior to starting chemotherapy and whether they are planning to have children after completing chemotherapy. In their response to clarification questions from the EAG,[101] Exact Sciences commented that the exclusion of this AE is a limitation of their economic analysis in the pre-menopausal LN+ subgroup and this limitation applies equally to the EAG's model. Other things being equal, the EAG's analysis of NHB (Table 44) indicates that any uncaptured negative health effects (e.g., infertility) would need to result in 0.24 to 0.27 QALYs lost per woman tested in order for Oncotype DX to achieve an ICER of £20,000 or £30,000 per QALY gained in the pre-menopausal subgroup (see Table 44). This is equivalent to an AE-related QALY loss of 0.69 to 0.78 QALYs per woman treated with adjuvant chemotherapy (calculated as the NHB shortfall divided by the proportion of women spared chemotherapy with tumour profile testing).

# 5 DISCUSSION AND CONCLUSIONS

## 5.1 Statement of principal findings

### *5.1.1 Clinical effectiveness – principal findings*

*Overview of evidence*

The search identified 4,057 articles. In total, 54 articles were included, 42 relating to prognostic and predictive ability, and 12 relating to impact on chemotherapy decisions. Studies of prognostic and predictive ability included prospective RCTs, retrospective reanalyses of trials and cohorts, and observational studies of prospective use of tests. Two prospective RCTs reported results: RxPONDER[28] for Oncotype DX and MINDACT[29] for MammaPrint. In RxPONDER,[28] patients with an Oncotype DX RS of ≤25 were randomised to chemotherapy vs. no chemotherapy. In RxPONDER, 65% of patients had 1 positive node, 25% had 2 positive nodes, and 9% had 3 positive nodes. The MINDACT study[29] assessed patients' genomic risk (via MammaPrint) and clinical risk (via mAOL). Patients who were low-risk on both measures were allocated to no chemotherapy, those who were high-risk on both were allocated to chemotherapy, and patients with discordant risk were randomised to chemotherapy vs. no chemotherapy. The ongoing OPTIMA RCT compares Prosigna test-directed chemotherapy use vs. standard chemotherapy use; however, results are not yet available.[33]

*Prognostic ability*

The prognostic ability of a test describes its ability to differentiate between patients with good versus poor outcomes. For all four tests, within re-analyses of trials and cohorts, the HR for distant recurrence between risk groups indicated statistically significant prognostic ability for most (though not all) analyses, both with and without adjustment for clinical factors. An analysis of the Clalit registry[64] reported that Oncotype DX was significantly prognostic for distant recurrence using both the RS <18 and >30 cut-offs and the RS <11 and >25 cut-offs, despite greater chemotherapy use in higher-risk patients. In the RxPONDER[28] prospective RCT, within the study population (RS 0-25), Oncotype DX was significantly prognostic for 5-year IDFS after adjusting for clinical factors, overall and in the pre-menopausal and post-menopausal subgroups. In the MINDACT RCT,[29] within LN+ patients at high clinical risk, 8-year DMFI was 92.3% for MammaPrint low-risk vs. 80.9% for MammaPrint high-risk, despite higher chemotherapy use for high-risk patients; however, no HRs or significance tests were reported for prognostic ability.

*Prediction of chemotherapy benefit: Oncotype DX*

Whether a test is predictive concerns whether the effect of chemotherapy vs. no chemotherapy on patient outcomes differs between test risk groups or ranges, generally assessed via an interaction test. Some data assessing predictive ability were identified for Oncotype DX and MammaPrint. No predictive data in a LN+ population were identified for Prosigna or EPclin.

In a reanalysis of the SWOG-8814 RCT,[31] Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy vs. no chemotherapy. For 10-year DFS, using cut-offs of RS <18 and >30, adjusted HRs indicated no effect of chemotherapy in the low-risk group (HR 1.02; 95% CI 0.54 to 1.93; *p*=0.97); a non-significant effect in the intermediate-risk group (HR 0.72; 95% CI 0.39 to 1.31; *p*=0.48); and a borderline statistically significant effect in the high-risk group (HR 0.59; 95% CI 0.35 to 1.01; *p*=0.033). Interaction tests for chemotherapy effect and risk group were statistically significant in some analyses but not others. The RxPONDER RCT[28] reported no benefit of chemotherapy in post-menopausal patients with RS 0-25 (difference in 5-year DRFI of 0.8% favouring no chemotherapy; adjusted HR 1.12; 95% CI 0.82 to 1.52; *p*=0.49). Conversely, there was chemotherapy benefit in pre-menopausal patients with RS 0-25 (difference of 2.4% favouring chemotherapy; adjusted HR 0.64; 95% CI 0.43 to 0.95; *p*=0.026). A test for interaction between RS (within the range 0-25) and effect of chemotherapy on IDFS was not statistically significant across all patients (HR 1.02; 95% 0.98 to 1.05; *p*=0.35) or in the pre-menopausal or post-menopausal subgroups, indicating no significant predictive effect within RS 0-25. Within registry data for Oncotype DX, the relationship between Oncotype DX risk group and effect of chemotherapy was unclear, and no interaction tests were reported. The NCDB database[69, 71, 72, 80] reported 5-year OS within post-menopausal or older subgroups with an RS of ≤25; some showed a statistically significant chemotherapy benefit while others did not; therefore the results did not clearly either support or refute the RxPONDER findings.

*Prediction of chemotherapy benefit: MammaPrint*

A reanalysis of two cohorts from 2009[32] reported a non-significant interaction test between MammaPrint score and effect of chemotherapy on BCSS (*p*=0.95) indicating no predictive effect. In the MINDACT[29] prospective RCT, within the clinical high-risk, MammaPrint low-risk, LN+, HR+ HER2- subgroup, 8-year DMFS was 91.2% with chemotherapy vs. 89.9% with no chemotherapy, an absolute difference of 1.3% favouring chemotherapy, with a non-significant HR (HR 0.84; 95% CI 0.51 to 1.37; *p*=NR). Since all patients in the clinical high-risk, MammaPrint high-risk group were offered chemotherapy, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit.

*Decision impact*

Evidence on chemotherapy decisions pre- and post-testing in LN+ populations included twelve studies of Oncotype DX (five in the UK and seven in other European countries). No decision impact studies were identified for EPclin, Prosigna or MammaPrint. The net change in the percentage of patients with a chemotherapy recommendation or decision (pre-test to post-test) was a reduction of 28% to 75% across five UK studies,[34-38] and a reduction of 12% to 73% across seven European studies.[39-45] Within

studies reporting data by Oncotype DX risk group, there were greater reductions in chemotherapy recommendation in the low-risk and intermediate-risk groups than in the high-risk groups.

*HRQoL and anxiety*

No studies reported HRQoL or anxiety associated with use of tumour profiling tests in a LN+ population. Across studies in a LN0 or mixed population, some reported significant improvements in anxiety after testing, while others reported no significant change. Some studies reported a decrease in anxiety after a low-risk test result or when treatment was downgraded to no chemotherapy, but an increase in anxiety after a high-risk test result or when treatment was upgraded to chemotherapy.

*Evidence on clinical subgroups*

The NICE scope[11] for this appraisal specified a number of patient subgroups. Data availability for these subgroups was as follows. For menopausal status, some subgroup data were available; in particular, the RxPONDER study indicated chemotherapy benefit in pre-menopausal patients with an RS of 0-25, but little chemotherapy benefit in post-menopausal patients with an RS of 0-25. For clinical risk, most studies did not subgroup patients by clinical risk, whilst the MINDACT study of MammaPrint reported separate data for people at high- or low-risk via mAOL (the low-mAOL subgroup was small for the LN+ population). No studies directly compared the genomic tests against clinical risk tools such as PREDICT, and the decision impact studies did not provide comparisons between genomic testing and specific clinical risk tools. In terms of sex, there were limited data in male-only subgroups or cohorts, though a subgroup analysis of the SEER database[67] reported significant prognostic ability of Oncotype DX in both men and women. In terms of ethnicity, one RxPONDER publication[49] reported that 5-year IDFS within RS 0-25 was slightly worse in black patients (87.0%) and slightly better in Asian patients (93.9%) compared with white patients (91.5%), but overall rates were similar, and no data were reported by ethnicity for prognostic or predictive ability. A subgroup analysis of the SEER database[65] reported statistically significant prognostic ability of Oncotype DX in white patients but non-significant results in black or other ethnicities, though these subgroups were based on small numbers. In terms of comorbidities, including people who may be affected by the side effects of chemotherapy, no specific clinical data were identified.

### *5.1.2 Cost-effectiveness – principal findings*

The EAG developed a *de novo* health economic model to assess the cost-effectiveness of Oncotype DX, MammaPrint, Prosigna, and EndoPredict (EPclin), each compared against current decision-making. The health economic analysis was undertaken from the perspective of the NHS and PSS and was largely based on the model developed to inform NICE DG34 in 2018, with updates to reflect changes in the breast cancer treatment pathway and updated evidence on the tests identified from the clinical effectiveness review. The EAG model adopts a hybrid decision tree/Markov structure. The model

parameters were informed by a number of sources, including the RxPONDER, TransATAC, SWOG-8814, and MINDACT trials, a recent unpublished UK decision impact study of Oncotype DX in LN+ women (Holt *et al.*), previous economic models, routine costing sources and other literature.
The results of the EAG's probabilistic base case analyses are summarised below.

*Oncotype DX*
Within the pre-menopausal LN+ population, Oncotype DX is dominated by usual care. These results are driven by the estimated reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment.

Within the post-menopausal LN+ subgroup, Oncotype DX dominates current decision-making, provided the assumption of predictive benefit holds. These results are driven by an estimated reduction in the use of adjuvant chemotherapy in women who would not have benefitted from treatment. As was the case with the economic analyses in the LN+ subgroup undertaken to inform DG34, removing this assumption of predictive benefit results in a situation whereby Oncotype DX is dominated by current decision-making. The assumption that Oncotype DX is predictive of chemotherapy benefit remains subject to some uncertainty and strongly influences the conclusions of the economic analysis in the post-menopausal subgroup.

*Prosigna*
The incremental cost-effectiveness ratio (ICER) for Prosigna versus current decision-making is expected to be £39,357 per quality-adjusted life year (QALY) gained. The model suggests that the use of Prosigna will result in a small decrease in the use of chemotherapy, a small reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. The EAG's systematic review did not identify any evidence to support a predictive benefit for Prosigna in the LN+ population.

*EndoPredict (EPclin)*
The ICER for EPclin versus current decision-making is expected to be £4,113 per QALY gained. The model suggests that the use of EPclin will result in a small decrease in the use of chemotherapy, a reduction in the lifetime probability of developing DM and additional net costs due to the cost of the test. The EAG's systematic review did not identify any evidence to support a predictive benefit for EPclin in the LN+ population.

*MammaPrint*
MammaPrint is dominated by current decision-making. These results are driven by the large estimated reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment, an increase in the lifetime probability of developing DM and additional net costs due to the cost of the

test. The EAG's systematic review did not identify sufficient evidence to support a predictive benefit for MammaPrint in the LN+ population.

## 5.2 Strengths and limitations of the assessment

### *5.2.1 Strengths and limitations in the clinical evidence base*

Strengths of the clinical evidence base include the fact that there is fairly substantial evidence for prognostic ability of all four tests. A major limitation is that it is difficult to collect new data on predictive ability because it is not considered ethical to randomise patients who are high-risk on any of the tests to chemotherapy vs. no chemotherapy. Therefore, although there are prospective RCTs for the effect of chemotherapy within low- to intermediate-risk patients, data for high-risk patients are limited to retrospective reanalyses of trials, plus observational data in which test results may have influenced treatment. Decision impact data in a LN+ population were available for Oncotype DX, but not for the other three tests. Anxiety and HRQoL data were not identified in a LN+ population.

### *5.2.2 Strengths and limitations relating to the health economic analysis*

The EAG's model is subject to several strengths. In particular: the economic analysis is consistent with the NICE Reference Case and relates specifically to the LN+ population under consideration within this appraisal; the model structure is consistent with most published economic models of tumour profiling tests as well as the two economic models submitted by the test manufacturers; where data permit, risk classification probabilities and DRFI estimates for each individual test have been taken from same source, which improves consistency and avoids the potential for spectrum bias; the analysis uses a recent UK decision impact study undertaken in LN+ women, and a broad assessment of uncertainty around all key model inputs has been presented, including testing assumptions around whether Oncotype DX is predictive of chemotherapy benefit. The EAG notes that under similar assumptions around the benefits of each tumour profiling test, the EAG's model and the Exact Sciences model indicate similar economic conclusions.

The EAG's economic analyses are subject to several weaknesses: the EAG's analyses of Oncotype DX based on RxPONDER indirectly assume a predictive benefit which reflects a plausible clinical assumption about the effect of chemotherapy in women who were excluded from the trial (those with an RS of >25), rather than a statistical test of interaction across the full RS spectrum; there are inconsistencies in RS cut-offs between sources used in the model; the analyses rely on a decision impact study of Oncotype DX to estimate post-test probabilities for all 2- and 3-level tests, which is highly uncertain; and there is insufficient evidence to allow for the economic analysis of EPclin and MammaPrint in an exclusively pre-menopausal subgroup. There is uncertainty around the potential negative effects of chemotherapy on infertility which may not be fully captured in the analyses of Oncotype DX in the pre-menopausal LN+ subgroup. The EAG's analyses of NHB provide a means for

the Appraisal Committee to decide whether any missing health effects are likely to impact on the conclusions drawn from the economic analysis.

### 5.3 Uncertainties

As was the case when NICE DG34 was undertaken, evidence relating to the impact on patient outcomes where the test is used in clinical practice remains largely absent, and is impeded by the long-term follow-up required, the large sample sizes required, and ethical problems associated with withholding chemotherapy from clinically high-risk patients.

Evidence relating to key subgroups defined in the scope is generally lacking. Where possible, separate data and analyses have been presented for pre-menopausal and post-menopausal women. Limited data were available by clinical risk subgroups as defined by risk assessment tools such as NPI or PREDICT. There were limited data in male-only subgroups or cohorts, and data relating to people of different ethnicities were difficult to interpret due to differences in treatment practices in different countries. No data were identified which could allow for a separate analysis of the value of tumour profiling tests in people with comorbidities who would be particularly affected by the adverse effects of chemotherapy.

There were no relevant decision impact studies on the use of MammaPrint, Prosigna or EPclin in a UK or European LN+ population. This remains a key area of uncertainty.

### 5.4 Generalisability

The economic analyses of EPclin and Prosigna are informed by the TransATAC trial which relates only to a post-menopausal population. It is expected that EPclin may also be used in pre-menopausal women. It was not possible to undertake separate economic analyses for MammaPrint or EPclin in a pre-menopausal LN+ population.

### 5.5 Implications for service provision

Oncotype DX, Prosigna and EPclin are already recommended by NICE for use in the NHS for women with ER+ (and/or PR+), HER2-, LN0 early breast cancer. The EAG's model suggests that all of the tumour profiling tests are expected to result in fewer women receiving adjuvant chemotherapy, thereby reducing costs and increasing capacity. However, for some of the tests, these initial benefits may lead to more women later requiring further treatment for DM, thereby offsetting cost savings and capacity reductions for chemotherapy services.

MammaPrint is not currently recommended for use in the NHS. MammaPrint testing can be undertaken either as an off-site service with samples sent to a laboratory in the US, or a through a decentralised testing service for laboratories with NGS capability. The per-sample pricing of MammaPrint remains the same regardless of where the testing is performed. Not all laboratories will have NGS capability

which will impact how testing services are delivered. For the other tests, only a single testing option is available – for Oncotype DX, samples are processed centrally, whereas for Prosigna and EPclin, samples are processed in local laboratories.

### 5.6 Suggested research priorities

Research priorities include the following:

- There remains some uncertainty around whether Oncotype DX is predictive of chemotherapy benefit. Further studies demonstrating a statistical interaction between Oncotype DX RS and long-term chemotherapy benefit across the full range of RS would help to address this uncertainty. However, such studies would require significant time and resources. Such studies may not be considered ethical as they may require chemotherapy to be withheld from some patients who are high-risk.
- The review of HRQoL studies did not identify any new relevant studies which quantify the negative impact of adjuvant chemotherapy. Future longer-term studies are required to estimate short-term toxicity as well as longer-term negative health effects, including temporary and permanent effects on fertility in pre-menopausal women. Such studies should include the use of a preference-based HRQoL instrument (e.g., the EQ-5D).
- The review did not identify any relevant decision impact studies for the use of Prosigna, EPclin or MammaPrint in a LN+ population. Further UK studies assessing the impact of tumour profiling tests on recommendations for adjuvant chemotherapy may help to reduce uncertainty around the cost-effectiveness of these tests.

# 6 REFERENCES


1. Cancer Research UK. Breast cancer incidence (invasive) statistics - breast cancer incidence by gender and UK country. Available from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-Zero (date accessed 07/08/2023).
2. Mayrovitz HN. Breast cancer. Brisbane, Australia: Exon Publications; 2022.
3. National Institute for Health and Care Excellence. NICE Guideline 101. Early and locally advanced breast cancer: diagnosis and management. NICE: London; 2018.
4. Chua B, Ung O, Taylor R, Boyages J. Frequency and predictors of axillary lymph node metastases in invasive breast cancer. *ANZ Journal of Surgery* 2001;71:723-8.
5. NHS Digital. Cancer survival in England, cancers diagnosed 2016 to 2020, followed up to 2021. Available from https://digital.nhs.uk/data-and-information/publications/statistical/cancer-survival-in-england/cancers-diagnosed-2016-to-2020-followed-up-to-2021 (date accessed 08/08/2023). 2023.
6. Cancer Research UK. Breast cancer incidence (invasive) statistics - breast cancer incidence by age. Available from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-One (date accessed 07/08/2023). 2023.
7. Cancer Research UK. Breast cancer mortality statistics - breast cancer mortality by age and UK country. Available from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/mortality#heading-Zero (date accessed 08/08/2023). 2023.
8. Rosen RD, Sapra A. TNM Classification. StatPearls. Available from https://www.ncbi.nlm.nih.gov/books/NBK553187/ (date accessed 08/08/2023). 2023.
9. Amin MB, Gress DM, Meyer Vega LR, Edge SB, Greene FL, Byrd DR, *et al.* AJCC cancer staging manual, Eighth edition: Springer; 2018.
10. Harnan S, Tappenden P, Cooper K, Stevens J, Bessey A, Rafia R, *et al.* Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer: a systematic review and economic analysis. *Health Technology Assessment* 2019;23:1-328.
11. National Institute for Health and Care Excellence. Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node-positive early breast cancer. Final scope. NICE: London; 2023.
12. Candido Dos Reis FJ, Wishart GC, Dicks EM, Greenberg D, Rashbass J, Schmidt MK, *et al.* An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. *Breast Cancer Research* 2017;19.
13. National institute for Health and Care Excellence. Diagnostics guidance 34. Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer. NICE: London; 2018.
14. National Institute for Health and Care Excellence. NICE Technology Appraisal Guidance 810. Abemaciclib with endocrine therapy for adjuvant treatment of hormone receptor-positive, HER2-negative, node-positive early breast cancer at high risk of recurrence. NICE: London; 2022.
15. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between different polychemotherapy regimens for early breast cancer: Meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet* 2012;379:432-44.
16. National Cancer Registration and Analysis Service. Bespoke data request - chemotherapy use amongst women with early breast cancer (data held on file). NCRAS: London; 2017.

17. Holt S (Peony Breast Cancer Unit submission to NICE). A UK prospective multicentre decision impact, decision conflict and economic evaluation of the 21-gene assay in women with node+ve, hormone receptor+ve, HER2-ve breast cancer. Manuscript provided as academic-in-confidence. 2023:1-42.

18. Ewertz M, Jensen AB. Late effects of breast cancer treatment and potentials for rehabilitation. *Acta Oncologica* 2011;50:187-93.

19. Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C*, et al.* Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: A secondary analysis of a randomized clinical trial. *JAMA Oncology* 2018;4:545-53.

20. Cancer Research UK. Number stages of breast cancer. Available from https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/number-stages (date accessed 09/03/2023). 2023.

21. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP*, et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *British Medical Journal* 2009;339.

22. Exact Sciences. Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer. Company evidence submission. Exact Sciences: London; 2023.

23. Agendia. MammaPrint: Reimbursement dossier United Kingdom. Agendia: Amsterdam; 2023.

24. Myriad. Additional information provided to NICE. Myriad: Nottingham; 2023.

25. Veracyte I. Additional information provided to NICE. Veracyte: San Francisco; 2023.

26. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I*, et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. B*ritish Medical Journal* 2019;366:l4898.

27. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS*, et al.* PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine* 2019;170:51-8.

28. Kalinsky K, Barlow WE, Gralow JR, Meric-Bernstam F, Albain KS, Hayes DF*, et al.* 21-gene assay to inform chemotherapy benefit in node-positive breast cancer. *New England Journal of Medicine* 2021;385:2336-47.

29. Piccart M, van 't Veer LJ, Poncet C, Lopes Cardozo JMN, Delaloge S, Pierga J-Y*, et al.* 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncology* 2021;22:476-88.

30. Ballman KV. Biomarker: Predictive or Prognostic? *Journal of Clinical Oncology* 2015;33:3968-71.

31. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT*, et al.* Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: A retrospective analysis of a randomised trial. *Lancet Oncology* 2010;11:55-65.

32. Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A*, et al.* The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study. *Breast Cancer Research and Treatment* 2009;116:295-302.

33. Stein RC, Makris A, MacPherson IR, Hughes-Davies L, Marshall A, Dotchin G*, et al.* Optima: Optimal personalised treatment of early breast cancer using multi-parameter analysis, an international randomized trial of tumor gene expression testdirected chemotherapy treatment in a largely node-positive population. *Journal of Clinical Oncology* 2021;39.

34. Battisti NML, Gutteridge E, Mylvaganam S, Parton M, Ring A, McGrath SE. The Oncotype DX Breast test to guide management of node-positive oestrogen receptor-positive HER2-negative breast cancer patients: the United Kingdom experience. *Breast* 2019;44:S17.

35. Holt SD, Sai-Giridhar P, Verrill M, Pettit L, Rigg A, Hickish T*, et al.* A UK prospective multicentre decision impact, decision conflict and economic evaluation of the use of Oncotype DX to guide chemotherapy in 680 women with hormone receptor positive, HER2 negative breast cancer and 1 to 3 nodes involved. *Cancer Research* 2023;83.

36. Loncaster J, Armstrong A, Howell S, Wilson G, Welch R, Chittalia A*, et al.* Impact of Oncotype DX breast Recurrence Score testing on adjuvant chemotherapy use in early breast cancer: Real world experience in Greater Manchester, UK. *European Journal of Surgical Oncology* 2017;43:931-7.

37. Malam Y, Rabie M, Geropantas K, Alexander S, Pain S, Youssef M. The impact of Oncotype DX testing on adjuvant chemotherapy decision making in 1-3 node positive breast cancer. *Cancer Reports* 2022;5:e1546.

38. Nanda A, Yarwood A, Dillon M, Roy PG. P120. Combined experience of utility of genomic profiling in lymph node-positive breast cancer: Reduced prescription of chemotherapy and follow-up. *European Journal of Surgical Oncology* 2021;47:e327.

39. Cognetti F, Masetti R, Fabi A, Bianchi G, Santini D, Rognone A*, et al.* PONDx: real-life utilization and decision impact of the 21-gene assay on clinical practice in Italy. *NPJ Breast Cancer* 2021;7:47.

40. Dieci MV, Guarneri V, Giarratano T, Mion M, Tortora G, De Rossi C*, et al.* First prospective multicenter Italian study on the impact of the 21-gene recurrence score in adjuvant clinical decisions for patients with ER positive/HER2 negative breast cancer. *The Oncologist* 2018;23:297-305.

41. Dieci MV, Guarneri V, Zustovich F, Mion M, Morandi P, Bria E*, et al.* Impact of 21-gene breast cancer assay on treatment decision for patients with T1-T3, N0-N1, estrogen receptor-positive/human epidermal growth receptor 2-negative breast cancer: Final results of the prospective multicenter ROXANE study. *The Oncologist* 2019;24:1424-31.

42. Fernandez-Perez I, Antolin Novoa S, De Paz Arias L, Vazquez Tunas L, Perez Lopez ME, Varela Ferreiro S*, et al.* 60P Real-life use of Oncotype DX Breast Recurrence Score test for the management of patients with node-negative and node-positive breast cancer in the autonomous community of Galicia (Spain). *Annals of Oncology* 2021;32:S46.

43. Llombart-Cussac A, Anton-Torres A, Rojas B, Andres R, Martinez N, Rodriguez CA*, et al.* Impact of the 21-gene assay in patients with high-clinical risk ER-positive and HER2-negative early breast cancer: Results of the KARMA Dx study. *Cancers* 2023;15.

44. Zambelli A, Simoncini E, Giordano M, La Verde N, Farina G, Torri V*, et al.* Prospective observational study on the impact of the 21-gene assay on treatment decisions and resources optimization in breast cancer patients in Lombardy: The BONDX study. *Breast* 2020;52:1-7.

45. Eiermann W, Rezai M, Kümmel S, Kühn T, Warm M, Friedrichs K*, et al.* The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. *Annals of Oncology* 2013;24:618-24.

46. Sestak I, Dowsett M, Cuzick J. NICE Request – TransATAC data analysis. 2017.

47. Mamounas EP, Tang G, Paik S, Baehner FL, Liu Q, Jeong J-H*, et al.* 21-gene recurrence score for prognosis and prediction of taxane benefit after adjuvant chemotherapy plus endocrine therapy: results from NSABP B-28/NRG Oncology. *Breast Cancer Research and Treatment* 2018;168:69-77.

48. Penault-Llorca F, Filleron T, Asselain B, Baehner FL, Fumoleau P, Lacroix-Triki M*, et al.* The 21-gene Recurrence Score R assay predicts distant recurrence in lymph node-positive, hormone

receptor-positive, breast cancer patients treated with adjuvant sequential epirubicin- and docetaxel-based or epirubicin-based chemotherapy (PACS-01 trial). *BMC Cancer* 2018;18:526.

49. Abdou Y, Barlow WE, Gralow JR, Meric-Bernstam F, Albain KS, Hayes DF*, et al.* Race and clinical outcomes in the RxPONDER trial (SWOG S1007). *Cancer Research* 2023;83.

50. Lopes Cardozo JMN, Drukker CA, Rutgers EJT, Schmidt MK, Glas AM, Witteveen A*, et al.* Outcome of patients with an ultralow-risk 70-gene signature in the MINDACT trial. *Journal of Clinical Oncology* 2022;40:1335-45.

51. Drukker CA, van Tinteren H, Schmidt MK, Rutgers EJ, Bernards R, van de Vijver MJ*, et al.* Long-term impact of the 70-gene signature on breast cancer outcome. *Breast Cancer Research and Treatment* 2014;143:587-92.

52. Vliek SB, Retel V, Drukker C, Bueno-De-Mesquita JM, Rutgers E, Van Tinteren H*, et al.* The 70-gene signature in node positive breast cancer: 10-year follow-up of the observational RASTER study. *Annals of Oncology* 2017;28:v60.

53. Jackisch C, Pronin D, Dimpfl T, Buttner R, Kunz G, Langwieder C*, et al.* MammaPrint 10-year follow up results from a German breast cancer cohort study. *Annals of Oncology* 2022;33:S612.

54. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z*, et al.* Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 risk of recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Annals of Oncology* 2014;25:339-45.

55. Filipits M, Nielsen TO, Rudas M, Greil R, Stöger H, Jakesz R*, et al.* The PAM50 Risk-of-Recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. *Clinical Cancer Research* 2014;20:1298-305.

56. Laenkholm A-V, Jensen M-B, Eriksen JO, Rasmussen BB, Knoop AS, Buckingham W*, et al.* PAM50 risk of recurrence score predicts 10-year distant recurrence in a comprehensive Danish cohort of postmenopausal women allocated to 5 years of endocrine therapy for hormone receptor-positive early breast cancer. *Journal of Clinical Oncology* 2018;36:735-40.

57. Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L*, et al.* Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study. *Breast Cancer Research and Treatment* 2016;156:81-9.

58. Martin M, Brase JC, Calvo L, Krappmann K, Ruiz-Borrego M, Fisch K*, et al.* Clinical validation of the EndoPredict test in node-positive, chemotherapy-treated ER+/HER2- breast cancer patients: results from the GEICAM 9906 trial. *Breast Cancer Research* 2014;16:R38.

59. Pu M, Messer K, Davies SR, Vickery TL, Pittman E, Parker BA*, et al.* Research-based PAM50 signature and long-term breast cancer survival. *Breast Cancer Research and Treatment* 2020;179:197-206.

60. Lundgren C, Bendahl P-O, Church SE, Ekholm M, Ferno M, Forsare C*, et al.* PAM50 subtyping and ROR score add long-term prognostic information in premenopausal breast cancer patients. *NPJ Breast Cancer* 2022;8:61.

61. Filipits M, Dubsky P, Rudas M, Greil R, Balic M, Bago-Horvath Z*, et al.* Prediction of distant recurrence using EndoPredict among women with ER+, HER2- node-positive and node-negative breast cancer treated with endocrine therapy only. *Clinical Cancer Research* 2019;25:3865-72.

62. Sestak I, Filipits M, Buus R, Rudas M, Balic M, Knauer M*, et al.* Prognostic value of EndoPredict in women with hormone receptor-positive, HER2-negative invasive lobular breast cancer. *Clinical Cancer Research* 2020;26:4682-7.

63. Constantinidou A, Marcou Y, Toss MS, Simmons T, Bernhisel R, Hughes E*, et al.* Clinical validation of EndoPredict in pre-menopausal women with ER-positive, HER2-negative primary breast cancer. *Clinical Cancer Research* 2022;28:4435-43.

64. Stemmer SM, Steiner M, Rizel S, Geffen DB, Nisenbaum B, Peretz T*, et al.* Clinical outcomes in ER+ HER2 -node-positive breast cancer patients who were treated according to the Recurrence Score results: evidence from a large prospectively designed registry. *NPJ Breast Cancer* 2017;3:32.

65. Petkov VI, Miller DP, Howlader N, Gliner N, Howe W, Schussler N*, et al.* Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. *NPJ Breast Cancer* 2016;2:16017.

66. Roberts MC, Miller DP, Shak S, Petkov VI. Breast cancer-specific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX Recurrence Score results in the SEER database. *Breast Cancer Research and Treatment* 2017;163:303-10.

67. Massarweh SA, Sledge GW, Miller DP, McCullough D, Petkov VI, Shak S. Molecular characterization and mortality from breast cancer in men. *Journal of Clinical Oncology* 2018;36:1396-404.

68. Ibraheem A, Olopade OI, Huo D. Propensity score analysis of the prognostic value of genomic assays for breast cancer in diverse populations using the National Cancer Data Base. *Cancer* 2020;126:4013-22.

69. Ibraheem AF, Press DJ, Olopade OI, Huo D. Community clinical practice patterns and mortality in patients with intermediate oncotype DX recurrence scores: Who benefits from chemotherapy? *Cancer* 2019;125:213-22.

70. Nash AL, Ren Y, Plichta JK, Rosenberger LH, van den Bruele AMB, DiNome ML*, et al.* Survival benefit of chemotherapy according to 21-gene recurrence score in young women with breast cancer. *Annals of Surgical Oncology* 2023;30:2130-9.

71. Weiser R, Polychronopoulou E, Hatch SS, Haque W, Ghani HA, He J*, et al.* Adjuvant chemotherapy in patients with invasive lobular carcinoma and use of the 21-gene recurrence score: A National Cancer Database analysis. *Cancer* 2022;128:1738-47.

72. Weiser R, Haque W, Polychronopoulou E, Hatch SS, Kuo Y-F, Gradishar WJ*, et al.* The 21-gene recurrence score in node-positive, hormone receptor-positive, HER2-negative breast cancer: a cautionary tale from an NCDB analysis. *Breast Cancer Research and Treatment* 2021;185:667-76.

73. Poorvu PD, Gelber SI, Rosenberg SM, Ruddy KJ, Tamimi RM, Collins LC*, et al.* Prognostic impact of the 21-gene recurrence score assay among young women with node-negative and node-positive ER-positive/HER2-negative breast cancer. *Journal of Clinical Oncology* 2020;38:725-33.

74. Nitz U, Gluz O, Christgen M, Kates RE, Clemens M, Malter W*, et al.* Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial. *Breast Cancer Research and Treatment* 2017;165:573-83.

75. Braun M, Kriegmair A, Szeterlak N, Andrulat A, Schrodi S, Kriner M*, et al.* Validation of the 21-gene recurrence score assay in patients with hormone receptor-positive, HER2-negative breast cancer and 0 to 3 positive lymph nodes: Risk pattern and outcomes on a community level. *Breast Care* 2022;17:288-95.

76. Kalinsky KM, Barlow WE, Gralow JR, Meric-Bernstam F, Albain KS, Hayes DF*, et al.* Updated results from a phase 3 randomized clinical trial in participants (pts) with 1-3 positive lymph nodes (LN), hormone receptor-positive (HR+) and HER2-negative (HER2-) breast cancer (BC) with recurrence score (RS) < 25 randomized to endocrine therapy (ET) +/- chemotherapy (CT): SWOG S1007 (RxPONDER). *Cancer Research* 2022;82.

77. Rotem O, Peretz I, Leviov M, Kuchuk I, Itay A, Tokar M*, et al.* Clinical outcomes in ER+ breast cancer patients with recurrence score 26-30-guided therapy: Real-world data. *Annals of Oncology* 2022;33:S607-S8.

78. Petkov VI, Kurian AW, Jakubowski DM, Shak S. Breast cancer-specific mortality (BCSM) in patients age 50 years or younger with nodepositive (N+) breast cancer (BC) treated based on the 21-gene assay in clinical practice. *Cancer Research* 2020;80.

79. Abel MK, Shui AM, Chien AJ, Rugo HS, Melisko M, Baehner F*, et al.* The 21-gene recurrence score in clinically high-risk lobular and ductal breast cancer: A National Cancer Database study. *Annals of Surgical Oncology* 2022;29:7739-47.

80. Cao L, Towe CW, Luo X, Stabellini N, Amin AL, Montero AJ. Adjuvant chemotherapy is associated with an overall survival benefit regardless of age in patients with ER+/HER2-breast cancer with 1-3 positive nodes and Oncotype DX recurrence score 20 to 25: A National Cancer Database analysis. *Journal of Clinical Oncology* 2022;40.

81. Iorgulescu JB, Freedman RA, Lester SC, Mittendorf EA, Brock JE. 21-gene recurrence score adds significant value for grade 3 breast cancers: Results from a national cohort. *JCO Precision Oncology* 2019;3.

82. Kumar PA, Wang D, Huang D, Sivapiragasam A. The impact of adjuvant chemotherapy on overall survival in hormone and node positive breast cancer patients with an Oncotype DX score of 25 or less. A NCDB analysis. *Cancer Research* 2023;83.

83. Cytel CEA report (produced on behalf of Agendia). Cost-effectiveness analysis of the MammaPrint gene profiling test to guide adjuvant chemotherapy use in early breast cancer in the United Kingdom, 2023.

84. Evans CN, Brewer NT, Vadaparampil ST, Boisvert M, Ottaviano Y, Lee MC*, et al.* Impact of genomic testing and patient-reported outcomes on receipt of adjuvant chemotherapy. *Breast Cancer Research and Treatment* 2016;156:549-55.

85. Lo SS, Mumby PB, Norton J, Rychlik K, Smerage J, Kash J*, et al.* Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. *Journal of Clinical Oncology* 2010;28:1671-6.

86. Retel VP, Groothuis-Oudshoorn CG, Aaronson NK, Brewer NT, Rutgers EJ, van Harten WH. Association between genomic recurrence risk and well-being among breast cancer patients. *BMC Cancer* 2013;13:295.

87. Fallowfield L, Matthews L, May S, Jenkins V, Bloomfield D. Enhancing decision-making about adjuvant chemotherapy in early breast cancer following EndoPredict testing. *Psycho-oncology* 2018;27:1264-9.

88. Martin M, Gonzalez-Rivera M, Morales S, de la Haba-Rodriguez J, Gonzalez-Cortijo L, Manso L*, et al.* Prospective study of the impact of the Prosigna assay on adjuvant clinical decision-making in unselected patients with estrogen receptor positive, human epidermal growth factor receptor negative, node negative early-stage breast cancer. *Current Medical Research Opinion* 2015;31:1129-37.

89. Wuerstlein R, Sotlar K, Gluz O, Otremba B, von Schumann R, Witzel I*, et al.* The West German Study Group Breast Cancer Intrinsic Subtype study: a prospective multicenter decision impact study utilizing the Prosigna assay for adjuvant treatment decision-making in estrogen-receptor-positive, HER2-negative early-stage breast cancer. *Current Medical Research Opinion* 2016;32:1217-24.

90. Berdunov V, Millen S, Paramore A, Hall P, Perren T, Brown R*, et al.* Cost-effectiveness analysis of the Oncotype DX Breast Recurrence Score test in node-positive early breast cancer. *Journal of Medical Economics* 2022;25:591-604.

91. Hinde S, Theriou C, May S, Matthews L, Arbon A, Fallowfield L*, et al.* The cost-effectiveness of EndoPredict to inform adjuvant chemotherapy decisions in early breast cancer. *Health Policy and Technology* 2019;8:75-83.

92. Masucci L, Torres S, Eisen A, Trudeau M, Tyono I, Saunders H*, et al.* Cost-utility analysis of 21-gene assay for node-positive early breast cancer. *Current Oncology* 2019;26:307-18.

93. Hall PS, Smith A, Hulme C, Vargas-Palacios A, Makris A, Hughes-Davies L*, et al.* Value of information analysis of multiparameter tests for chemotherapy in early breast cancer: The OPTIMA Prelim trial. *Value in Health* 2017;20:1311-8.

94. Stein RC, Dunn JA, Bartlett JM, Campbell AF, Marshall A, Hall P*, et al.* OPTIMA Prelim: A randomised feasibility study of personalised care in the treatment of women with early breast cancer. *Health Technology Assessment* 2016;20:xxiii-xxix, 1-201.

95. Hannouf MB, Xie B, Brackstone M, Zaric GS. Cost effectiveness of a 21-gene recurrence score assay versus Canadian clinical practice in post-menopausal women with early-stage estrogen or progesterone-receptor-positive, axillary lymph-node positive breast cancer. *PharmacoEconomics* 2014;32:135-47.

96. Blohmer JU, Rezai M, Kümmel S, Kühn T, Warm M, Friedrichs K*, et al.* Using the 21-gene assay to guide adjuvant chemotherapy decision-making in early-stage breast cancer: A cost-effectiveness evaluation in the German setting. *Journal of Medical Economics* 2013;16:30-40.

97. Lamond NW, Skedgel C, Rayson D, Lethbridge L, Younis T. Cost-utility of the 21-gene recurrence score assay in node-negative and node-positive breast cancer. *Breast Cancer Research and Treatment* 2012;133:1115-23.

98. Hall PS, McCabe C, Stein RC, Cameron D. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. *Journal of the National Cancer Institute* 2012;104:56-66.

99. Wong WB, Ramsey SD, Barlow WE, Garrison LP Jr, Veenstra DL. The value of comparative effectiveness research: projected return on investment of the RxPONDER trial (SWOG S1007). *Contemporary Clinical Trials* 2012;33:1117-23.

100. Vanderlaan BF, Broder MS, Chang EY, Oratz R, Bentley TG. Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. *American Journal of Managed Care* 2011;17:455-64.

101. Exact Sciences. Company's response to clarification questions from the EAG. Exact Sciences: London; 2023.

102. Ara R, Brazier JE. Populating an economic model with health state utility values: Moving toward better practice. *Value in Health* 2010;13:509-18.

103. Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P*, et al.* Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: A systematic review and cost-effectiveness analysis. *Health Technology Assessment* 2013;17:1-302.

104. Sledge GW Jr, Toi M, Neven P, Sohn J, Inoue K, Pivot X*, et al.* The effect of abemaciclib plus fulvestrant on overall survival in hormone receptor-positive, ERBB2-negative breast cancer that progressed on endocrine therapy-MONARCH 2: A randomized clinical trial. *JAMA Oncology* 2020;6:116-24.

105. Petrelli F, Borgonovo K, Cabiddu M, Lonati V, Barni S. Mortality, leukemic risk, and cardiovascular toxicity of adjuvant anthracycline and taxane chemotherapy in breast cancer: a meta-analysis. *Breast Cancer Research and Treatment* 2012;135:335-46.

106. Wolff AC, Blackford AL, Visvanathan K, Rugo HS, Moy B, Goldstein LJ*, et al.* Risk of marrow neoplasms after adjuvant breast cancer therapy: the national comprehensive cancer network experience. *Journal of Clinical Oncology* 2015;33:340-8.

107. National Institute for Health and Care Excellence. NICE Technology Appraisal No. 552. Liposomal cytarabine–daunorubicin for untreated acute myeloid leukaemia. Committee papers. NICE: London; 2018.

108. de Bock GH, Putter H, Bonnema J, van der Hage JA, Bartelink H, van de Velde CJ. The impact of loco-regional recurrences on metastatic progression in early-stage breast cancer: A multistate model. *Breast Cancer Research and Treatment* 2009;117:401-8.

109. Lidgren M, Wilking N, Jönsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Quality of Life Research* 2007;16:1073-81.

110. Campbell HE, Epstein D, Bloomfield D, Griffin S, Manca A, Yarnold J*, et al.* The cost-effectiveness of adjuvant chemotherapy for early breast cancer: A comparison of no chemotherapy and first, second, and third generation regimens for patients with differing prognoses. *European Journal of Cancer* 2011;47:2517-30.

111. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by age and sex for the UK. NICE Decision Unit: Sheffield; 2022.

112. Kurosky SK, Mitra D, Zanotti G, Kaye JA. Treatment patterns and outcomes of patients with metastatic ER+/HER-2- breast cancer: A multicountry retrospective medical record review. *Clinical Breast Cancer* 2018;18:e529-e38.

113. Zeidan AM, Mahmoud D, Kucmin-Bemelmans IT, Alleman CJ, Hensen M, Skikne B*, et al.* Economic burden associated with acute myeloid leukemia treatment. *Expert Review of Hematology* 2016;9:79-89.

114. Ellis P, Barrett-Lee P, Johnson L, Cameron D, Wardley A, O'Reilly S*, et al.* Sequential docetaxel as adjuvant chemotherapy for early breast cancer (TACT): An open-label, phase III, randomised controlled trial. *Lancet* 2009;373:1681-92.

115. NHS England. National Schedule of NHS Costs 2018/19. NHSE: London; 2020.

116. NHS England. National Schedule of NHS Costs 2019/20. NHSE: London; 2021.

117. Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT). Commercial Medicines Unit: London; 2021.

118. Joint Formulary Committee. British National Formulary (online). BMJ and Pharmaceutical Press: London; 2020.

119. Curtis LA, Burns A. Unit costs of health and social care. PSSRU: Kent; 2020.

120. Stemmer SM, Leviov M, Tokar M, Ben-Baruch N, Uziely B, Fried G*, et al.* Ten-year clinical outcomes in >1000 node-negative (N0) Estrogen Receptor (ER)+ breast cancer (BC) patients (Pts) where treatment decisions incorporated the Recurrence Score Results: a registry analysis using TAILORx categorization. *Breast* 2019;44:S102.

121. Genomic Health Inc. Oncotype DX breast recurrence score test - Instructions for Use. Genomic Health: California, USA; 2022.

122. Karnon J, Kerr GR, Jack W, Papo NL, Cameron DA. Health care costs for the treatment of breast cancer recurrent events: estimates from a UK-based patient-level analysis. *British Journal of Cancer* 2007;97:479-85.

123. Lancet JE, Uy GL, Cortes JE, Newell LF, Lin TL, Ritchie EK*, et al.* CPX-351 (cytarabine and daunorubicin) liposome for injection versus conventional cytarabine plus daunorubicin in older patients with newly diagnosed secondary acute myeloid leukemia. *Journal of Clinical Oncology* 2018;36:2684-92.

124. Agendia. Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer. Company's response to clarification questions from the EAG. Agendia: Amsterdam; 2023.

125. National institute for Health and Care Excellence. NICE Diagnostics Guidance 10. Gene expression profiling and expanded immunohistochemistry tests for guiding adjuvant

chemotherapy decisions in early breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat. NICE: London; 2013.

126. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF*, et al.* Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *New England Journal of Medicine* 2018;379:111-21.

127. NHS England. NHS England Oncotype DX Access Scheme dataset. NHSE: London; 2016.

128. Kuijer A, Straver M, den Dekker B, van Bommel ACM, Elias SG, Smorenburg CH*, et al.* Impact of 70-gene signature use on adjuvant chemotherapy decisions in patients with estrogen receptor-positive early breast cancer: Results of a prospective cohort study. *Journal of Clinical Oncology* 2017;35:2814-9.

129. Crolley VE, Marashi H, Rawther S, Sirohi B, Parton M, Graham J*, et al.* The impact of Oncotype DX breast cancer assay results on clinical practice: a UK experience. *Breast Cancer Research and Treatment* 2020;180:809-17.

130. Wang R, Zhu Y, Liu X, Liao X, He J, Niu L. The clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer. *BMC Cancer* 2019;19.

131. Edlin R, Connock M, Tubeuf S, Round J, Fry-Smith A, Hyde C*, et al.* Azacitidine for the treatment of myelodysplastic syndrome, chronic myelomonocytic leukaemia and acute myeloid leukaemia. *Health Technology Assessment* 2010;Suppl 1:69-74.

132. Geurts YM, Witteveen A, Bretveld R, Poortmans PM, Sonke GS, Strobbe LJA*, et al.* Patterns and predictors of first and subsequent recurrence in women with early breast cancer. *Breast Cancer Research and Treatment* 2017;165:709-20.

133. Younis T, Rayson D, Skedgel C. The cost-utility of adjuvant chemotherapy using docetaxel and cyclophosphamide compared with doxorubicin and cyclophosphamide in breast cancer. *Current Oncology* 2011;18:e288-e96.

134. Thomas RJ, Williams M, Marshall C, Glen J, Callam M. The total hospital and community UK costs of managing patients with relapsed breast cancer. *British Journal of Cancer* 2009;100:598-600.

135. Russell-Smith TA, Brockbank J, Mamolo C, Knight C. Cost effectiveness of gemtuzumab ozogamicin in the first-line treatment of acute myeloid leukaemia in the UK. *PharmacoEconomics Open* 2021;5:677-91.

136. Roché H, Fumoleau P, Spielmann M, Canon JL, Delozier T, Serin D*, et al.* Sequential adjuvant epirubicin-based and docetaxel chemotherapy for node-positive breast cancer patients: The FNCLCC PACS 01 trial. *Journal of Clinical Oncology* 2006;24:5664-71.

137. Martín M, Ruiz Simón A, Ruiz Borrego M, Ribelles N, Rodríguez-Lescure Á, Muñoz-Mateu M*, et al.* Epirubicin plus cyclophosphamide followed by docetaxel versus epirubicin plus docetaxel followed by capecitabine as adjuvant therapy for node-positive early breast cancer: Results from the GEICAM/2003-10 study. *Journal of Clinical Oncology* 2015;33:3788-95.

138. Aigner J, Marmé F, Smetanay K, Schuetz F, Jaeger D, Schneeweiss A. Nab-paclitaxel monotherapy as a treatment of patients with metastatic breast cancer in routine clinical practice. *Anticancer Research* 2013;33:3407-13.

139. Jones S, Holmes FA, O'Shaughnessy J, Blum JL, Vukelja SJ, McIntyre KJ*, et al.* Docetaxel with cyclophosphamide is associated with an overall survival benefit compared with doxorubicin and cyclophosphamide: 7-year follow-up of US Oncology Research Trial 9735. . *Journal of Clinical Oncology* 2009;27:1177-83.

140. Liu D, Wu J, Lin C, Ding S, Lu S, Fang Y*, et al.* The comparative safety of epirubicin and cyclophosphamide versus docetaxel and cyclophosphamide in lymph node-negative, HR-positive, HER2-negative breast cancer (ELEGANT): A randomized trial. *Cancers* 2022;14:3221.

141. NHS England. National schedule of NHS costs 2021/22. NHSE: London; 2023.

142. Commercial Medicines Unit. Drugs and pharmaceutical electronic market information tool (eMIT). Available from https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit; 2021.

143. Haymarket Media Group Ltd. Monthly Index of Medical Specialities (MIMS). Haymarket: Gloucester, UK; 2022.

144. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. Nuffield Trust: London; 2014.

145. Peasgood T, Ward SE, Brazier J, . Health-state utility values in breast cancer. *Expert Review of Pharmacoeconomics and Outcomes Research* 2010;10:553-66.

146. Tufts-New England Medical Center. The CEA Registry; 2010.

147. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Medical Care* 2000;38:583-637.

148. Jones K, Burns A. Unit costs of health and social care. PSSRU: Kent; 2021.

149. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S*, et al.* 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *New England Journal of Medicine* 2016;375:717-29.

150. Lyon AR, López-Fernández T, Couch LS, Asteggiano R, Aznar MC, Bergler-Klein J*, et al.* 2022 ESC guidelines on cardio-oncology developed in collaboration with the European Hematology Association (EHA), the European Society for Therapeutic Radiology and Oncology (ESTRO) and the International Cardio-Oncology Society (IC-OS). *European Heart Journal* 2022;43:4229-361.

151. Suri G, Mistry R, Young KC, Hettle R, May JR, Brixner D*, et al.* Cost effectiveness of ribociclib plus letrozole versus palbociclib plus letrozole for the treatment of post-menopausal women with hormone receptor-positive (HR+), human epidermal growth factor receptor 2-negative (HER2-) advanced or metastatic breast cancer from a us private third-party payer perspective. *Value in Health* 2017;20(9):A436.

152. Bewersdorf JP, Patel KK, Goshua G, Shallis RM, Podoltsev NA, Huntington SF*, et al.* Cost-effectiveness of liposomal cytarabine/daunorubicin in patients with newly diagnosed acute myeloid leukemia. *Blood* 2022;139:1766-70.

153. Office for National Statistics. National life tables: England. ONS: London; 2021.

154. Commercial Medicines Unit. Drugs and pharmaceutical electronic market information tool (eMIT). Available from https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (date accessed 09/03/2023). 2023.

155. Joint National Formulary. British National Formulary (online). BMJ Group and Pharmaceutical Press: London. Available from https://bnf.nice.org.uk/ (date accessed 09/03/2023); 2023.

156. Pan H, Gray R, Braybrooke J, Davies C, Taylor C, McGale P*, et al.* 20-year risks of breast cancer recurrence after stopping endocrine therapy at 5 years. *New England Journal of Medicine* 2017;377:1836-46.

157. Hortobagyi GN, Stemmer SM, Burris HA, Yap YS, Sonke GS, Hart L*, et al.* Overall survival with ribociclib plus letrozole in advanced breast cancer. *New England Journal of Medicine* 2022;386:942-50.

158. Finn RS, Boer K, Bondarenko I, Patel R, Pinter T, Schmidt M*, et al.* Overall survival results from the randomized phase 2 study of palbociclib in combination with letrozole versus letrozole alone for first-line treatment of ER+/HER2- advanced breast cancer (PALOMA-1, TRIO-18). *Breast Cancer Research and Treatment* 2020;183:419-28.

159. Martin PJ, Counts GW Jr, Appelbaum FR, Lee SJ, Sanders JE, Deeg HJ*, et al.* Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. *Journal of Clinical Oncology* 2010;28:1011-6.

160. Verrill M, Wardley AM, Retzler J, Smith AB, Bottomley C, Ni Dhochartaigh S*, et al.* Health-related quality of life and work productivity in UK patients with HER2-positive breast cancer: a cross-sectional study evaluating the relationships between disease and treatment stage. *Health & Quality of Life Outcomes* 2020;18:353.

161. Filipits M, Dubsky P, Rudas M, Greil R, Balic M, Fitzal F*, et al.* Prediction of distant recurrence using EndoPredict among women with ER-positive, HER2-negative breast cancer with a maximum follow-up of 16 years. *Cancer Research* 2019;79.

162. National Institute for Health and Care Excellence. NICE health technology evaluations: The manual. NICE: London; 2022.

163. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford University Press: New York; 2005.

164. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C*, et al.* Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value in Health* 2003;6:9-17.

# 7 APPENDICES

## Appendix 1: Literature search strategies

**Searches**

- Clinical effectiveness searches
- Cost effectiveness searches
- EQ-5D searches

**CLINICAL EFFECTIVENESS SEARCHES**

**Sources searched**

| Host | Database | Dates covered | Results |
|---|---|---|---|
| Ovid | Epub Ahead of Print, In-Process & Other Non-Indexed Citations, MEDLINE(R) Daily and MEDLINE(R) | 1946-Present | 1191 |
| Ovid | Embase | 1974-Present | 3184 |
| Wiley | Cochrane Database of Systematic Reviews (Cochrane Library) | 1996-Present | 132 |
| Wiley | Cochrane Central Register of Controlled Trials (Cochrane Library) | 1898-Present | 507 |
| INAHTA | INAHTA | 1989-Present | 77 |
| Clarivate | Web of Science Science Citation Index Expanded (1900-), Conference Proceedings Citation Index - Science (1990-) | 1900-Present | 1846 |
| NIH | ClinicalTrials.gov | | 58 |
| WHO | WHO International Clinical Trials Registry Platform | | 43 |
| | Total | | 7038 |
| | Unique records | | 4195 |

**Search strategies**

**Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to April 25, 2023**
21st April 2023
1191 records

| # | Searches | Results |
|---|---|---|
| 1 | exp Breast Neoplasms/ | 339439 |
| 2 | exp mammary neoplasms/ | 23367 |
| 3 | exp breast/ | 52644 |
| 4 | exp neoplasms/ | 3822842 |
| 5 | 3 and 4 | 32428 |
| 6 | (breast* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)).mp. | 471637 |
| 7 | (mammar* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)).mp. | 44846 |
| 8 | 1 or 2 or 5 or 6 or 7 | 495731 |
| 9 | (endopredict or epclin or "ep score").mp. | 150 |
| 10 | (mammaprint or 70-gene or "70 gene").mp. | 882 |
| 11 | (oncotype or "recurrence score" or 21-gene or "21 gene").mp. | 1967 |

| | | |
|---|---|---|
| 12 | (prosigna or pam50 or 50-gene or "50 gene").mp. | 805 |
| 13 | or/9-12 | 3400 |
| 14 | 8 and 13 | 1978 |
| 15 | limit 14 to yr="2017 -Current" | 1191 |

Search strategy adapted from Harnan et al., (2019) © Queen's Printer and Controller of HMSO 2019.

**Embase 1974 to 2023 Week 16**
21st April 2023
3184 records

| # | Searches | Results |
|---|---|---|
| 1 | breast tumor/ | 94110 |
| 2 | exp breast/ | 127654 |
| 3 | exp neoplasm/ | 5482710 |
| 4 | 2 and 3 | 82927 |
| 5 | (breast* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)).mp. | 748134 |
| 6 | (mammar* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)).mp. | 43926 |
| 7 | 1 or 4 or 5 or 6 | 766300 |
| 8 | (endopredict or epclin or "ep score").mp. | 390 |
| 9 | (mammaprint or 70-gene or "70 gene").mp. | 2013 |
| 10 | (oncotype or "recurrence score" or 21-gene or "21 gene").mp. | 4933 |
| 11 | (prosigna or pam50 or 50-gene or "50 gene").mp. | 2103 |
| 12 | or/8-11 | 8314 |
| 13 | 7 and 12 | 5481 |
| 14 | limit 13 to yr="2017 -Current" | 3184 |

Search strategy adapted from Harnan et al., (2019) © Queen's Printer and Controller of HMSO 2019.

**Cochrane (CDSR and CENTRAL)**
21st April 2023
639 records

| # | Searches | Results |
|---|---|---|
| #1 | MeSH descriptor: [Breast Neoplasms] explode all trees | 17635 |
| #2 | MeSH descriptor: [Neoplasms, Ductal, Lobular, and Medullary] explode all trees | 865 |
| #3 | MeSH descriptor: [Breast] explode all trees | 1428 |
| #4 | MeSH descriptor: [Neoplasms] explode all trees | 110452 |
| #5 | #3 and #4 | 563 |
| #6 | (breast* near/5 (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)) | 44803 |
| #7 | (mammar* near/5 (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)) | 354 |
| #8 | #1 or #2 or #5 or #6 or #7 | 45124 |
| #9 | (endopredict or epclin or "ep score") | 31 |

| | | |
|---|---|---|
| #10 | (mammaprint or "70 gene") | 138 |
| #11 | (oncotype or "recurrence score" or "21 gene") | 289 |
| #12 | (prosigna or pam50 or "50 gene").mp. | 19573 |
| #13 | #9 or #10 or #11 or #12 | 19949 |
| #14 | #8 and #13 with Cochrane Library publication date Between Jan 2017 and Jan 2023, in Cochrane Reviews, Trials | 639 |

Search strategy adapted from Harnan et al., (2019) © Queen's Printer and Controller of HMSO 2019.

**Web of Science Science Citation Index Expanded (1900-), Conference Proceedings Citation Index - Science (1990-)**
26th April 2023
1846 records

| # | Searches | Results |
|---|---|---|
| 1 | (breast* NEAR/5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)) (Topic) | 613,247 |
| 2 | (mammar* NEAR/5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)) (Topic) | 24,383 |
| 3 | #1 OR #2 | 626,070 |
| 4 | (endopredict OR epclin OR "ep score") (Topic) | 188 |
| 5 | (mammaprint OR 70-gene OR "70 gene") (Topic) | 1,676 |
| 6 | (oncotype OR "recurrence score" OR 21-gene OR "21 gene") (Topic) | 4,301 |
| 7 | (prosigna OR pam50 OR 50-gene OR "50 gene") (Topic) | 1,777 |
| 8 | #7 OR #6 OR #5 OR #4 | 7,417 |
| 9 | #8 AND #3 | 3,322 |
| 10 | #8 AND #3 and 2023 or 2022 or 2021 or 2020 or 2019 or 2018 or 2017 (Publication Years) | 1,846 |

Search strategy adapted from Harnan et al., (2019) © Queen's Printer and Controller of HMSO 2019.

**INAHTA HTA**
26th April 2023
77 records

| # | Searches | Results |
|---|---|---|
| 1 | endopredict | 7 |
| 2 | epclin | 1 |
| 3 | "ep score" | 231 |
| 4 | mammaprint | 21 |
| 5 | oncotype | 31 |
| 6 | "recurrence score" | 2 |
| 7 | prosigna | 13 |
| 8 | pam50 | 2 |
| 9 | breast* | 903 |
| 10 | mammar* | 9 |
| 11 | #10 OR #9 | 908 |
| 12 | "70-gene" | 371 |
| 13 | "70 gene" | 371 |
| 14 | "21-gene" | 371 |

| | | |
|---|---|---|
| 15 | "21 gene" | 371 |
| 16 | "50-gene" | 371 |
| 17 | "50 gene" | 371 |
| 18 | #17 OR #16 OR #15 OR #14 OR #13 OR #12 | 371 |
| 19 | #18 AND #11 | 58 |
| 20 | #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 | 272 |
| 21 | #20 OR #19 | 309 |
| 22 | ((pam50) OR (prosigna) OR ("recurrence score") OR (oncotype) OR (mammaprint) OR ("ep score") OR (epclin) OR (endopredict)) OR ((("50 gene") OR ("50-gene") OR ("21 gene") OR ("21-gene") OR ("70 gene") OR ("70-gene")) AND ((mammar*) OR (breast*))) 2017 to 2023 | 77 |

Search strategy adapted from Harnan et al., (2019) © Queen's Printer and Controller of HMSO 2019.

**WHO International Clinical Trials Registry Platform**
28th April 2023
43 records

| # | Searches | Results |
|---|---|---|
| 1 | endopredict OR epclin OR "ep score" | 7 |
| 2 | mammaprint OR 70-gene OR "70 gene" | 19 |
| 3 | oncotype OR "recurrence score" OR 21-gene OR "21 gene" | 35 |
| 4 | prosigna or pam50 or "50 gene" | 20 |
| 5 | or/1-4 (limit to 2017-present) | 43 |

**Clinicaltrials.gov**
28th April 2023
58 records

| # | Searches | Results |
|---|---|---|
| 1 | endopredict OR epclin OR "ep score" | 7 |
| 2 | mammaprint OR 70-gene OR "70 gene" | 26 |
| 3 | oncotype OR "recurrence score" OR 21-gene OR "21 gene" | 77 |
| 4 | prosigna or pam50 or "50 gene" | 2 |
| 5 | or/1-4 (limit to 2017-present) | 58 |

**Conference websites searches**

**American Society of Clinical Oncology (ASCO) https://www.asco.org/**
19th May 2023

| # | Searches | Results |
|---|---|---|
| 1 | endopredict | 2 |
| 2 | epclin | 3 |
| 3 | "ep score" | 9 |
| 4 | mammaprint | 160 |
| 5 | oncotype | 212 |
| 6 | "recurrence score" | 465 |
| 7 | prosigna | 14 |
| 8 | pam50 | 57 |

| | | |
|---|---|---|
| 9 | "70-gene" | 23 |
| 11 | "21-gene" | 57 |
| 13 | "50-gene" | 12 |

**European Society for Medical Oncology (ESMO) https://www.esmo.org/**
23rd May 2023

| # | Searches | Results |
|---|---|---|
| 1 | endopredict | 24 |
| 2 | epclin | 16 |
| 3 | "ep score" | 9 |
| 4 | mammaprint | 32 |
| 5 | oncotype | 57 |
| 6 | "recurrence score" | 63 |
| 7 | prosigna | 29 |
| 8 | pam50 | 90 |
| 9 | "70-gene" | 27 |
| 11 | "21-gene" | 32 |
| 13 | "50-gene" | 10 |

**American Association for Cancer Research (AACR) https://www.aacr.org/**
25th May 2023

| # | Searches | Results |
|---|---|---|
| 1 | endopredict | 0 |
| 2 | epclin | 0 |
| 3 | "ep score" | 0 |
| 4 | mammaprint | 3 |
| 5 | oncotype | 7 |
| 6 | "recurrence score" | 7 |
| 7 | prosigna | 0 |
| 8 | pam50 | 3 |
| 9 | "70-gene" | 2 |
| 11 | "21-gene" | 4 |
| 13 | "50-gene" | 2 |

**European Cancer Organization (ECO). https://www.europeancancer.org/**
25th May 2023

| # | Searches | Results |
|---|---|---|
| 1 | endopredict | 0 |
| 2 | epclin | 0 |
| 3 | ep score | 9 |
| 4 | mammaprint | 0 |
| 5 | oncotype | 0 |
| 6 | "recurrence score" | 1 |
| 7 | prosigna | 0 |
| 8 | pam50 | 0 |

| | | |
|---|---|---|
| 9 | 70-gene | 0 |
| 11 | "21-gene" | 0 |
| 13 | "50-gene" | 0 |

**Manufacturer website search**

**Myriad Genetics https://myriad.com/publications/**
1st June 2023
21 records

**Agendia https://agendia.com/**
30th May 2023
45 records

**Exact Sciences (aka Genomic Health) https://www.exactsciences.com/**
26th May 2023
5 records

**NanoString https://nanostring.com/**
26th May 2023
132 records

**COST EFFECTIVENESS SEARCHES**

**Sources searched**

| Host | Database | Dates covered | Results |
|---|---|---|---|
| Ovid | Epub Ahead of Print, In-Process & Other Non-Indexed Citations, MEDLINE(R) Daily and MEDLINE(R) | 1946-Present | 77 |
| Ovid | Embase | 1974-Present | 317 |
| Clarivate | Web of Science Science Citation Index Expanded (1900-), Conference Proceedings Citation Index - Science (1990-) | 1900-Present | 155 |
| | Total retrieved | - | 549 |
| | Unique records | | 404 |

**Search strategies**

**Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to May 03, 2023**
4th May 2023
77 records

| # | Searches | Results |
|---|---|---|
| 1 | exp Breast Neoplasms/ | 339611 |
| 2 | exp mammary neoplasms/ | 23370 |
| 3 | exp breast/ | 52667 |
| 4 | exp neoplasms/ | 3824859 |
| 5 | 3 and 4 | 32448 |

| | | |
|---|---|---|
| 6 | (breast* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)).mp. | 471982 |
| 7 | (mammar* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)).mp. | 44866 |
| 8 | 1 or 2 or 5 or 6 or 7 | 496080 |
| 9 | (endopredict or epclin or "ep score").mp. | 149 |
| 10 | (mammaprint or 70-gene or "70 gene").mp. | 880 |
| 11 | (oncotype or "recurrence score" or 21-gene or "21 gene").mp. | 1963 |
| 12 | (prosigna or pam50 or 50-gene or "50 gene").mp. | 805 |
| 13 | or/9-12 | 3396 |
| 14 | 8 and 13 | 1977 |
| 15 | limit 14 to yr="2017 -Current" | 1190 |
| 16 | exp "Costs and Cost Analysis"/ | 264079 |
| 17 | Economics/ | 27499 |
| 18 | exp Economics, Hospital/ | 25708 |
| 19 | exp Economics, Medical/ | 14388 |
| 20 | Economics, Nursing/ | 4013 |
| 21 | exp models, economic/ | 16199 |
| 22 | Economics, Pharmaceutical/ | 3101 |
| 23 | exp "Fees and Charges"/ | 31352 |
| 24 | exp Budgets/ | 14104 |
| 25 | budget*.tw. | 35158 |
| 26 | ec.fs. | 442581 |
| 27 | cost*.ti. | 142156 |
| 28 | (cost* adj2 (effective* or utilit* or benefit* or minimi*)).ab. | 188331 |
| 29 | (economic* or pharmacoeconomic* or pharmaco-economic*).ti. | 59859 |
| 30 | (price* or pricing*).tw. | 51979 |
| 31 | (financial or finance or finances or financed).tw. | 120944 |
| 32 | (fee or fees).tw. | 21211 |
| 33 | (value adj2 (money or monetary)).tw. | 2985 |
| 34 | quality-adjusted life years/ | 15581 |
| 35 | (qaly or qalys).af. | 14091 |
| 36 | (quality adjusted life year or quality adjusted life years).af. | 23657 |
| 37 | or/16-36 | 895282 |
| 38 | 15 and 37 | 77 |

**Embase 1974 to 2023 Week 17**
4th May 2023
317 records

| # | Searches | Results |
|---|---|---|
| 1 | breast tumor/ | 94109 |
| 2 | exp breast/ | 127715 |
| 3 | exp neoplasm/ | 5486388 |
| 4 | 2 and 3 | 82947 |

| | | |
|---|---|---|
| 5 | (breast* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)).mp. | 748784 |
| 6 | (mammar* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)).mp. | 43946 |
| 7 | 1 or 4 or 5 or 6 | 766964 |
| 8 | (endopredict or epclin or "ep score").mp. | 390 |
| 9 | (mammaprint or 70-gene or "70 gene").mp. | 2013 |
| 10 | (oncotype or "recurrence score" or 21-gene or "21 gene").mp. | 4934 |
| 11 | (prosigna or pam50 or 50-gene or "50 gene").mp. | 2102 |
| 12 | or/8-11 | 8316 |
| 13 | 7 and 12 | 5481 |
| 14 | limit 13 to yr="2017 -Current" | 3184 |
| 15 | exp breast tumor/ | 642405 |
| 16 | exp breast/ | 127715 |
| 17 | exp neoplasm/ | 5486388 |
| 18 | 16 and 17 | 82947 |
| 19 | Socioeconomics/ | 159524 |
| 20 | Cost benefit analysis/ | 93753 |
| 21 | Cost effectiveness analysis/ | 179610 |
| 22 | Cost of illness/ | 21158 |
| 23 | Cost control/ | 75866 |
| 24 | Economic aspect/ | 123726 |
| 25 | Financial management/ | 120747 |
| 26 | Health care cost/ | 222179 |
| 27 | Health care financing/ | 13847 |
| 28 | Health economics/ | 35524 |
| 29 | Hospital cost/ | 25189 |
| 30 | (fiscal or financial or finance or funding).tw. | 286082 |
| 31 | Cost minimization analysis/ | 3974 |
| 32 | (cost adj estimate*).mp. | 4184 |
| 33 | (cost adj variable*).mp. | 320 |
| 34 | (unit adj cost*).mp. | 5524 |
| 35 | or/19-34 | 1107927 |
| 36 | 14 and 35 | 317 |

**Web of Science Science Citation Index Expanded (1900-), Conference Proceedings Citation Index - Science (1990-)**
4th May 2023
155 records

| # | Searches | Results |
|---|---|---|
| 1 | (breast* NEAR/5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)) (Topic) | 614,142 |
| 2 | (mammar* NEAR/5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)) (Topic) | 24,399 |

| | | |
|---|---|---|
| 3 | #1 OR #2 | 626,970 |
| 4 | (endopredict OR epclin OR "ep score") (Topic) | 1,778 |
| 5 | (mammaprint OR 70-gene OR "70 gene") (Topic) | 4,308 |
| 6 | (oncotype OR "recurrence score" OR 21-gene OR "21 gene") (Topic) | 1,682 |
| 7 | (prosigna OR pam50 OR 50-gene OR "50 gene") (Topic) | 188 |
| 8 | #7 OR #6 OR #5 OR #4 | 7,430 |
| 9 | #8 AND #3 and 2023 or 2022 or 2021 or 2020 or 2019 or 2018 or 2017 (Publication Years) | 1,854 |
| 10 | TS=(cost* and (effective* or utilit* or benefit* or minimi*)) OR TS=(cost*) OR TI=(economic* or pharmacoeconomic* or pharmaco-economic*) OR TS=(price* or pricing*) OR TS=(financial or finance or finances or financed) OR TS=(fee or fees) OR TS=(value and (money or monetary)) OR TS=(economic*) OR TS=(economic* and (hospital or medical or nursing or pharmaceutical)) OR TS=("quality adjusted life year" or "quality adjusted life years") OR TS=(qaly or qalys) OR TS=(budget*) | 2,940,280 |
| 11 | #9 AND #10 | 155 |

**EQ-5D SEARCHES**

**Sources searched**

| Host | Database | Dates covered | Results |
|---|---|---|---|
| Ovid | Epub Ahead of Print, In-Process & Other Non-Indexed Citations, MEDLINE(R) Daily and MEDLINE(R) | 1946-Present | 139 |
| Ovid | Embase | 1974-Present | 391 |
| Clarivate | Web of Science Science Citation Index Expanded (1900-), Conference Proceedings Citation Index - Science (1990-) | 1900-Present | 139 |
| | | | |
| | Total retrieved | - | 669 |
| | Unique records | | 404 |

**Search Strategies**

**Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to May 03, 2023**
16th May 2023
139 records

| # | Searches | Results |
|---|---|---|
| 1 | exp Breast Neoplasms/ | 340146 |
| 2 | exp mammary neoplasms/ | 23375 |
| 3 | exp breast/ | 52745 |
| 4 | exp neoplasms/ | 3830835 |
| 5 | 3 and 4 | 32512 |
| 6 | (breast* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)).ti. | 256329 |
| 7 | (mammar* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* | 15265 |

| | | |
|---|---|---|
| | or lobular or medullar)).ti. | |
| 8 | 1 or 2 or 5 or 6 or 7 | 402600 |
| 9 | (euroqol or euro qol or eq5d or "eq 5d" or eq-5d).tw. | 16230 |
| 10 | 8 and 9 | 203 |
| 11 | limit 10 to yr="2017 -Current" | 139 |

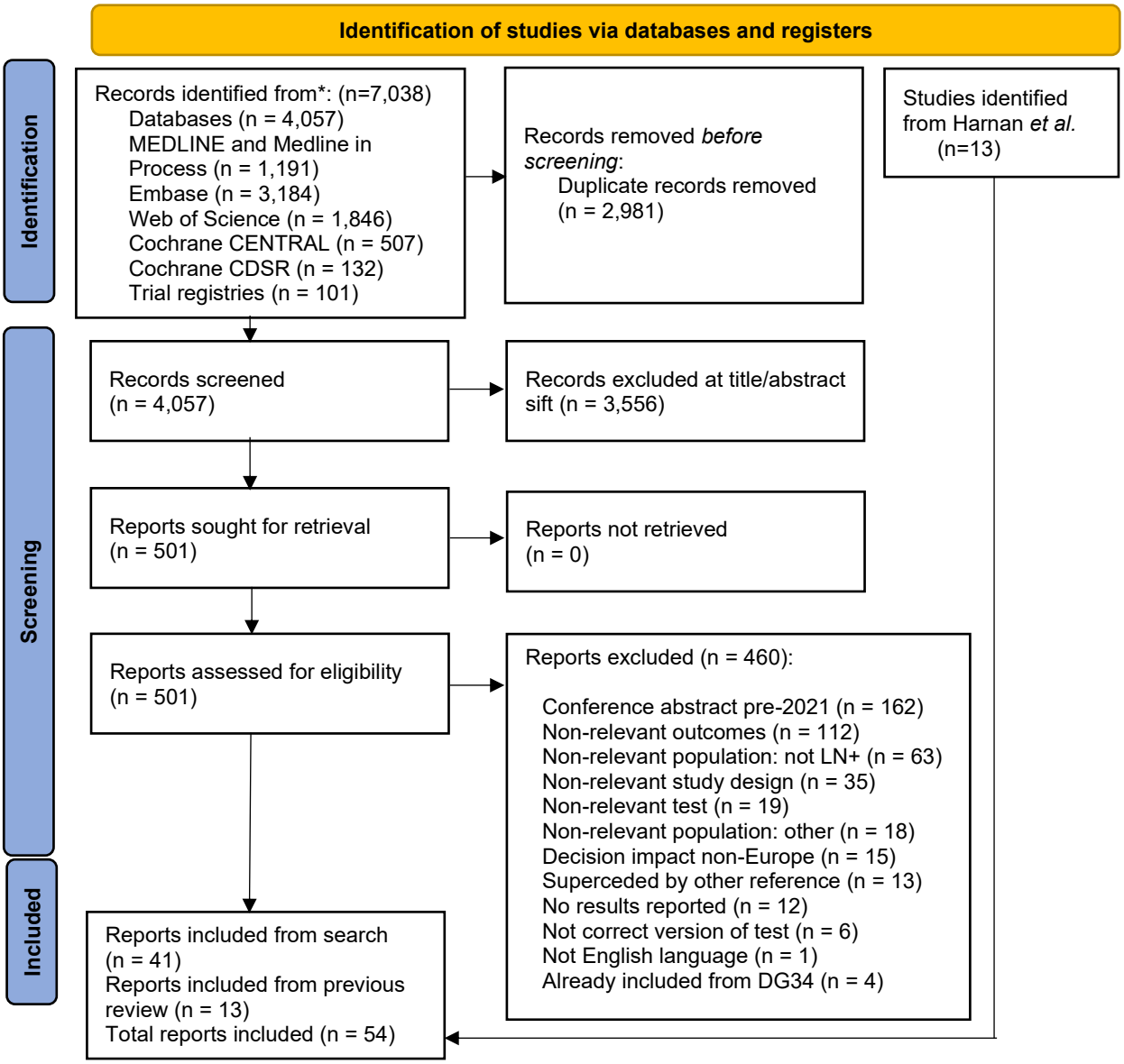
**Embase 1974 to 2023 Week 19**
16th May 2023
391 records

| # | Searches | Results |
|---|---|---|
| 1 | exp breast tumor/ | 643804 |
| 2 | exp breast/ | 127820 |
| 3 | exp neoplasm/ | 5497634 |
| 4 | 2 and 3 | 83006 |
| 5 | (breast* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)).ti. | 360376 |
| 6 | (mammar* adj5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar)).ti. | 16188 |
| 7 | 1 or 4 or 5 or 6 | 687386 |
| 8 | (euroqol or euro qol or eq5d or "eq 5d" or eq-5d).tw. | 29905 |
| 9 | 7 and 8 | 597 |
| 10 | limit 9 to yr="2017 -Current" | 391 |

**Web of Science Science Citation Index Expanded (1900-), Conference Proceedings Citation Index - Science (1990-)**
16th May 2023
139 records

| # | Searches | Results |
|---|---|---|
| 1 | ((breast* NEAR/5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary))) (Title) | 348,943 |
| 2 | ((mammar* NEAR/5 (neoplasm* or cancer* or tumo?r* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullar))) (Title) | 11,684 |
| 3 | #1 OR #2 | 359,016 |
| 4 | #1 OR #2 | 359,016 |
| 5 | (euroqol or euro qol or eq5d or eq 5d or "eq-5d") (Topic) | 19,973 |
| 6 | #4 AND #5 | 213 |
| 7 | #4 AND #5 and 2022 or 2023 or 2021 or 2020 or 2019 or 2018 or 2017 (Publication Years) | 139 |

## Appendix 2: PRISMA flow diagram for clinical studies




*From:* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

## Appendix 3: Risk of bias assessment

*Risk of bias assessment strategy*

Studies were assessed using risk of bias assessment tools relevant to the study design. Prospective RCTs were assessed using the Cochrane Risk of Bias tool (RoB2).[26] Prognostic and prediction studies were assessed using the Prediction model study Risk Of Bias Assessment Tool (PROBAST);[27] items from each domain were selected based on their relevance to this review, and definitions of high or low risk for each item specific to this review were defined a priori. Each study, cohort or registry was assessed once, rather than assessing each publication separately. Decision impact studies did not undergo formal quality assessment, but design and relevance were considered narratively. The impact of the quality of studies on the evidence base was considered within the narrative synthesis.

*Definition of items in PROBAST for this review*

For assessment of prognostic and prediction studies, items from each domain of PROBAST were selected based on their relevance to this review, and definitions of high or low risk for each item specific to this review were defined a priori, as shown in Table 45 below.

**Table 45: Risk of bias and applicability (adapted from PROBAST)**

| **Risk of bias** | | |
|---|---|---|
| **Domain** | **Criterion** | **Scoring for this review** |
| Domain 1 Participants | Were appropriate data sources used? | • Yes (prognosis): reanalysis of RCT or cohort or nested case control AND patients did not receive chemotherapy<br>• Yes (predicting chemotherapy benefit): RCT or reanalysis of RCT<br>• No (prognostic): non-nested case control or case series AND/OR some/all patients had chemotherapy<br>• No (predicting chemotherapy benefit): patients not randomised to chemotherapy vs. no chemotherapy |
| Domain 1 Participants | Were all inclusions and exclusions of participants appropriate? | • Yes: all eligible patients from trial or consecutive eligible patients from prospective registry<br>• No: some eligible patients excluded (e.g. not sent for testing, insufficient tissue, test failures, missing data, AND/OR non-prospective registry)<br>• Unclear: if unclear |
| Domain 2 Predictors [tests] | Were the tests [predictors] defined and assessed in a similar way for all participants? | • Yes: If test assessed in similar way for all participants [most/all studies in this review likely to score Yes as uses standardised test]<br>• No: Test not assessed in similar way for all participants |
| Domain 2 Predictors [tests] | Were the tests [predictor assessments] made without knowledge of outcome data? | • Yes: If test assessors blinded to clinical outcomes<br>• No: If not blinded<br>• Unclear: if unclear |
| Domain 3 Outcomes | Were the outcome definitions standardised or defined a priori? | • Yes: At least one outcome was standardised (e.g. DRFS, OS) or defined a priori<br>• No: All outcomes non-standardised and not defined a priori<br>• Unclear: if unclear |
| Domain 3 Outcomes | Were the outcomes determined without knowledge of test [predictor] information? | • Yes: If outcome assessors blinded to test results<br>• No: If not blinded<br>• Unclear: if unclear |

| **Risk of bias** | | |
|---|---|---|
| **Domain** | **Criterion** | **Scoring for this review** |
| Domain 3 Outcomes | Was chemotherapy decision made before test result known? | • Yes: Test did not influence use of chemotherapy [Yes if retrospective use of test on stored tumour samples, i.e. reanalyses of RCTs or cohorts]<br>• No: Test result may have influenced use of chemotherapy [No for observational studies of prospective use of test]<br>• [This item is not in PROBAST but is important for this review] |
| Domain 4 Analysis | Were there a reasonable number of participants with outcome data? | • Yes: At least 100 patients with outcome data<br>• No: Less than 100 patients with outcome data |
| Domain 4 Analysis | Were all enrolled participants included in the analysis? | • Yes: If all enrolled participants included in the analysis<br>• No: If some enrolled patients not analysed |
| **Applicability** | | |
| **Number** | **Criterion** | **Scoring for this review** |
| Domain 1 Participants | Did the included participants match the review question? | • Yes: all patients in scope (HR+, HER2–, LN1–3)<br>• Mostly: < 20% out of scope<br>• No: > 20% out of scope<br>• Unclear: if unclear |
| Domain 2 Predictors [tests] | Did the definition and assessment of tests [predictors] match the review question? | • Yes: same as commercially available tests<br>• No: different from commercially available tests (e.g. FFPE vs. fresh samples, test methods) |
| Domain 3 Outcomes | Did the outcomes match the review question? | • Yes: At least one outcome matched the review question<br>• No: No outcomes matched the review question |

*Results: risk of bias in prospective RCTs*

The risk of bias in the two prospective RCTs, assessed using the Cochrane RoB2 tool,[26] is shown in Table 46. The two RCTs scored low risk of bias on all domains, and low risk of bias overall.

**Table 46: Risk of bias in prospective RCTs (using Cochrane RoB2)**

| **RCT** | **Risk of bias due to …** | | | | | |
|---|---|---|---|---|---|---|
| | **Randomisation process** | **Deviations from intended interventions** | **Missing outcome data** | **Measurement of the outcome** | **Selection of the reported result** | **Overall risk of bias** |
| RxPONDER Kalinsky 2021[28] | Low | Low | Low | Low | Low | **Low** |
| MINDACT Piccart 2021[29] | Low | Low | Low | Low | Low | **Low** |

*Results: risk of bias in prognostic studies*

The risk of bias in prognostic studies, assessed using the PROBAST tool,[27] is presented in Table 47 for RCT reanalyses and cohort reanalyses (within which the test was used retrospectively), and in Table 48 for observational studies (within which the test was used prospectively).

The following factors may have affected results to some extent. For Domain 1 (participants), studies varied in terms of whether participants received chemotherapy or not; studies are therefore reported separately according to chemotherapy use in the section on prognostic ability (Section 3.4). In some studies, some participants did not match the review question (either not ER+, not HER2- or not LN1-3); these factors were taken into account when selecting studies for use in the economic model. Most studies excluded a proportion of patients for various reasons including insufficient tissue, missing data, failed tests and others, which may have influenced results to some extent, though the impact is difficult to assess. For Domain 3 (outcomes), chemotherapy decisions were not influenced by the test result in studies of retrospective use of the test (i.e., reanalyses of RCTs and cohorts), whereas in observational studies in which the test was used prospectively, chemotherapy decisions may have been influenced by the test result; therefore, observational studies are reported separately in the section on prognostic ability (Section 3.4).

The following factors were either judged low risk or were unlikely to have affected results. For Domain 2 (predictors, i.e., the tests themselves), all studies used the same version of the test for all participants (as the tests are standardised). Some studies blinded test assessors to patient outcomes while for other studies this was unclear; however, since the tests are based on objective measures of gene expression, this is unlikely to have affected interpretation of test results. For Domain 3 (outcomes), all studies used standardised outcomes relating to recurrence or survival. It was assumed that blinding of outcome assessors to test results applied within studies of retrospective use of the test, while in studies of prospective use, blinding to test results was generally unclear; however, as most outcomes were standardised cancer outcomes, this is unlikely to have affected outcome reporting. For Domain 4 (analysis), most studies included a reasonable number of participants (over 100). In terms of applicability to the review question, the test and outcomes matched the review question in all studies.

**Table 47: Risk of bias in prognostic studies (retrospective reanalyses of RCTs and cohorts)**

| Reference | Cohort | Design | Risk of Bias | | | | | | | | | Applicability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Domain 1 Participants | | Domain 2 Predictors | | Domain 3 Outcomes | | | Domain 4 Analysis | | Participants | Predictors | Outcomes |
| | | Derivation or validation? | Appropriate data sources? | Appropriate exclusions? | Tests same for all participants? | Blinded test assessors to outcomes? | Outcomes standardised or a priori? | Blinded outcome assessors to test? | CT decision made before test result known? | Participants N>100? | All analysed? | Participants match review question? | Tests match review question? | Outcomes match review question? |
| Albain 2010[31] | SWOG-8814 | RCT-R<br>V | Y (RCT-R, ET only) | N (InT, TF) | Y | Y | Y | Y | Y | Y | N | N (>20% LN4+) | Y | Y |
| Constantinidou 2022[63] | Cyprus + Notts | Cohort-R<br>V | Y (cohort-R, ET only) | N (InT, MD) | Y | Y | Y | Y | Y | N | N | Y | Y | Y |
| Drukker 2014[51] | VdV cohort , Netherlands | Cohort-R<br>V (21% also in derivation set) | N (cohort-R, some CT) | Y | Y | UC | Y | Y | Y | Y | Y | N (>20% ER- and >20% LN4+) | Y | Y |
| Filipits 2019[61] | ABCSG-6/8 | RCT-R<br>V | Y (RCT-R, ET only) | UC | Y | UC | Y | Y | Y | Y | Y | Y | Y | Y |
| Gnant 2014,[54] Filipits 2014[55] | ABCSG-8, Austria | RCT-R<br>V | Y (RCT-R, ET only) | N (InT, MS, TF, no consent) | Y | Y | Y | Y | Y | Y | N | Mostly (11% LN4+) | Y | Y |
| Jackisch 2022 (abst)[53] | Germany, PATH | Cohort-R<br>V | N (cohort-R, some CT) | N (reason NR) | Y | UC | Y | Y | Y | N | N | UC | Y | Y |
| Laenkholm 2018[56] | DBCG, Denmark | Cohort-R<br>V | Y (cohort-R, ET only) | N (FT, MD) | Y | Y | Y | Y | Y | Y | N | Y | Y | Y |
| Lundgren 2022[60] | SBII:pre trial | RCT-R<br>V | Y (RCT-R, ET only) | N (InT, FT, MD) | Y | N | Y | Y | Y | Y | Y | Y | Y | Y |
| Mamounas 2018[47] | NSABP-28 | RCT-R<br>V | N (RCT-R, all CT) | N (InT, MS) | Y | UC | Y | Y | Y | Y | Y | N (HER2 NR) | Y | Y |

| Reference | Cohort | Design | Risk of Bias | | | | | | | | | Applicability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Domain 1 Participants | | Domain 2 Predictors | | Domain 3 Outcomes | | | Domain 4 Analysis | | Participants | Predictors | Outcomes |
| | | Derivation or validation? | Appropriate data sources? | Appropriate exclusions? | Tests same for all participants? | Blinded test assessors to outcomes? | Outcomes standardised or a priori? | Blinded outcome assessors to test? | CT decision made before test result known? | Participants N>100? | All analysed? | Participants match review question? | Tests match review question? | Outcomes match review question? |
| Martin 2016,[57] Martin 2014[58] | GEICAM 9906, Spain | RCT-R<br>V | N (RCT-R, adjuvant CT) | N (MD) | Y | Y | Y | Y | Y | Y | N | N (>20% LN4+) | Y | Y |
| Mook 2009[32] | NKI and Italy | Cohort-R<br>V | N (cohort-R, some CT) | N (InT, RNA quality) | Y | Y | Y | Y | Y | Y | Y | N (>20% ER-, 16% HER2+) | Y | Y |
| Penault-Llorca 2018[48] | PACS01 | RCT-R<br>V | N (RCT-R, some CT) | N (FT, InT, MS) | Y | Y | Y | Y | Y | Y | Y | N (>20% LN4+) | Y | Y |
| Pu 2020[59] | WHEL Study | RCT-R | N (RCT-R, some CT) | N (InT, MS, TF) | Y | UC | Y | Y | Y | Y | N | UC (NR N nodes) | Y | Y |
| Sestak 2018;[19] 2017[46] | TransATAC | RCT-R<br>V | Y (RCT-R, ET only) | N (InT; FT ) | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Sestak 2020[62] | Lobular subgroup (TransATAC + ABCSG-6/8) | RCT-R<br>V | Y (RCT-R, ET only) | UC | Y | Y | Y | Y | Y | Y | N | Mostly (20% LN4+) | Y | Y |
| Vliek 2017[52] | RASTER | Cohort-R<br>V | N (cohort-R, some CT) | N (InT, MS, no consent) | Y | UC | Y | Y | Y | Y | Y | Mostly (17% ER-, 15% HER2+) | Y | Y |

*Cohort-R - reanalysis of cohort study; CT - chemotherapy; D - development study; ET - endocrine therapy; FT - failed test; InT - insufficient tissue; MD - missing data; MS - missing samples; N - no; NR - not reported; LN - number of positive lymph nodes; RCT-R - reanalysis of RCT; TF - test failure; UC - unclear; V - validation study; Y - yes*

**Table 48: Risk of bias in prognostic studies (observational studies of prospective use of test)**

| Reference | Cohort | Design | Risk of Bias | | | | | | | | | Applicability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Domain 1 Participants | | Domain 2 Predictors | | Domain 3 Outcomes | | | Domain 4 Analysis | | Participants | Predictors | Outcomes |
| | | Derivation or validation? | Appropriate data sources? | Appropriate exclusions? | Tests same for all participants? | Blinded test assessors to outcomes? | Outcomes standardised or a priori? | Blinded outcome assessors to test? | CT decision made before test result known? | Participants N>100? | All analysed? | Participants match review question? | Tests match review question? | Outcomes match review question? |
| Braun 2022[75] | Red Cross Hospital, Munich, Germany | Observational<br>V | N (prospective use of test, some CT) | Y | Y | Y | Y | UC | N | Y | N | Mostly (20% LNmic) | Y | Y |
| Ibraheem 2020[68] | NCDB | Observational<br>V | N (prospective use of test, CT) | N (MD, SFT) | Y | Y | Y | UC | N | Y | Y | Y | Y | Y |
| Massarweh 2018[67]<br>Petkov 2016[65]<br>Roberts 2017[66] | SEER | Observational<br>V | N (prospective use of test, some CT) | N (InT, MS, SFT, no consent) | Y | Y | Y | Y | N | Y | Y | UC (% LNmic NR) | Y | Y |
| Nitz 2017[74] | WSG PlanB | Observational<br>V | N (prospective use of test, some CT) | N (dropout, screening failure) | Y | Y | Y | UC | N | Y | Y | Y | Y | Y |
| Poorvu 2020[73] | Young Women's Breast Cancer Study | Observational<br>V | N (part prospective use of test, part stored samples, some CT) | N (InT, MS, no consent) | Y | Y | Y | UC | N | Y | N | UC (% LNmic NR) | Y | Y |

*CT - chemotherapy; D - development study; InT - insufficient tissue; MD - missing data; MS - missing samples; N - no; NCDB - National Cancer Database; NR - not reported; LN - number of positive lymph nodes; SFT - only those sent for test included; UC - unclear; V - validation study; WSG - West German Study Group; Y - yes*

*Results: risk of bias in prediction studies*

The risk of bias in prediction studies, assessed using the PROBAST tool,[27] is presented in Table 49.

The following factors may have affected results to some extent. For Domain 1 (participants), only the SWOG-8814 study[31] was a reanalysis of an RCT in which chemotherapy use was randomised; in the remaining studies, chemotherapy use was not randomised. This limitation is reflected in the section on prediction of chemotherapy benefit (Section 3.5). In some studies, some participants did not match the review question (either not ER+, not HER2- or not LN1-3). Most studies excluded a proportion of patients for various reasons including insufficient tissue, missing data, failed tests and others, which may have influenced results to some extent, though the impact is difficult to assess. For Domain 3 (outcomes), chemotherapy decisions were not influenced by the test result in the two studies of retrospective use of the test, whereas in the three observational registries in which the test was used prospectively, chemotherapy decisions may have been influenced by the test result; therefore, observational studies are reported separately in the section on prediction of chemotherapy benefit (Section 3.5).

The following factors were either judged low risk or were unlikely to have affected results. For Domain 2 (predictors, i.e., the tests themselves), all studies used the same version of the test for all participants (as the tests are standardised), and all studies blinded test assessors to patient outcomes. For Domain 3 (outcomes), all studies used standardised outcomes relating to recurrence or survival, and in all studies outcome assessors were blinded to test results. For Domain 4 (analysis), all studies included a reasonable number of participants (over 100). In terms of applicability to the review question, the test and outcomes matched the review question in all studies.

**Table 49: Risk of bias in prediction studies**

| Reference | Cohort | Derivation or validation? | Risk of Bias | | | | | | | | | Applicability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Domain 1 Participants | | Domain 2 Predictors | | Domain 3 Outcomes | | | Domain 4 Analysis | | Participants | Predictors | Outcomes |
| | | | Appropriate data sources? | Appropriate exclusions? | Tests same for all participants? | Blinded test assessors to outcomes? | Outcomes standardised or a priori? | Blinded outcome assessors to test? | CT decision made before test result known? | Participants N>100? | All analysed? | Participants match protocol? | Tests match review question? | Outcomes match review question? |
| Albain 2010[31] | SWOG-8814 | RCT-R<br>V | Y (RCT-R) | N (InT, TF) | Y | Y | Y | Y | Y | Y | N | N (>20% LN4+) | Y | Y |
| Mook 2009[32] | NKI and Italy | Cohort-R<br>V | N (not RCT) | N (InT, RNA qual) | Y | Y | Y | Y | Y | Y | Y | N (>20% ER-, 16% HER2+) | Y | Y |
| Abel 2022[79]<br>Cao 2022 (abst)[80]<br>Ibraheem 2019[69]<br>Iorgulescu 2019[81]<br>Kumar 2023 (abst)[82]<br>Nash 2023[70]<br>Weiser 2021[72]<br>Weiser 2022[71] | NCDB | Observational<br>V | N (not RCT) | N (MD, SFT) | Y | Y | Y | Y | N | Y | N | Y | Y | Y |
| Petkov 2020 (abst)[78] | SEER | Observational<br>V | N (not RCT) | N (InT, MS, SFT, no consent) | Y | Y | Y | Y | N | Y | Y | UC (% LNmic NR) | Y | Y |
| Rotem 2022 (abst)[77]<br>Stemmer 2017[64] | Clalit, Israel | Observational<br>V | N (not RCT) | N (SFT) | Y | Y | Y | Y | N | Y | N | N (>20% LNmic) | Y | Y |

*Cohort-R - reanalysis of cohort study; D - development study; InT - insufficient tissue; MD - missing data; MS - missing samples; NCDB - National Cancer Database; N - no; NR - not reported; LN - number of positive lymph nodes; LNmic - lymph node micrometastases; RCT-R - reanalysis of RCT; SFT - only those sent for test included; TF - test failure; UC - unclear; V - validation study; Y - yes*

## Appendix 4: Additional tables for prognostic ability

### Table 50: Prognostic data (Oncotype DX)

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| **Oncotype DX: Distant recurrence, ET monotherapy** | | | | | | | | | | | | | | | | |
| Sestak 2018;[19] 2017[46] TransATAC | DRFI | n=183 ET mono RCT-R | LN1-3 100% HR+ 100% HER2- | Post-meno | 18, 30 | 57 | 32 | 11 | 95.9 | 84.8 | 83.6 | 0-10y 80.6<br>5-10y 82.1 | 0-10y 70.9<br>5-10y 80.5 | 0-10y 62.0<br>5-10y 72.5 | 0-5yr: Int vs. low: HR 3.84 (1.31 to 11.23)<br>0-5yr: High vs. low: HR 4.45 (1.19 to 16.58)<br>0-10yr: Int vs. low: HR 1.66 (0.86 to 3.23)<br>0-10yr: High vs. low: HR 2.35 (0.99 to 5.60)<br>0-10yr: Per 1SD change: 1.39 (1.05-1.85)<br>*Adj: LR vs. CTS ($p$=0.06) and NPI ($p$ = 0.1) | Y<br>Y<br>N<br>N<br>Y<br>N* |
| **Oncotype DX: Distant recurrence, variable ET/CT** | | | | | | | | | | | | | | | | |
| Mamounas 2018[47] NSABP-28 | DRFI | n=722 All CT+ET RCT-R | LN1-3 100% ER+ NR HER2 | All meno | 18, 30 | 37 | 34 | 28 | - | - | - | 84.7 | 71.5 | 63.1 | 0-10yr: $p$<0.001<br>*0-10yr: Adj HR per 50-RS: 2.42 (NR); $p$<0.001 | Y<br>Y* |
| Penault-Llorca 2018[48] PACS01 | DRFI | n=530 All CT 74% ET RCT-R | LN1-3: 60% LN4+: 40% 100% HR+ 90% HER2- | All meno (39% post) | 18, 30 | 39 | 30 | 31 | 93.7 | 87.3 | 69.3 | - | - | - | 0-5yr: HR per 50-RS: 4.14 (2.67 to 6.43); $p$<0.001<br>*0-5yr: Adj HR 3.36 (1.88 to 6.00), $p$<0.001 | Y<br>Y* |
| **Oncotype DX: DFS** | | | | | | | | | | | | | | | | |
| Albain 2010[31] SWOG-8814 | DFS | n=148 ET mono RCT-R | LN+ 100% LN4+: 37% 100% HR+, 91% HER2– | Post-meno | 18, 30 | 37 | 31 | 32 | - | - | - | 60 | 49 | 43 | 0-5yr: HR 5.55 (2.32 to 3.28); $p$=0.0002<br>0-10yr: Between risk groups: $p$=0.017<br>0-10yr: HR per 50-RS: 2.64 (1.33 to 5.27); $p$=0.006<br>5-10yr: HR 0.86 (0.27 to 2.74); $p$=0.80 | Y<br>Y<br>Y<br>N |
| Mamounas 2018[47] NSABP-28 | DFS | n=722 All CT+ET RCT-R | LN1-3 100% ER+ NR HER2 | All meno | 18, 30 | 37 | 34 | 28 | - | - | - | 79.8 | 64.8 | 57 | 0-10yr: $p$<0.001<br>*0-5yr: Adj HR per 50-RS 3.81 (2.67 to 5.43); $p$<0.001<br>*0-10yr: Adj HR per 50-RS 2.53 (1.90 to 3.38); $p$<0.001<br>*5-10yr: Adj HR per 50-RS 1.39 (0.88 to 2.19); $p$=0.16 | Y<br>Y*<br>Y*<br>N* |

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| Penault-Llorca 2018[48] PACS01 | DFS | n=530 All CT 74% ET RCT-R | LN1-3: 60% LN4+: 40% 100% HR+ 90% HER2- | All meno (39% post) | 18, 30 | 39 | 30 | 31 | 90.8 | 84.9 | 64.6 | - | - | - | 0-5yr: HR per 50-RS: 3.28 (2.18 to 4.94); *p*<0.001<br>*0-5yr: Adj HR 2.66 (1.62 to 4.37), *p*<0.001 | Y<br>Y* |
| Kalinsky 2021[28] RxPONDER | IDFS | n=5,018 CT+ET vs. ET Prosp RCT | LN1-3 100% HR+ 100% HER2- | All meno (67% post) | All ≤25 | - | - | - | See prediction tables for outcomes per risk group | | | - | - | - | *0-5yr: HR per unit-RS (within RS 0-25): 1.05 (1.04 to 1.07), *p*<0.001 (adj meno and CT) | Y* |
| | | | | Post-meno | All ≤25 | - | - | - | - | - | - | - | - | - | *0-5yr: HR per unit-RS (within RS 0-25): 1.05 (1.03 to 1.07), *p*<0.001 (adj CT, nodes, grade, tumour size, age) | Y* |
| | | | | Pre-meno | All ≤25 | - | - | - | - | - | - | - | - | - | *0-5yr: HR per unit-RS (within RS 0-25): 1.06 (1.02 to 1.09), *p*=0.001 (adj CT, nodes, grade, tumour size, age) | Y* |
| Abdou 2023[49] RxPONDER | IDFS | n=4,015 CT+ET vs. ET Prosp RCT | LN1-3 100% HR+ 100% HER2- | White n=2,833 | All ≤25 | - | - | - | 91.5 | | | - | - | - | - | - |
| | | | | Black n=248 | All ≤25 | - | - | - | 87.0 | | | - | - | - | - | - |
| | | | | Asian n=324 | All ≤25 | - | - | - | 93.9 | | | - | - | - | - | - |
| | | | | Hispanic n=610 | All ≤25 | - | - | - | 91.4 | | | - | - | - | - | - |
| **Oncotype DX: OS and BCSS** | | | | | | | | | | | | | | | | |
| Albain 2010[31] SWOG-8814 | OS | n=148 ET mono RCT-R | LN+ 100% LN4+: 37% 100% HR+, 91% HER2– | Post-meno | 18, 30 | 37 | 31 | 32 | - | - | - | 77 | 68 | 51 | 0-10yr: Between risk groups: p=0.003<br>0-10yr: HR per RS-50: 4.42 (1.96 to 9.97), *p*=0.0006 | Y<br>Y |
| Penault-Llorca 2018[48] PACS01 | OS | n=530 All CT 74% ET RCT-R | LN1-3: 60% LN4+: 40% 100% HR+ 90% HER2- | All meno (39% post) | 18, 30 | 39 | 30 | 31 | 99 | 95.6 | 85.6 | - | - | - | 0-5yr: HR per 50-RS: 5.0 (3.01 to 8.28); *p*<0.001 | Y |

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| Mamounas 2018[47] NSABP-28 | OS | n=722 All CT+ET RCT-R | LN1-3 100% ER+ NR HER2 | All meno | 18, 30 | 37 | 34 | 28 | - | - | - | 93.3 | 79.2 | 70.7 | 0-10yr: p<0.001<br>*0-10yr: Adj HR per 50-RS: 3.09 (CI NR); *p*<0.001 | Y<br>Y* |
| Mamounas 2018[47] NSABP-28 | BCSS | n=722 All CT+ET RCT-R | LN1-3 100% ER+ NR HER2 | All meno | 18, 30 | 37 | 34 | 28 | - | - | - | 98 | 82.9 | 75.6 | 0-10yr: p<0.001<br>*0-10yr: Adj HR per 50-RS: 3.38 (CI NR); *p*<0.001 | Y<br>Y* |

*[a]The last column indicates whether each hazard ratio between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*
*Adj - adjusted; BCSS - breast cancer-specific survival; CI - confidence interval; CT - chemotherapy; CTS - Clinical Treatment Score (set of clinical factors); DFS - disease-free survival; DRFI - distant recurrence-free interval; ER - oestrogen receptor; ET - endocrine therapy; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor; IDFS - invasive disease-free survival; int - intermediate; LN - lymph nodes (number positive); LR - likelihood ratio; meno - menopausal; NPI - Nottingham Prognostic Index; NR - not reported; OS - overall survival; prosp - prospective; RCT - randomised controlled trial; RCT-R - RCT reanalysis; RS - Recurrence Score (Oncotype DX); SD - standard deviation; sig - significant; y/yr - year*

**Table 51: Prognostic data (MammaPrint)**

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| **MammaPrint: Distant recurrence, ET monotherapy** | | | | | | | | | | | | | | | | |
| No studies | | | | | | | | | | | | | | | | |
| **MammaPrint: Distant recurrence, variable ET/CT** | | | | | | | | | | | | | | | | |
| Piccart 2021[29] MINDACT (Not on prognostics summary table since CT use per risk group was influenced by test result) | DMFS | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | - | 31 | 95.7 (50% CT) | - | 89.0 (all CT) | 8y 91.0 (50% CT) | - | 8y 79.1 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | - | 8 | 96.3 (no CT) | - | - | 8y 94.0 (no CT) | - | - | - | - |
| | DMFI | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | - | 31 | 96.3 (50% CT) | - | 89.3 (all CT) | 8y 92.3 (50% CT) | - | 8y 80.9 (all CT) | - | - |

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | - | 8 | 97.5 (no CT) | - | - | 8y 95.2 (no CT) | - | - | - | - |
| Lopes Cardozo 2022[50] MINDACT | DMFI | N=201 (ultra-low) Var ET/CT Prosp-RCT | LN1-3 99% ER+ 97% HER2- | - | >0.355 ultra-low | Ultra-low: 15 | - | - | Ultra-low: 97.4 | - | - | 8y Ultra-low: 95.2 | - | - | - | - |
| Drukker 2014[51] VdV cohort , Netherlands | DMFS | n=144 Var ET/CT Cohort-R | LN1-3: 74% LN4+: 26% 77% ER+ NR HER2 | Age <53y | 0.4 | 38 | - | 62 | 94.5 | - | 64.7 | 10y 78.6 25y NE | - | 10y 54.3 25y 44.5 | 0-25 yr: HR 2.24 (1.25 to 4.00); *p*=0.01 | Y |
| Mook 2009[32] NKI and Italy | DMFS | n=241 Var ET/CT Cohort-R | LN1-3: 100% inc micromets 79% ER+ 84% HER2- | All meno | NR | 41 | - | 59 | 98 | - | 80 | 91 | - | 76 | 0-10 yr: HR 4.13 (1.72 to 9.96); *p*=0.002<br>*0-10 yr: Adj HR: 2.99 (0.996 to 8.99); *p*=0.051 | Y<br>N* |
| Vliek 2017[52] RASTER | DRFI | N=134 Var ET/CT Cohort-R | LN1-3 83% ER+ 85% HER2- | All ages | NR | 48 | - | 52 | 98.4 | - | 86.9 | 94.9 | - | 80.7 | 0-10 yr: Low vs high: HR 4.7 (1.3 to 16.2); *p*=0.008 | Y |
| | | | | All ages High mAOL (n=109) | NR | 40 | - | 60 | 97.7 | - | 86.1 | 95.2 | - | 79.5 | 0-10 yr: Low vs high: HR 4.8 (1.1 to 21.4), *p*=0.022 | Y |
| **MammaPrint: DFS** | | | | | | | | | | | | | | | | |
| Piccart 2021[29] MINDACT | DFS | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | - | 31 | 91.6 (50% CT) | - | 85.9 (all CT) | 8y 84.5 (50% CT) | - | 8y 74.5 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | - | 8 | 92.6 (no CT) | - | - | 8y 85.6 (no CT) | - | - | - | - |

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| **MammaPrint: OS and BCSS** | | | | | | | | | | | | | | | | |
| Piccart 2021[29] MINDACT | OS | n=1,176 CT+ET vs. ET Prosp-RCT | LN1-3 100% HR+ 100% HER2- | High mAOL (n=989) | >0 low, ≤0 high | 69 | - | 31 | 98.3 (50% CT) | - | 95.8 (all CT) | 8y 95.1 (50% CT) | - | 8y 89.1 (all CT) | - | - |
| | | | | Low mAOL (n=187) | >0 low, ≤0 high | 92 | - | 8 | 98.1 (no CT) | - | - | 8y 98.1 (no CT) | - | - | - | - |
| Drukker 2014[51] VdV cohort, Netherlands | OS | n=144 Var ET/CT Cohort-R | LN1-3: 74% LN4+: 26% 77% ER+ NR HER2 | Age <53y | 0.4 | 38 | - | 62 | 98.2 | - | 76.9 | 10y 92.5 25y 42.2 | - | 10y 58.7 25y 47.1 | 0-25 yr: HR 1.83 (1.07 to 3.11), *p*=0.03 | Y |
| Jackisch 2022 (abst)[53] Germany, PATH | OS | n=38 Var ET/CT Cohort-R | LN+ NR | All meno (assumed) | NR | 53 | - | 47 | - | - | - | 93.3 | - | 40.4 | - | - |
| Mook 2009[32] NKI and Italy | BCSS | n=241 Var ET/CT Cohort-R | LN1-3: 100% inc micromets 79% ER+ 84% HER2- | All meno | NR | 41 | - | 59 | 99 | - | 88 | 96 | - | 76 | 0-10 yr: HR 5.70 (2.01 to 16.23), *p*=0.001<br>*0-10 yr: Adj HR: 7.17 (1.81 to 28.43), *p*=0.005 | Y<br>Y* |
| | | | | All meno High AOL (n=209) | NR | - | - | - | - | - | - | 94 | - | 76 | 0-10 yr: HR 4.12 (1.45 to 11.76); *p*=0.008 | Y |

[a]*The last column indicates whether each hazard ratio between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*

*Adj - adjusted; AOL - Adjuvant! Online; BCSS - breast cancer-specific survival; CI - confidence interval; cohort-R - cohort reanalysis; CT - chemotherapy; DFS - disease-free survival; DMFI - distant metastasis-free interval; DMFS - distant metastasis-free survival; DRFI - distant recurrence-free interval; ER - oestrogen receptor; ET - endocrine therapy; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor; int - intermediate; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; OS - overall survival; prosp - prospective; RCT - randomised controlled trial; sig - significant; var - variable; y/yr - year*

**Table 52: Prognostic data (Prosigna)**

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| **Prosigna: Distant recurrence, ET monotherapy** | | | | | | | | | | | | | | | | |
| Sestak 2018;[19] 2017[46] TransATAC | DRFI | n=183 ET mono RCT-R | LN1-3 100% HR+ 100% HER2- | Post-meno | NR; assume 16, 40 | 8 | 32 | 60 | 100 | 91.7 | 87.4 | 0-10y 100<br>5-10y 100 | 0-10y 79.3<br>5-10y 87.0 | 0-10y 69.3<br>5-10y 75.0 | 0-5yr: Int vs. high: HR 1.30 (0.47 to 3.60)<br>0-10yr: Int vs. high: HR 1.37 (0.69 to 2.72)<br>HR per 1SD change: 1.58 (1.16-2.15)<br>*LR vs. CTS (p=0.04) and NPI (*p* = 0.09) | N<br>N<br>Y<br>Y, N* |
| Gnant 2014,[54] Filipits 2014[55] ABCSG-8, Austria | DMFS | n=413 ET mono RCT-R | LN1-3: 89% LN4+: 11% 100% ER+ 100% HER2- | Post-meno | 16, 40 | 4 | 34 | 62 | - | - | - | 0-10y 100<br>5-15y 100 | 0-10y 93.6<br>5-15y 87.0 | 0-10y 76.1<br>5-15y 75.0 | 5-15yr: Low risk: No events<br>5-15yr: Int vs. high: HR 3.15 (1.20 to 8.24); *p*=0.020<br>*0-10yr: Prognostic over clinical factors (*p*<0.0001)<br>*5-15yr: Prognostic over clinical factors (*p*=0.003) | -<br>Y<br>Y*<br>Y* |
| Laenkholm 2018[56] DBCG, Denmark | DRFS | n=1,395 ET mono Cohort-R | LN1-3 100% HR+ 100% HER2- | Post-meno | Bespoke Varies by N nodes | 26 | 28 | 46 | - | - | - | 96.5 | 88.5 | 77.9 | 0-10yr: Unadj: p<0.001<br>*0-10yr: Low vs. int: Adj HR 0.39 (0.20 to 0.77)<br>*0-10yr: High vs. int: Adj HR 1.54 (1.04 to 2.26), *p*<0.001 | Y<br>Y*<br>Y* |
| | | | | | 40 only | - | - | - | - | - | - | 95.2 (low to int) | | 78.1 | - | - |
| **Prosigna: Distant recurrence, variable ET/CT** | | | | | | | | | | | | | | | | |
| Martin 2016,[57] Martin 2014[58] GEICAM 9906, Spain | DMFS | n=536 All CT+ET RCT-R | LN1-3: 64% LN4+: 36% 100% ER+ 100% HER2- | All meno (46% post) | 18, 65 | 19 | 56 | 26 | - | - | - | 92 | 74 | 66 | 0-10yr: Low vs. int: HR 4.4 (NR)<br>0-10yr: Low vs. high: HR 5.8 (NR), *p*<0.0001<br>*Prosigna v.s EPclin + clinical factors (*p*=0.567) | -<br>Y<br>N* |
| **Prosigna: DFS** | | | | | | | | | | | | | | | | |
| Pu 2020[59] WHEL Study | DFS | n=344 Var ET/CT RCT-R | LN+ 100% ER+ 100% HER2- | All meno | NR | 26 | 53 | 21 | - | - | - | 81 | 64 | 56 | 0-10yr: *p*=0.02 | Y |

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno status | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| **Prosigna: OS and BCSS** | | | | | | | | | | | | | | | | |
| Lundgren 2022[60] SBII:pre trial | OS | n=123 ET/ none RCT-R | LN1-3 100% ER+ 100% HER2- | Pre-meno | 16, 40 | 2 | 42 | 57 | - | - | - | - | - | - | 0-10yr: Int vs. high: HR 1.84 (0.91–3.74); *p*=0.09<br>*0-10yr: Int vs. high: Adj HR 1.32 (0.61–2.88); *p*=0.48<br>>10yr: Int vs. high: HR 1.02 (0.54–1.93); *p*=0.96<br>*>10yr: Int vs. high: Adj HR 1.29 (0.66–2.53); *p*=0.46 | N<br>N*<br>N<br>N* |
| Lundgren 2022[60] SBII:pre trial | BCFI | n=123 ET/ none RCT-R | LN1-3 100% ER+ 100% HER2- | Pre-meno | 16, 40 | 2 | 42 | 57 | - | - | - | - | - | - | 0-10yr: Int vs. high: HR 1.99 (1.08–3.66); *p*=0.03<br>*0-10yr: Int vs. high: Adj HR 1.85 (0.95–3.58); *p*=0.07<br>>10yr: Int vs. high: HR 1.19 (0.50–2.80); *p*=0.70<br>*>10yr: Int vs. high: Adj HR 1.13 (0.43–2.95); *p*=0.81 | Y<br>N*<br>N<br>N* |
| Pu 2020[59] WHEL Study | BCSS | n=344 Var ET/CT RCT-R | LN+ 100% ER+ 100% HER2- | All meno | NR | 26 | 53 | 21 | - | - | - | 90 | 84 | 77 | 0-10yr: *p*=0.003 | Y |

[a]*The last column indicates whether each hazard ratio between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*
*Adj - adjusted; BCFI - breast cancer-free interval; BCSS - breast cancer-specific survival; CI - confidence interval; cohort-R - cohort reanalysis; CT - chemotherapy; CTS - Clinical Treatment Score (set of clinical factors); DFS - disease-free survival; DMFS - distant metastasis-free survival; DRFI - distant recurrence-free interval; DRFS - distant recurrence-free survival; ER - oestrogen receptor; ET - endocrine therapy; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor positive; int - intermediate; LN - lymph nodes (number positive); LR - likelihood ratio; meno - menopausal; NPI - Nottingham Prognostic Index; NR - not reported; OS - overall survival; RCT - randomised controlled trial; RCT-R - RCT reanalysis; SD - standard deviation; sig - significant; y/yr - year*

**Table 53: Prognostic data (EPclin)**

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno Clin risk | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| **EPclin: Distant recurrence, ET monotherapy** | | | | | | | | | | | | | | | | |
| Sestak 2018;[19] 2017[46] TransATAC | DRFI | n=183 ET mono RCT-R | LN1-3 100% HR+ 100% HER2- | Post-meno | 3.3 | 23 | - | 77 | 97.9 | - | 87.6 | 0-10y 94.4<br>5-10y 96.7 | 0-10y -<br>5-10y - | 0-10y 69.7<br>5-10y 76.4 | 0-5yr: High vs. low: HR 6.00 (0.80 to 44.93)<br>0-10yr: High vs. low: HR 6.77 (1.63 to 28.07)<br>0-10yr: Per 1SD change: 1.69 (1.29-2.22)<br>*LR vs CTS (*p*=0.20) or NPI (*p*=0.02) | N<br>Y<br>Y<br>N, Y* |

| Reference Study/cohort | Outcome | N, ET/CT Design | Nodal status HR, HER2 | Meno Clin risk | Test cut-offs | Distribution % | | | Risk 0-5yr % | | | Risk 0-10yr/other % | | | HR between test groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Low | Int | High | Low | Int | High | Low | Int | High | | |
| Filipits 2019[61] ABCSG-6/8 | DRFR | n=453 ET mono RCT-R | LN1-3 100% ER+ 100% HER2- | Post-meno | 3.3 | 35 | - | 65 | - | - | - | 0-10y 95.6 0-15y 84.7 5-10y 98.2 5-15y 87.0 | - | 0-10y 80.9 0-15y 75.1 5-10y 90.5 5-15y 84.0 | 0-10yr: HR 3.65 (1.73 to 7.68), *p*=0.0003<br>*0-10yr: Adj HR: 2.68 (1.77 to 4.08), *p*<0.0001<br><br>5-15yr: HR 3.00 (1.03 to 8.71), *p*=0.034<br>*5-15yr: Adj HR 3.43 (1.74 to 6.76), *p*=0.0005 | Y<br>Y*<br><br>Y<br>Y* |
| Sestak 2020[62] Lobular (from TransATAC + ABCSG-6/8) | DRFS | n=144 ET mono RCT-R | LN1-3: 80% LN4+: 20% 100% HR+ 100% HER2- | Post-meno Lobular | 3.3 | 26 | - | 74 | - | - | - | 93.6 | - | 68.8 | HR 3.70 (2.49 to 5.50), *p*<0.0001<br>*EPclin vs. clinical factors (*p*=0.0026) | Y<br>Y* |
| Constantinidou 2022[63] Cyprus + Notts | DRFS | n=62 ET mono Cohort-R | LN1-3 100% ER+ 100% HER2- | Pre-meno | 3.3 | 19 | - | 81 | - | - | - | 100 | - | 75 | High vs low: *p*=0.066<br>*Adj HR (cont score): 2.91 (1.70 to 4.97), *p*<0.001 | N<br>Y* |
| **EPclin: Distant recurrence, variable ET/CT** | | | | | | | | | | | | | | | | |
| Martin 2016,[57] Martin 2014[58] GEICAM 9906, Spain | DMFS | n=555 All CT+ET RCT-R | LN1-3: 64% LN4+: 36% 100% ER+ 100% HER2- | All meno (46% post) | 3.3 | 13 | - | 87 | - | - | - | 100 | - | 72 | Low vs. high: HR not estimable, *p*<0.0001<br>*EPclin vs. clinical factors (*p*=0.0018) | Y<br>Y* |
| | | | | Pre-meno | 3.3 | 12 | - | 88 | - | - | - | 100 | - | 70 | Low vs. high: HR NR, *p*=0.0006 | Y |
| | | | | Post-meno | 3.3 | 13 | - | 87 | - | - | - | 100 | - | 76 | Low vs. high: HR NR, *p*=0.0023 | Y |

*[a]The last column indicates whether each hazard ratio between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*

*Adj - adjusted; CI - confidence interval; cohort-R - cohort reanalysis; cont - continuous; CT - chemotherapy; CTS - Clinical Treatment Score (set of clinical factors); DMFS - distant metastasis-free survival; DRFI - distant recurrence-free interval; DRFR - distant recurrence-free rate; DRFS - distant recurrence-free survival; ER - oestrogen receptor; ET - endocrine therapy; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor; int - intermediate; LN - lymph nodes (number positive); LR - likelihood ratio; meno - menopausal; NPI - Nottingham Prognostic Index; NR - not reported; RCT - randomised controlled trial; RCT-R - RCT reanalysis; SD - standard deviation; sig - significant; y/yr - year*

## Appendix 5: Additional tables for observational data

### Table 54: Observational data for Oncotype DX (all outcomes and analyses)

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N ET/CT | Meno Age Clin | Test cut-offs | Distribution % | | | % risk of outcome | | | HR between test risk groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | Int | High | Low | Int | High | | |
| **Oncotype DX: Distant recurrence** | | | | | | | | | | | | | | |
| Clalit, Israel | Stemmer 2017[64] | LN1mic: 42%<br>LN1-3: 58%<br>100% ER+<br>100% HER2- | DRFI (0-5yr) | n=709 Var ET/CT | All meno | 18, 30 | 53 | 36 | 10 | 96.8 (7% CT) | 93.7 (40% CT) | 83.1 (86% CT) | 0-5yr: Low vs high: HR 0.19 (0.09 to 0.40)<br>0-5yr: Int vs. high: HR 0.39 (0.20 to 0.79), $p$<0.001<br>*0-5yr: Adj HR: Low vs high: HR 0.23 (0.11 to 0.50)<br>*0-5yr: Adj HR: Int vs. high: HR 0.42 (0.20 to 0.86), $p$=0.001 | Y<br>Y<br>Y*<br>Y* |
| | | | | | | 11, 25 | ≤25: 81 | | 19 | 95.7 (5% CT) | 96.0 (18% CT) | 86.9 (77% CT) | 0-5yr: $p$<0.001 | Y |
| | | | | | | ≤25, 26-30 | | | | 96.0 (15% CT) | | 91.5 (67% CT) | - | - |
| | | | | | | 18-25 | | | | | 94.4 (31% CT) | | - | - |
| | | | | n=109 Var ET/CT | Age <50 | 18, 30 | 48 | 37 | 16 | 96.2 (12% CT) | 100.0 (48% CT) | 64.2 (100% CT) | 0-5yr: $p$<0.001 | Y |
| | | | | n=464 Var ET/CT | Age 50-69 | 18, 30 | 54 | 37 | 9 | 97.6 (6% CT) | 93.5 (42% CT) | 87.8 (90% CT) | 0-5yr: $p$=0.017 | Y |
| | | | | n=136 Var ET/CT | Age ≥70 | 18, 30 | 57 | 33 | 10 | 94.7 (7% CT) | 88.7 (22% CT) | 92.9 (57% CT) | 0-5yr: $p$=0.458 | N |
| Young Women's Breast Cancer Study | Poorvu 2020[73] | LNmic, LN1-3<br>100% ER+<br>100% HER2- | DRFS (0-6yr) | n=163 Var ET/CT | Age ≤40 | 18, 30 | 33 | 42 | 25 | 0-6yr: 85.9 (83% CT) | 0-6yr: 87.3 (97% CT) | 0-6yr: 62.8 (98% CT) | 0-6yr: $p$=0.004 | Y |
| | | | | | | 11, 25 | 9 | 54 | 37 | 0-6yr: 92.3 (79% CT) | 0-6yr: 85.2 (92% CT) | 0-6yr: 71.3 (97% CT) | 0-6yr: $p$=0.10 | N |
| **Oncotype DX: DFS** | | | | | | | | | | | | | | |

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N ET/CT | Meno Age Clin | Test cut-offs | Distribution % | | | % risk of outcome | | | HR between test risk groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | Int | High | Low | Int | High | | |
| WSG PlanB | Nitz 2017[74] | LN1-3 100% HR+ 100% HER2- | DFS (0-5yr) | n=110 Var ET/CT | All meno | 0-10 | | | | 94.4 (No CT) | - | - | - | - |
| Red Cross Hospital, Munich, Germany | Braun 2022[75] | LNmic: 20% LN1-3: 80% 100% HR+ 100% HER2- | DFS (0-5yr) | n=217 Var ET/CT | All meno (63% post) | ≤25, 26+ | 86 | | 14 | RS 0-25: 90.3 (19% CT) | | 71.0 (93% CT) | - | - |
| **Oncotype DX: OS and BCSS** | | | | | | | | | | | | | | |
| Clalit, Israel | Stemmer 2017[64] | LN1mic: 42% LN1-3: 58% 100% ER+ 100% HER2- | BCSS (0-5yr) | n=709 Var ET/CT | All meno | 18, 30 | 53 | 36 | 10 | 99.5 (7% CT) | 96.6 (40% CT) | 94.3 (86% CT) | 0-5yr: $p<0.001$ | Y |
| | | | | | | 11, 25 | RS≤25: 81 | | 19 | 99.1 (5% CT) | 98.8 (18% CT) | 93.5 (77% CT) | 0-5yr: $p<0.001$ | Y |
| | | | | | | ≤25, 26-30 | | | | | 98.9 (15% CT) | 92.6 (67% CT) | - | - |
| | | | | | | 18-25 | | | | | 97.8 (31% CT) | | - | - |
| SEER registry | Petkov 2016[65] | LN1mic, LN1-3 100% HR+ 100% HER2- | BCSS (<5yr) | n=4,691 Var ET/CT | All | 18, 30 | 57 | 36 | 7 | 99.0 (23% CT) | 97.7 (47% CT) | 85.7 (75% CT) | <5yr: High vs. low: HR 11.0 (7.8 to 15.5)<br><5yr: Int vs. low: HR 3.1 (2.3 to 4.3), $p<0.001$<br>*<5yr: Adj HR: High vs. low: HR 7.8 (5.3 to 11.6)<br>*<5yr: Adj HR: Int vs. low: HR 3.0 (2.1 to 4.2), $p<0.001$ | Y<br>Y<br>Y*<br>Y* |
| | | | | n=328 Var ET/CT | Black ethnicity | 18, 30 | 54 | 36 | 9 | 99.4 (CT NR) | 98.9 | 91.3 | <5yr: $p=0.4117$ | N |
| | | | | n=4,021 Var ET/CT | White ethnicity | 18, 30 | 58 | 36 | 7 | 99 (CT NR) | 97.6 | 84.1 | <5yr: $p<0.0001$ | Y |
| | | | | n=320 Var ET/CT | Other ethnicity | 18, 30 | 57 | 34 | 8 | 98.5 (CT NR) | 99.1 | 100 | <5yr: $p=0.8427$ | N |

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N ET/CT | Meno Age Clin | Test cut-offs | Distribution % | | | % risk of outcome | | | HR between test risk groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | Int | High | Low | Int | High | | |
| SEER registry | Roberts 2017[66] | LN1mic, LN1-3 100% HR+ 100% HER2- | BCSS (0-5yr) | n=6,483 Var ET/CT | All | 18, 30 | 58 | 35 | 7 | 98.8 (CT NR) | 97.3 | 88.5 | 0-5yr: $p$<0.001<br>*0-5yr: Adj: $p$<0.001 | Y<br>Y* |
| | | | OS (0-5yr) | n=6,483 Var ET/CT | All | 18, 30 | 58 | 35 | 7 | 92.1 (CT NR) | 90.9 | 81.7 | 0-5yr: $p$<0.001<br>*0-5yr: Adj: $p$<0.001 | Y<br>Y* |
| SEER registry | Massarweh 2018[67] | LN1mic, LN1-3 100% HR+ 100% HER2- | BCSS (0-5yr) | n=6,437 Var ET/CT | Women | 18, 30 | 59 | 35 | 7 | 98.8 (23% CT) | 97.3 (48% CT) | 89.2 (77% CT) | 0-5yr: $p$<0.001 | Y |
| | | | | n=46 Var ET/CT | Men | 18, 30 | 52 | 26 | 22 | 100 (33% CT) | 100 (50% CT) | N/A (60% CT) | 0-5yr: $p$=0.02 | Y |
| | | | OS (0-5yr) | n=6,437 Var ET/CT | Women | 18, 30 | 59 | 35 | 7 | 92.2 (23% CT) | 90.8 (48% CT) | 83.2 (77% CT) | 0-5yr: $p$<0.001 | Y |
| | | | | n=46 Var ET/CT | Men | 18, 30 | 52 | 26 | 22 | 78.9 (33% CT) | 100 (50% CT) | N/A (60% CT) | 0-5yr: $p$=0.002 | Y |
| NCDB | Ibraheem 2020[68] | LN1-3 100% HR+ 100% HER2- | OS (0-5yr) | n=25,029 Var ET/CT | All meno | 11, 25 | 24 | 64 | 13 | - | - | - | 0-5yr:<br>Int vs low: HR 1.15 (0.97 to 1.36)<br>High vs low: HR 2.94 (2.43 to 3.56)<br>Per 10-RS: HR 1.38 (1.31 to 1.44) | <br>N<br>Y<br>Y |
| NCDB | Ibraheem 2019[69] | LN1-3: 97% LN4-9: 3% | OS (0-5yr) | n=13,163 Var ET/CT | All meno | 11, 25 | - | - | - | - | - | - | 0-5yr:<br>RS 18-25 vs 11-17: HR 1.20 (1.07-1.35), $p$<0.001<br>*RS 18-25 vs 11-17: Adj HR 1.15 (1.03-1.29), $p$<0.001<br>RS 26-30 vs 11-17: HR 1.91 (1.65-2.22), $p$<0.001<br>*RS 26-30 vs 11-17: Adj HR 1.62 (1.38-1.89), $p$<0.001 | <br>Y<br>Y*<br>Y<br>Y* |
| NCDB | Nash 2023[70] | LN1-3 100% HR+ 100% HER2- | OS (NR, med FU 5.5yr) | N=4,124 Var ET/CT | Age 40-50 | 11, 25 | - | - | - | - | - | - | *RS 26-30 vs 0-25: Adj HR 2.29 (1.49 to 4.86)<br>*RS 31-50 vs. 0-25: Adj HR 3.70 (2.03 to 6.75)<br>*RS 51-100 vs 0-25: Adj HR 2.31 (0.78–6.86)<br>$p$<0.001 | Y*<br>Y*<br>N*<br> |
| NCDB | Weiser 2022[71] | LN1-3 100% HR+ 100% HER2- | OS (0-5yr) | n=2,691 Var ET/CT | Lobular | 11, 25 | - | - | - | 95.5 | 95.5 | 83.8 | 0-5yr: $p$=0.0004<br>*Adj: sig | Y<br>Y* |

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N ET/CT | Meno Age Clin | Test cut-offs | Distribution % | | | % risk of outcome | | | HR between test risk groups (95% CI) | [a]Sig? *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | Int | High | Low | Int | High | | |
| NCDB | Weiser 2021[72] | LN1-3 100% HR+ 100% HER2- | OS (0-5yr) | n=28,591 Var ET/CT | All | ≤25 | - | - | - | - | - | - | *0-5yr: RS 18-25 vs. RS 12-17: Adj HR 1.30 (1.00 to 1.68) | Y* |

[a]*The last column indicates whether each hazard ratio between test risk groups is statistically significant at the 5% level. Asterisk (*) denotes analyses adjusted for clinical factors.*

*Adj - adjusted; BCSS - breast cancer-specific survival; CI - confidence interval; CT - chemotherapy; DFS - disease-free survival; DRFI - distant recurrence-free interval; DRFS - distant recurrence-free survival; ER - oestrogen receptor; ET - endocrine therapy; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor; int - intermediate; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; OS - overall survival; sig - significant; var - variable; y/yr - year*

**Appendix 6: Additional tables for chemotherapy effect within risk groups**

**Table 55: Chemotherapy effect within risk groups: Registry data for Oncotype DX (all outcomes)**

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno Age Clin | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| **Oncotype DX: Observational: Distant recurrence** | | | | | | | | | | | | | | | | | | | | |
| Clalit, Israel | Stemmer 2017[64] | LN1mic: 42%<br>LN1-3: 58%<br>100% ER+<br>100% HER2- | DRFI 0-5yr | n=709 | All meno | 18, 30 | 92.3 | 97.1 | 99 | 90.3 | 82 | 90 | -4.8 | 8.7 | -8.0 | *p*=0.245 | *p*=0.019 | - | - | - |
| | | | | | | 11, 25 | 83.3 | 96.3 | 98.8 | 95.4 | 97.5 | 79.7 | -13.0 | 3.4 | 17.8 | - | - | *p*=0.017 | - | - |
| | | | | | | All ≤25 | - | - | 97.7 | 95.6 | - | - | 2.1 | | - | *p*=0.521 | | - | - | - |
| | | | | | | 18-25 | - | - | 100 | 91.8 | - | - | - | 8.2 | - | - | *p*=0.058 | - | - | - |
| | Rotem 2022 (abst)[77] | LN+<br>100% ER+<br>100% HER2- | DRFS 0-7yr | n=140 | All meno | All 26-30 | - | - | - | - | 89.4 | 78.0 | - | - | 11.4 | - | - | Not sig | - | - |
| **Oncotype DX: Observational: BCSS and OS** | | | | | | | | | | | | | | | | | | | | |
| Clalit, Israel | Stemmer 2017[64] | LN1mic: 42%<br>LN1-3: 58%<br>100% ER+<br>100% HER2- | BCSS 0-5yr | n=709 | All meno | 18, 30 | 100.0 | 99.4 | 98.9 | 95.1 | 93.4 | 100 | 0.6 | 3.8 | -6.6 | - | - | - | - | - |
| | | | | | | 11, 25 | 100.0 | 99.1 | 100.0 | 98.6 | 97.1 | 84.0 | 0.9 | 1.4 | 13.1 | - | - | - | - | - |
| | | | | | | All ≤25 | - | - | 100.0 | 98.7 | - | - | 1.3 | | - | - | - | - | - | - |
| | | | | | | 18-25 | - | - | 100.0 | 96.8 | - | - | - | 3.2 | - | - | - | - | - | - |
| | Rotem 2022 (abst)[77] | LN+<br>100% ER+<br>100% HER2- | BCSS 0-7yr | n=140 | All meno | 26-30 | - | - | - | - | 98.7 | 93.8 | - | - | 4.9 | - | - | *p*=0.024 | - | - |
| SEER | Petkov 2020 (abst)[78] | LN1mic-LN3<br>100% HR+<br>100% HER2- | BCSS 0-5yr | n=2,588 | Age≤50 | 0-10 | 100 | 100 | - | - | - | - | 0 | - | - | - | - | - | - | - |
| | | | | | | 11-15 | - | - | 97.7 | 99.5 | - | - | - | -1.8 | - | - | - | - | - | - |
| | | | | | | 16-20 | - | - | 98.4 | 98.7 | - | - | - | -0.3 | - | - | - | - | - | - |
| | | | | | | 21-25 | - | - | 98.8 | 98.4 | - | - | - | 0.4 | - | - | - | - | - | - |
| | | | | | | 26-100 | - | - | - | - | 93.9 | 95.6 | - | - | -1.7 | - | - | - | | |

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno Age Clin | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| Clalit, Israel | Rotem 2022 (abst)[77] | LN+ 100% ER+ 100% HER2- | OS 0-7yr | n=140 | All meno | 26-30 | - | - | - | - | 96.3 | 93.8 | - | - | 2.5 | - | - | Not sig | - | - |
| NCDB | Abel 2022[79] | LN1-3 100% HR+ 100% HER2- | OS 0-5yr | n=21,370 | Ductal | All ≤25 | - | - | - | - | - | - | - | - | - | *p*=0.278 | | - | - | - |
| | | | | n=6,356 | Lobular | All ≤25 | - | - | - | - | - | - | - | - | - | *p*=0.532 | | - | - | - |
| | | | | n=4,251 | Age<50 Ductal | All ≤25 | - | - | - | - | - | - | - | - | - | Unadj: 0.44 (0.22 to 0.86), *p*=0.016 | | - | - | - |
| | | | | n=1,062 | Age<50 Lobular | All ≤25 | - | - | - | - | - | - | - | - | - | Unadj: 0.54 (0.14 to 2.18), *p*=0.39 | | - | - | - |
| NCDB (cont) | Cao 2022 (abst)[80] | LN1-3 100% ER+ 100% HER2- | OS NR | n=28,427 | Age≤50 | All 20-25 | - | - | - | - | - | - | - | - | - | - | Unadj: 0.334 (NR), *p*=0.002 | - | - | - |
| | | | | | Age>50 | All 20-25 | - | - | - | - | - | - | - | - | - | - | Unadj: 0.521 (NR), *p*=0.019 | - | - | - |
| NCDB (cont) | Ibraheem 2019[69] | LN1-3: 97% LN4-9: 3% 100% HR+ 100% HER2- | OS 0-5yr | n=13,163 | All meno | 11-17 | - | - | 97.7 | 96.5 | - | - | - | 1.2 | - | - | Adj: 0.63 (0.40 to 0.99), *p*=0.044<br><br>Threshold: RS >13 sig CT benefit | - | - | - |
| | | | | | | 18-25 | - | - | 96.0 | 92.7 | - | - | - | 3.3 | - | - | Adj: 0.53 (0.37 to 0.76), *p*=0.001 | - | - | - |
| | | | | | | 26-30 | | | | | 92.2 | 85.5 | | | 6.7 | | | Adj: 0.50 (0.28 to 0.89), *p*=0.018 | | |

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno Age Clin | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| | | | | n=3,101 | Age≤50 | All 11-25 | - | - | - | - | - | - | - | - | - | - | Adj: 0.68 (0.35 to 1.32), *p*=0.25 | - | - | - |
| | | | | n=8,886 | Age>50 | All 11-25 | - | - | - | - | - | - | - | - | - | - | Adj: 0.64 (0.47 to 0.86), *p*=0.004 | - | - | - |
| NCDB (cont) | Iorgulescu 2019[81] | LN1-3 100% ER+ 100% HER2- | OS 0-5yr | n=2,735 | All meno | 18, 30 | 93 | 92 | 93.2 | 85.7 | 92.4 | 66.9 | 1.0 | 7.5 | 25.5 | Unadj: p=0.27 Adj: 0.81 (0.33 to 1.98, *p*=0.64 | Unadj: p=0.02 Adj: 0.67 (0.35 to 1.27), *p*=0.22 | Unadj: p<0.001 Adj: 0.24 (0.13 to 0.47), *p*<0.001 | - | - |
| NCDB (cont) | Kumar 2023 (abst)[82] | LN1-3: >90% LN4+: <10% 100% HR+ 100% HER2- | OS 0-10yr | n=8,628 | Age≤50 | 0-11 | - | - | - | - | - | - | - | - | - | Adj: 0.56 (0.22 to 1.42) | - | - | - | - |
| | | | | | | 12-25 | - | - | - | - | - | - | - | - | - | - | Adj: 0.55 (0.38 to 0.80) | - | | |
| | | | | | | All ≤25 | - | - | 93.0 | 91.0 | - | - | 2.0 | | - | Unadj: 0.60 (0.48 to 0.75), *p*<0.0001 Adj: 0.54 (0.39 to 0.76), *p*=0.0004 | | - | - | - |
| | | | | n=8,628 | Age 18-40 | All ≤25 | - | - | 86.0 | 82.8 | - | - | 3.2 | | - | Adj: 0.43 (0.22 to 0.85) | | - | - | - |
| | | | | n=8,628 | Age 40-50 | All ≤25 | - | - | 94.7 | 92.2 | - | - | 2.5 | | - | Adj: 0.59 (0.39 to 0.87) | | - | - | - |
| NCDB (cont) | Nash 2023[70] | LN1-3 100% HR+ 100% HER2- | OS NR, med FU 5.5yr | N=4,124 | Age 40-50 | All ≤25 | - | - | - | - | - | - | - | - | - | Unadj: *p*=0.41 Adj: 0.72 (0.47 to 1.12), *p*=0.15 | | - | - | - |
| | | | | | | 25-30 | - | - | - | - | - | - | - | - | - | - | | Unadj: *p*=0.28 | | |

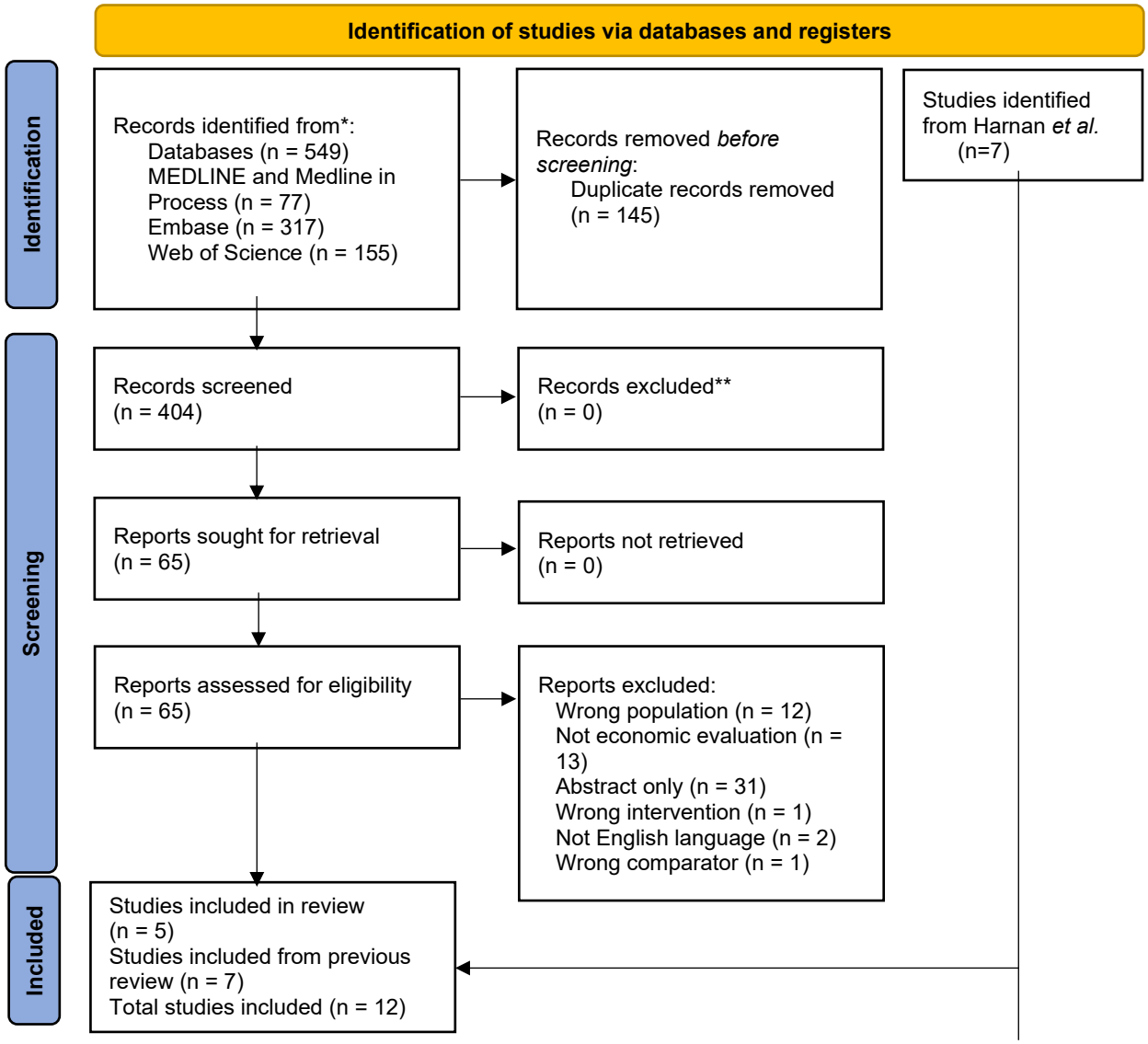
| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno Age Clin | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| | | | | | | 31-50 | - | - | - | - | - | - | - | - | - | - | | Unadj: p=0.002 Adj: 0.29 (0.10 to 0.85), *p*=0.02 | | |
| | | | | | | >50 | - | - | - | - | - | - | - | - | - | - | | Not sig (few events) | | |
| NCDB (cont) | Weiser 2022[71] | LN1-3 100% HR+ 100% HER2- | OS 0-5yr | n=16,646 | All meno Ductal | 11-25 | - | - | 96.7 | 95.1 | - | - | - | 1.6 | - | - | Unadj: *p*=0.004 Adj: non-sig | - | - | - |
| | | | | NR | Age<50 Ductal | All ≤25 | - | - | - | - | - | - | - | - | - | Adj: 2.32 (1.19 to 4.49) | | - | - | - |
| | | | | NR | Age 50-75 Ductal | All ≤25 | - | - | - | - | - | - | - | - | - | Adj: 1.12 (0.86 to 1.46) | | - | - | - |
| | | | | n=2,691 | All meno Lobular | 0-10 | 94.7 | 95.7 | - | - | - | - | -1.0 | - | - | Unadj: *p*=0.888 Adj: non-sig | - | - | - | - |
| | | | | | | 11-25 | - | - | 96.6 | 94.9 | - | - | - | 1.7 | - | - | Unadj: *p*=0.381 Adj: non-sig | - | | |
| NCDB (cont) | Weiser 2021[72] | LN1-3 100% HR+ 100% HER2- | OS 0-5yr | n=28,591 | All meno | All ≤25 | - | - | 96.6 | 93.2 | - | - | 3.4 | | - | Unadj: p<0.001 Adj: 1.63 (1.28 to 2.07) | | - | - | - |
| | | | | NR | Age≤50 | All ≤25 | - | - | - | - | - | - | 1.4 | | - | Adj: 1.88 (1.05 to 3.37), *p*=0.032 | | - | - | - |
| | | | | | | 12-17 | - | - | - | - | - | - | - | 1.3 | - | - | Adj: 2.49 (0.80 to 7.76) | - | - | - |

| Cohort | Ref | Nodal status HR, HER2 | Out-come | N | Meno Age Clin | Test cut-offs | % risk of outcome | | | | | | Abs diff CT v no CT | | | HR for CT vs. no CT (95% CI) | | | Inter-action | [a]Pred *Adj |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Low | | Int | | High | | Low | Int | High | Low | Int | High | | |
| | | | | | | | CT | No | CT | No | CT | No | | | | | | | | |
| | | | | | | 18-25 | - | - | - | - | - | - | - | 4.4 | - | - | Adj: 3.30 (1.38 to 7.84) | - | - | - |
| | | | | NR | Age 51-70 | All ≤25 | - | - | - | - | - | - | 1.6 | | - | Adj: 1.49 (1.12 to 1.97), *p*=0.006 | | - | - | - |
| | | | | | | 12-17 | - | - | - | - | - | - | - | 3.6 | - | - | Adj: 2.80 (1.45 to 5.24) | - | - | - |
| | | | | | | 18-25 | - | - | - | - | - | - | - | 3.2 | - | - | Adj: 1.37 (0.92–2.05) | - | - | - |
| | | | | NR | Age>70 | All ≤25 | - | - | - | - | - | - | - | - | - | Adj: 1.1 (0.68 to 1.78), *p*=0.69 | | - | - | - |
| | | | | NR | Age≤70 | 0-10 | - | - | - | - | - | - | - | - | - | *p*=0.44 | - | - | | |
| | | | | | | 12-25 | - | - | - | - | - | - | - | 3.0 | - | - | Adj: 1.91 (1.42 to 2.57) | - | - | - |
| | | | | | | 12-17 | - | - | - | - | - | - | - | 3.4 | - | - | Adj: 3.04 (1.78 to 5.21), *p*<0.001 | - | - | - |
| | | | | | | 18-25 | - | - | - | - | - | - | - | 3.8 | - | - | Adj: 2.02 (1.42 to 2.87), *p*<0.001 | - | - | - |

*[a]The Last column indicates whether interaction test (between risk group and CT use) indicates a significant predictive effect for CT benefit at the 5% level. Asterisk (*) denotes interaction adjusted for clinical factors. Abs diff - absolute difference; adj - adjusted; BCSS - breast cancer-specific survival; CI - confidence interval; CT - chemotherapy; DRFI - distant recurrence-free interval; DRFS - distant recurrence-free survival; ER - oestrogen receptor; HER2 - human epidermal growth factor receptor 2; HR - hazard ratio; HR - hormone receptor; int - intermediate; LN - lymph nodes (number positive); meno - menopausal; NR - not reported; OS - overall survival; prosp - prospective; pred - predictive of CT benefit; RCT - randomised controlled trial; RS – Recurrence Score (Oncotype DX); sig - significant; unadj - unadjusted; yr - year*

## Appendix 7: PRISMA flow diagrams for published economic evaluations and HRQoL studies

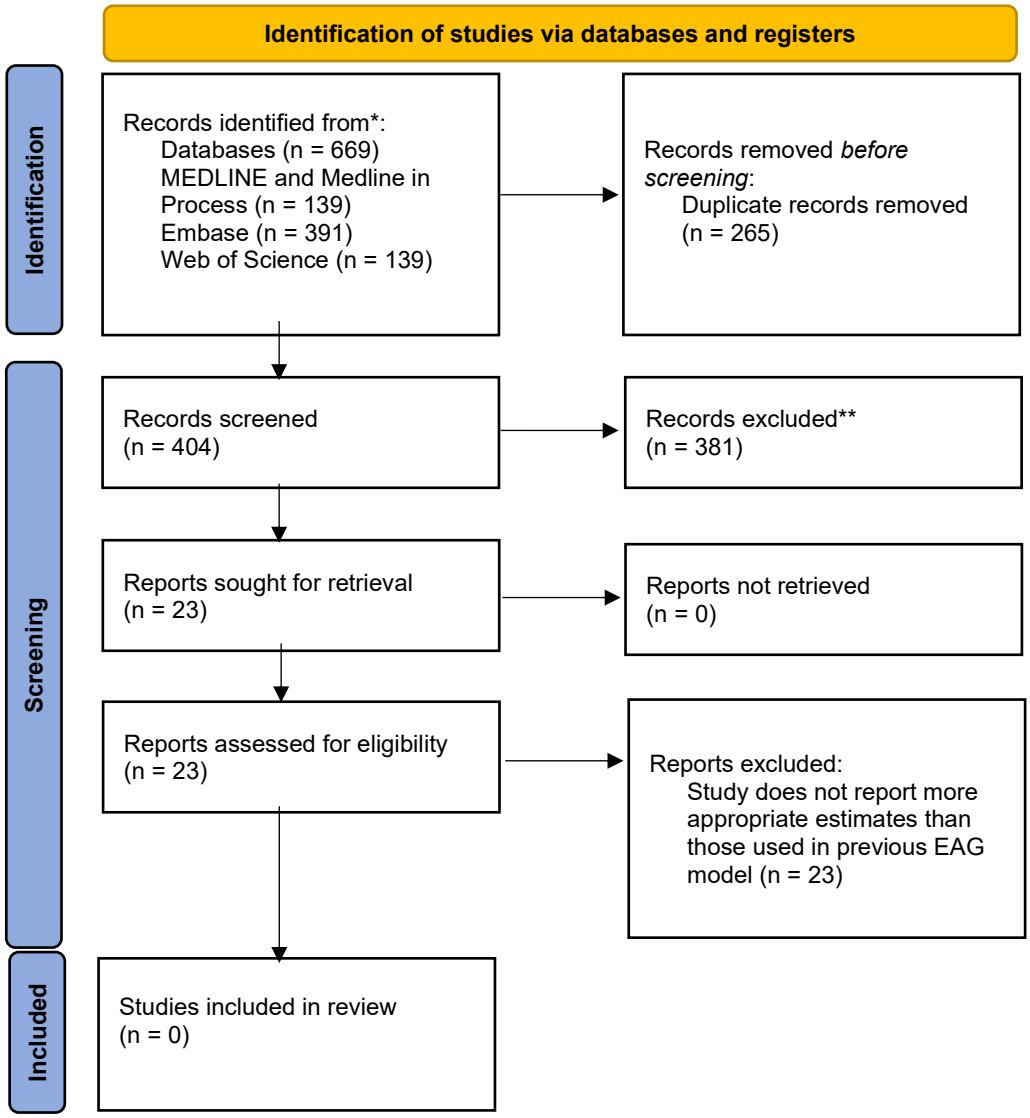
### PRISMA 2020 flow diagram for economic evaluations of tumour profiling tests




*From:* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

**PRISMA 2020 flow diagram for HRQoL associated with different health states for women with breast cancer**




*From:* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

## Appendix 8: Adjuvant chemotherapy infusion time by regimen

**Table 56: Infusion time for each chemotherapy regimen included in the EAG model**

| Regimen | Doses per course | Infusion time per dose (hours) | Infusion time per course |
|---|---|---|---|
| **FEC75 (6 cycles)** | | | |
| Fluorouracil 600 mg/m² | 6 | 0.08 | 0.50 |
| Epirubicin 75 mg/m² | 6 | 0.08 | 0.50 |
| Cyclophosphamide 600mg/m² | 6 | 0.50 | 3.00 |
| **FEC100+T (3+3 cycles)** | | | |
| Fluorouracil 500 mg/m² | 3 | 0.08 | 0.25 |
| Epirubicin 100 mg/m² | 3 | 0.08 | 0.25 |
| Cyclophosphamide 500mg/m² | 3 | 0.50 | 1.50 |
| Docetaxel 100mg/m² | 3 | 1.00 | 3.00 |
| **TC (4 cycles)** | | | |
| Cyclophosphamide 600mg/m² | 4 | 0.08 | 0.33 |
| Docetaxel 75mg/m² | 4 | 1.00 | 4.00 |
| **EC90/T75 (4+4 cycles)** | | | |
| Epirubicin 90 mg/m² | 4 | 0.08 | 0.33 |
| Cyclophosphamide 600mg/m² | 4 | 0.08 | 0.33 |
| Docetaxel 75mg/m² | 4 | 0.50 | 2.00 |
| **EC90 (4 cycles)** | | | |
| Epirubicin 90 mg/m² | 4 | 0.08 | 0.33 |
| Cyclophosphamide 600mg/m² | 4 | 0.08 | 0.33 |
| **Accelerated EC90/P (4+4 cycles)** | | | |
| Epirubicin 90 mg/m² | 4 | 0.08 | 0.33 |
| Cyclophosphamide 600mg/m² | 4 | 0.08 | 0.33 |
| Paclitaxel 175mg/m² | 4 | 1.00 | 4.00 |
| **C-D (6 cycles)** | | | |
| Carboplatin AUC 6 (assumed 600mg) | 6 | 1.00 | 6.00 |
| Docetaxel 75mg/m² | 6 | 1.00 | 6.00 |
| **TAC (6 cycles)** | | | |
| Docetaxel 75mg/m² | 6 | 1.00 | 6.00 |
| Doxorubicin 50mg/m² | 6 | 0.08 | 0.50 |
| Cyclophosphamide 500mg/m² | 6 | 0.08 | 0.50 |
| **Weekly P (12 weeks)** | | | |
| Paclitaxel 80mg/m² | 12 | 1.00 | 12.00 |
| **EC90 / weekly P (4 cycles, 12 weeks)** | | | |
| Epirubicin 90 mg/m² | 4 | 0.08 | 0.33 |
| Cyclophosphamide 600mg/m² | 4 | 0.08 | 0.33 |
| Paclitaxel 80mg/m² | 12 | 1.00 | 12.00 |

## Appendix 9: Cost-effectiveness acceptability curves for EAG base case scenarios

**Figure 7: CEACs, BC1 - Oncotype DX, RxPONDER pre-menopausal (predictive benefit)**

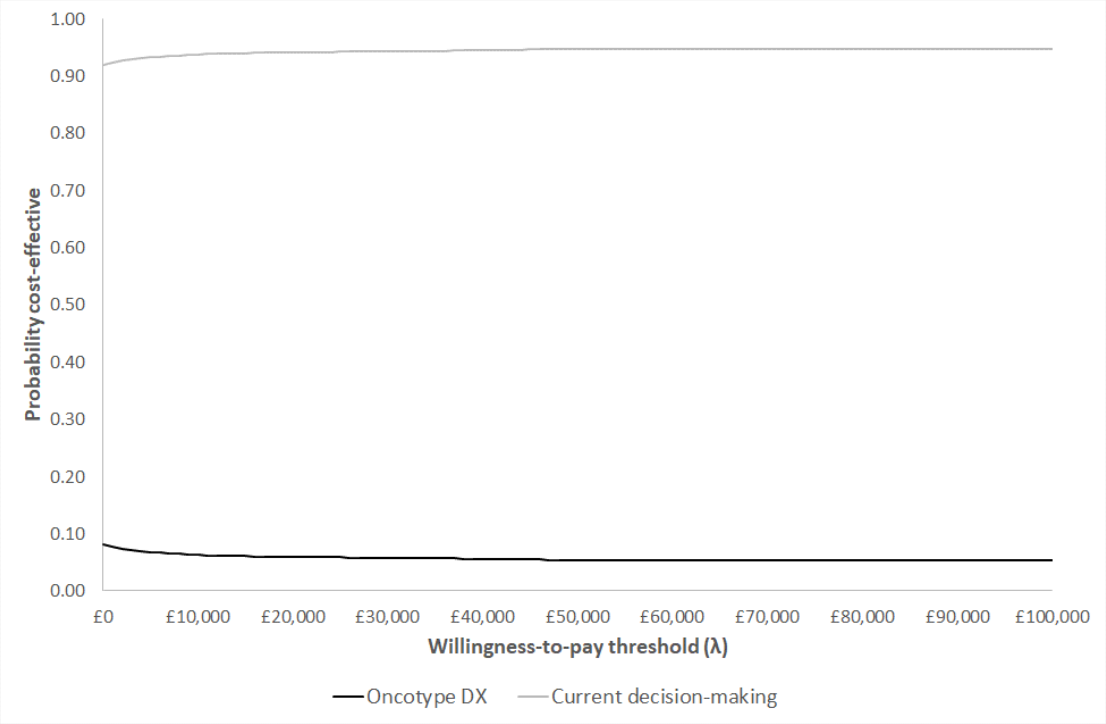



**Figure 8: CEACs, BC2 - Oncotype DX, RxPONDER post-menopausal (predictive benefit)**

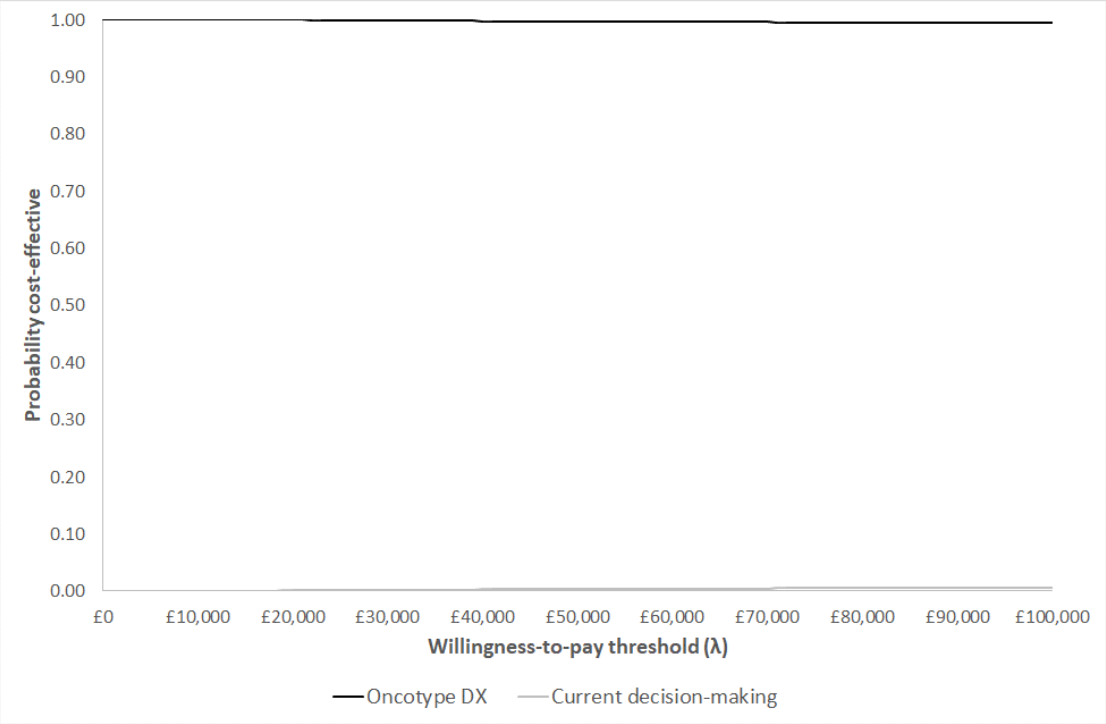


**Figure 9: CEACs, BC3 - Oncotype DX, TransATAC, post-menopausal (predictive benefit)**

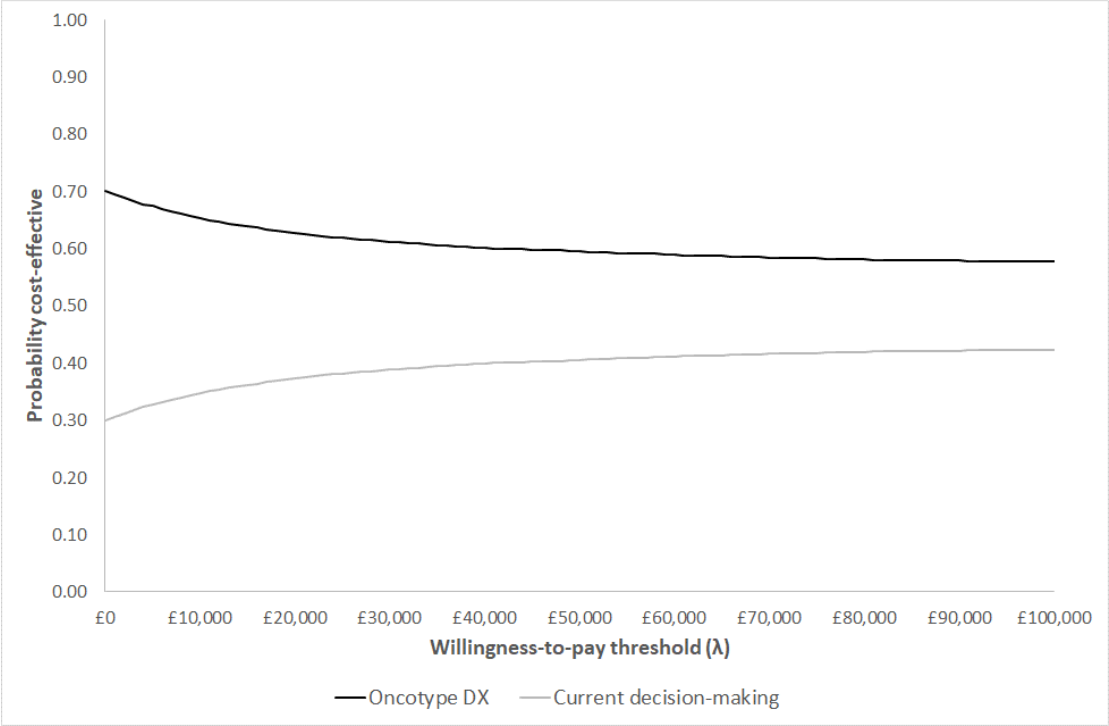



**Figure 10: CEACs, BC4 - Oncotype DX, TransATAC, post-menopausal (non-predictive benefit)**

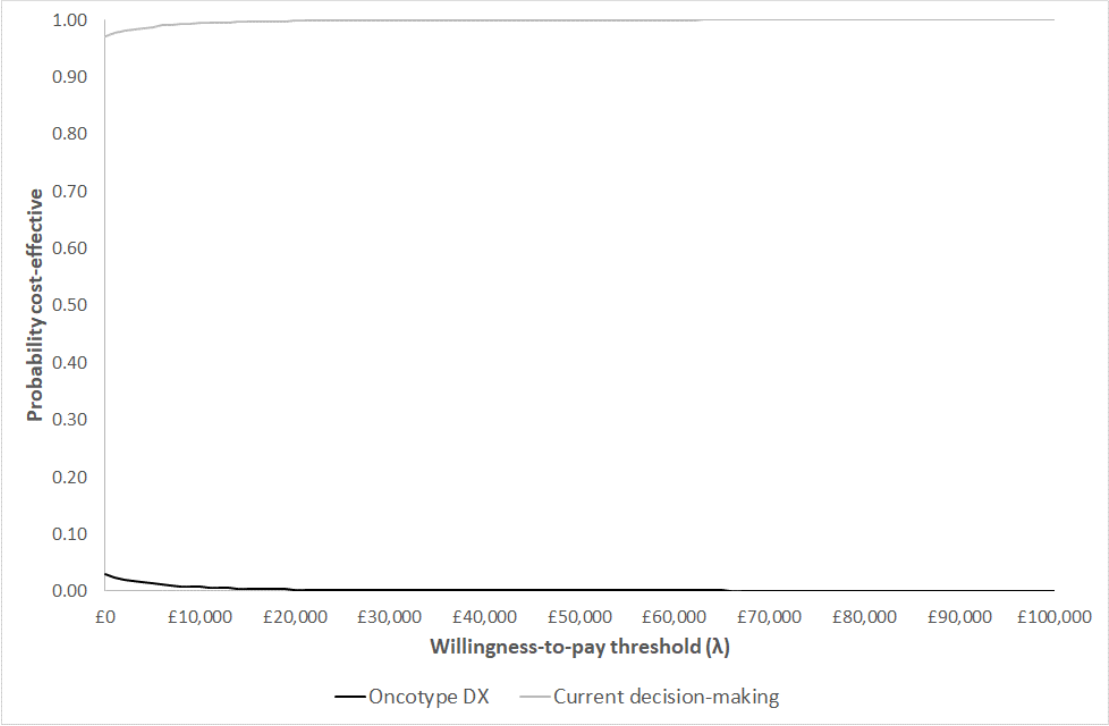


**Figure 11: CEACs, BC5 – Prosigna, TransATAC, post-menopausal (non-predictive benefit)**

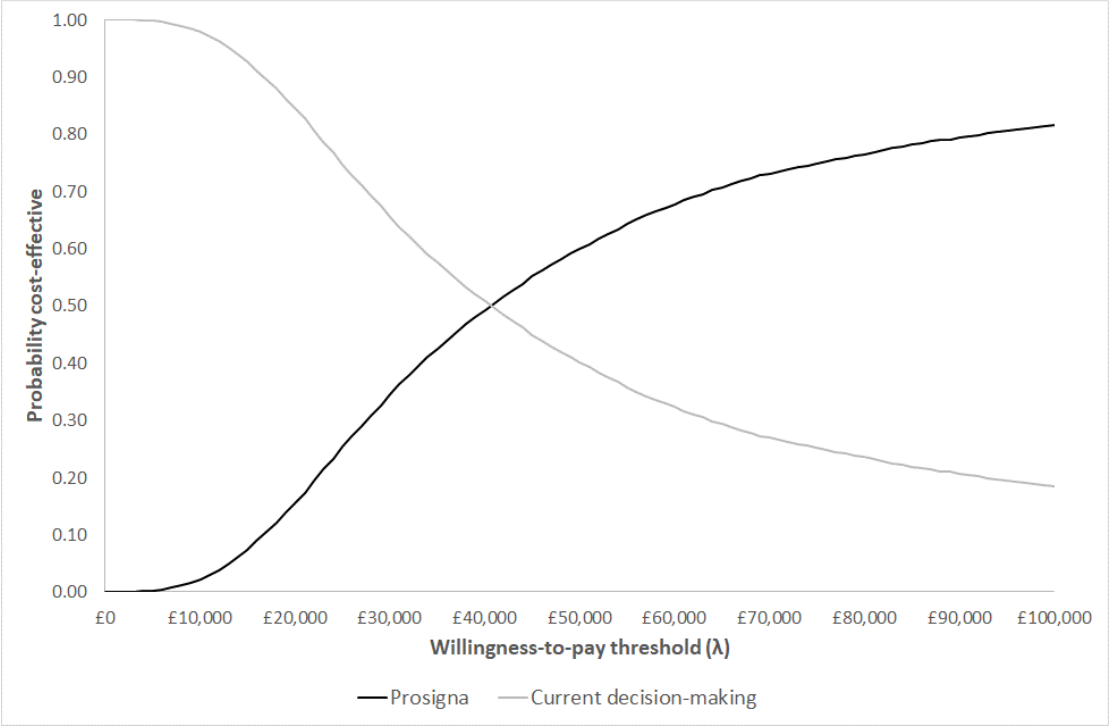



**Figure 12: CEACs, BC6 – EPclin, TransATAC, post-menopausal (non-predictive benefit)**

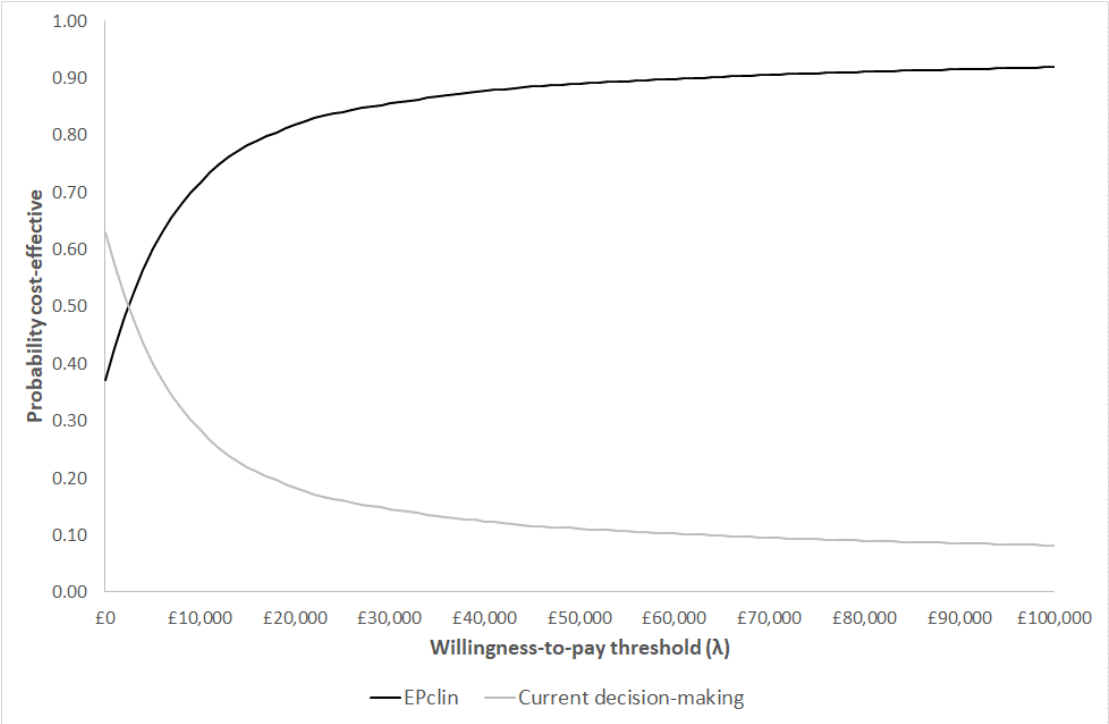
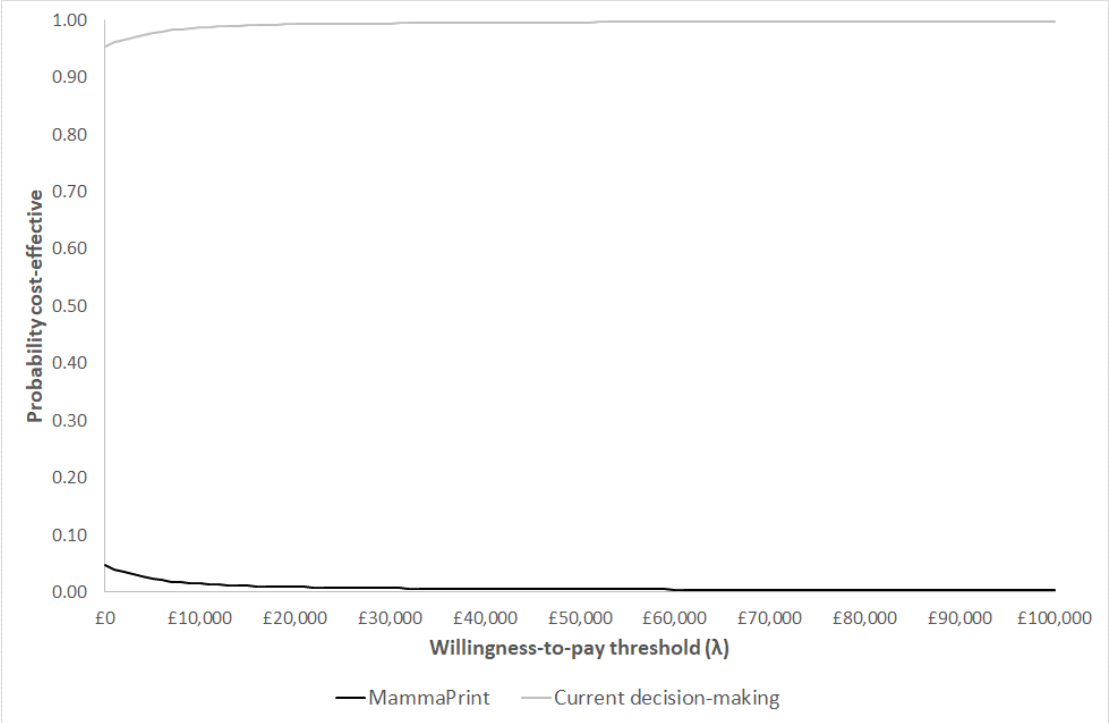


**Figure 13: CEACs, BC7 - MammaPrint, MINDACT, LN+ subgroup (non-predictive benefit)**



# Evidence overview: Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

This overview summarises the main issues the diagnostics advisory committee needs to consider. It should be read together with the final scope and the external assessment report.

## 1 Aims and scope

Tumour profiling tests are designed to provide information on the activity of genes within tumour samples from people with early breast cancer. The results of the tests provide a risk profile of an individual's breast cancer which can be combined with other clinical risk factors that are routinely assessed, such as nodal status and tumour size, to better predict the risk of disease recurrence in the future. Some tests may also predict the benefit a person may receive from chemotherapy. This information is intended to help treatment decision-making with regard to adjuvant chemotherapy use.

NICE diagnostics guidance 34 makes recommendations on the use of tests for people with oestrogen receptor (ER) positive, human epidermal growth factor receptor 2 (HER2) negative and lymph node (LN) negative early breast cancer (including micrometastatic disease). Use of the tests for LN positive disease was examined in this assessment, but there was not sufficient evidence to make a positive recommendation. Because there are no NICE recommendations on the use of tumour profiling tests for LN positive disease, current practice is variable across the country. New evidence is now available, and in 2023 NHS England submitted a specific request to re-examine the use of tumour profiling tests in this population.

The use of tumour profiling tests may improve the identification of people with LN positive early breast cancer who may not benefit from having adjuvant chemotherapy because they have a genomic low risk of disease recurrence.

These people could potentially avoid unnecessary treatment, and therefore they would not be exposed to the co-morbidities and negative impacts on quality of life that are associated with chemotherapy. The tests may also identify people with LN positive early breast cancer who have been identified as low risk of disease recurrence based on current clinical practice, but would actually benefit from chemotherapy. People with breast cancer and clinicians may also benefit from improved confidence in the appropriateness of the treatment they are having or recommending.

## Decision question

Do tumour profiling tests used for guiding adjuvant chemotherapy decisions for people with lymph node positive early breast cancer represent a clinically- and cost-effective use of NHS resources?

## Populations

People with ER positive and/or progesterone receptor (PR) positive, HER2 negative, early breast cancer with 1 to 3 positive lymph nodes, who are deciding whether to have adjuvant chemotherapy.

Where data permits, the following subgroups may be considered:

- Pre-menopausal women and post-menopausal women
- People predicted to be in low, intermediate or high risk groups using a risk assessment tool (such as PREDICT or the Nottingham Prognostic Index [NPI]), or using clinical and pathological features
- Sex
- People of different ethnicities
- People with comorbidities which mean that they could be particularly affected by the side effects of chemotherapy, such as cardiac, pulmonary or neurological conditions.

### Interventions

- EndoPredict EPclin score
- MammaPrint
- Oncotype DX Breast Recurrence Score
- Prosigna

in combination with current decision making.

### Comparator

Current decision making, which may include any tool, or clinical and pathological features, used to assess risk

### Healthcare setting

Secondary and tertiary care

Further details, including descriptions of the interventions, comparator, care pathway and outcomes, are in the final scope.

## 2 Clinical effectiveness evidence

The external assessment group (EAG) did a systematic review to identify evidence on the clinical effectiveness and diagnostic accuracy of tumour profiling tests in lymph node positive early breast cancer. Find the full systematic review results in section 3 (page 41) of the external assessment report.

### Overview of included studies

The EAG updated the systematic review carried out for the LN positive subgroup in NICE diagnostics guidance 34. Most included studies had a population that was at least 80% people with ER or PR positive, HER2 negative early breast cancer with 1 to 3 positive lymph nodes. Only studies using the commercial versions of the tumour profiling tests (see interventions) were included.

The clinical review aimed to identify the following types of data:

- End-to-end studies comparing the tests versus current decision-making
- Prognostic ability
- Ability to predict benefit from chemotherapy
- Impact of test results on chemotherapy decisions
- Health-related quality of life (HRQoL) and anxiety associated with use of the tests.

The data on prognostic and predictive ability included the following clinical outcomes:

- Distant recurrence-free survival (DRFS), distant recurrence-free interval (DRFI), distant metastasis-free survival (DMFS) and distant metastasis-free interval (DMFI)
- Disease-free survival (DFS)
- Overall survival (OS) and breast cancer-specific survival (BCSS).

More detail can be found in section 3.1 (page 41) of the external assessment report.

In total, 54 publications were included in the clinical review, of which 41 were newly identified. Forty-two publications related to clinical outcomes and 12 to decision impact studies. No studies were identified that assessed HRQoL or anxiety in the LN positive population, so a short summary of studies in a broader early breast cancer population was produced.

Two prospective randomised controlled trials were identified: MINDACT for MammaPrint and RxPONDER for Oncotype DX. Other key studies were reanalyses of the SWOG-8814 trial for Oncotype DX, and of the TransATAC trial (this data was also used in the evaluation for NICE diagnostics guidance 34). For full details see section 3.2 (page 44) in the external assessment report.

## MINDACT

MINDACT (Piccart et al. 2021) assessed the genomic risk (using MammaPrint) and clinical risk (using modified Adjuvant! Online, mAOL) of people with early breast cancer. People who were low-risk on both MammaPrint and mAOL were allocated to no chemotherapy, those who were high-risk on both were allocated to chemotherapy, and people with discordant risk were randomised to chemotherapy or no chemotherapy. Outcomes were reported for people according to high or low clinical risk and high or low MammaPrint risk. There are limitations in using MINDACT to assess prognostic ability because MammaPrint results influenced chemotherapy use (more people in the MammaPrint high-risk group received chemotherapy than in the low-risk group), and no hazard ratios or significance tests were reported for the difference in outcomes between test risk groups. The study also provided data on the effect of chemotherapy versus no chemotherapy on clinical outcomes. Results were presented for the clinical high, MammaPrint low group; however, data were not analysed for the clinical low, MammaPrint high group due to small numbers of people with LN positive disease. The study therefore provided data on chemotherapy benefit only for people with clinical high, MammaPrint low risk.

## RxPONDER

RxPONDER (Kalinsky et al. 2021) randomised people with an Oncotype DX recurrence score (RS) of up to 25 to chemo-endocrine therapy or endocrine monotherapy. The study assessed the prognostic ability of RS (between 0 and 25) for invasive disease-free survival. It also reported the effect of chemotherapy versus no chemotherapy, and whether RS was predictive of chemotherapy benefit. No prognostic or predictive data was available for people with RS above 25 due to the study design.

## TransATAC

The ATAC study evaluated the efficacy and safety of 2 different endocrine therapies in post-menopausal women with localised breast cancer.

TransATAC was a retrospective reanalysis of tumour samples collected from HR positive participants to evaluate the prognostic value of various tumour profiling tests for distant recurrence. Data from TransATAC was used to inform the economic model in NICE diagnostics guidance 34.

### SWOG-8814

The SWOG-8814 randomised controlled trial compared the efficacy of chemo-endocrine therapy with endocrine monotherapy in post-menopausal women with ER positive, LN positive breast cancer. More than 20% of the population had 4 or more positive nodes. Albain et al. 2010 reported a retrospective reanalysis in which Oncotype DX was done on tumour samples from the study to determine the prognostic and predictive ability of the test. Oncotype DX was used as a 3-level test in this study.

## Study quality

The 2 prospective randomised clinical trials MINDACT and RxPONDER were both scored as being at low risk of bias in all domains of the Cochrane RoB2 tool.

Risk of bias in prognostic and predictive studies was assessed using the PROBAST tool. The following factors may have affected results:

- Prognostic studies varied in terms of whether people received chemotherapy or not
- Only 1 study of predictive benefit (Albain et al. 2010) was a reanalysis of a study where chemotherapy use was randomised
- In some studies, some participants did not match the scope population (either not ER positive, not HER2 negative, or not LN 1 to 3)
- Most studies excluded a proportion of patients for various reasons including insufficient tissue, missing data, failed tests and others, although the impact of this was unclear

- Chemotherapy decisions were not influenced by the test result in studies of retrospective use of the test (reanalyses of randomised trials and cohorts). In observational studies in which the test was used prospectively, chemotherapy decisions may have been influenced by the test result.

For full details on the risk of bias assessment, please see section 3.3 (page 47) and appendix 3 (page 174) in the external assessment report.

## Intermediate outcomes

### Prognostic ability

The prognostic ability of a genomic test describes its ability to differentiate between people with good versus poor outcomes. Studies of prognostic ability also provide data on the proportion of people allocated to different risk groups by the test. In total, 23 reanalyses of randomised trials or cohorts with long-term follow up (18 studies) provided data on prognostic ability: 5 studies of EPclin, 5 studies of MammaPrint, 5 studies of Oncotype DX, and 6 studies of Prosigna. The randomised controlled trials MINDACT and RxPONDER also provided prognostic data, and 12 further publications reported on 6 prospective observational studies of Oncotype DX.

Across most reanalyses of randomised trials or cohorts with long-term follow up (including those adjusted for clinical factors and those which were not), all 4 tests showed statistically significant prognostic ability for 10-year distant recurrence. Oncotype DX (using cutoffs of RS less than 18 and greater than 30) tended to assign more people to the low-risk group than seen in the studies of EPClin, MammaPrint or Prosigna. For full detail please see table 5 (page 50) in the external assessment report.

Data from MINDACT was confounded as the result of the MammaPrint test influenced chemotherapy prescribing, but within the mAOL high-risk group, outcomes were generally better for people with MammaPrint low-risk versus MammaPrint high-risk. No significance tests or hazard ratios were reported.

RxPONDER was limited to people with RS of 25 or less. In this population, Oncotype DX was significantly prognostic for 5-year invasive disease-free survival (DFS) after adjusting for clinical factors, both in the overall population and in the pre- and post-menopausal subgroups.

Observational data on Oncotype DX was similarly confounded by the influence of test results on chemotherapy prescribing. Despite greater chemotherapy use in people with higher recurrence scores, use of the test was significantly prognostic for 5-year distant recurrence-free interval (DRFI) in the Clalit registry (Stemmer et al. 2017), although this was not seen in a subgroup of people 70 years or older. Other registries (Petkov et al. 2016; Roberts et al. 2017; Ibraheem et al. 2019; Ibraheem et al. 2020) found significant prognostic effect for breast cancer-specific survival and overall survival.

For full prognostic data, see section 3.4 (page 48) and appendix 4 (page 182) in the external assessment report.

### Prediction of chemotherapy benefit

The predictive ability of a test is determined by whether the effect of chemotherapy versus no chemotherapy on clinical outcomes differs between test risk groups or ranges. Predictive ability is generally assessed using a statistical test for interaction between chemotherapy effect and risk score. A significant result on an interaction test indicates that risk score or risk category is predictive of chemotherapy benefit. Only 3 publications presented an interaction test for prediction of chemotherapy benefit (1 for MammaPrint and 2 for Oncotype DX). The effect of chemotherapy according to test result was also presented from the MINDACT randomised controlled trial and from 3 registries using Oncotype DX. No predictive data was identified for EPclin or Prosigna in a LN positive population. For full details see section 3.5 (page 57) and appendix 6 (page 194) in the external assessment report.

### MammaPrint

Mook et al. 2009 reported a reanalysis of a European cohort study. There was no significant interaction between MammaPrint score and the effect of chemotherapy on breast cancer-specific survival (p=0.95).

The MINDACT trial (Piccart et al. 2021) reported data for the effect of chemotherapy on 8-year distant metastasis-free survival (DMFS) for people with HR positive, HER2 negative, LN positive disease who scored high-risk on mAOL but low-risk with MammaPrint. The effect of chemotherapy had a non-significant hazard ratio (HR 0.84; 95% CI 0.51 to 1.37; p not reported). Data for the high-risk mAOL, high-risk MammaPrint group was not reported, so it was not possible to determine whether MammaPrint was predictive for chemotherapy benefit from the MINDACT data.

### Oncotype DX

Albain et al. 2010 was a reanalysis of a randomised controlled trial in which Oncotype DX was done retrospectively on samples from post-menopausal women randomised to chemo-endocrine therapy or endocrine monotherapy. Results were reported for people in low (RS 0 to 17), intermediate (18 to 30) and high-risk (31+) groups. For 10-year DFS, adjusted hazard ratios for chemotherapy versus no chemotherapy indicated a significant effect of chemotherapy for people in the high-risk group, but not for those in the intermediate or low-risk groups. Similar results were seen for DFS at different time points and for breast cancer-specific survival (BCSS) and overall survival (OS). However, the statistical significance of the interaction of chemotherapy effect and risk group was variable. For DFS and OS, interactions in the first 5 years were statistically significant while those events between 5 to 10 years were not. Over 0 to 10 years, the interaction for DFS was almost statistically significant if adjusted for the number of positive lymph nodes (p=0.053) and was stated to be statistically significant when adjusted for various clinical factors. But it was not significant when adjusted for Allred-scored ER status (p=0.15).

In the RxPONDER prospective randomised controlled trial (Kalinsky et al. 2021), chemotherapy showed no benefit for post-menopausal women with an RS of 0 to 25 on measures of distant recurrence (5-year DRFS or DRFI) or invasive DFS. However, in pre-menopausal women with an RS of 0 to 25 there was a statistically significant benefit of chemotherapy. There was no significant interaction between RS and the effect of chemotherapy on invasive DFS ($p>0.05$) for the whole cohort or in the pre- and post-menopausal subgroups. So, there was no statistically significant predictive effect within the RS 0 to 25 group. Since people with RS of 26 or higher were not included in the trial, it could not provide evidence on whether there is a predictive effect between groups with RS 0 to 25 and RS 26 to 100.

There was evidence from the Clalit registry in Israel (n=709) and the SEER (n=2,588) and NCDB (n=28,591) registries in the US on the outcomes of people with and without chemotherapy alongside Oncotype DX RS. However, since use of chemotherapy was not randomised, the data should be interpreted with caution. No publications reported an interaction test between RS and chemotherapy benefit. The relationship between RS and chemotherapy benefit was unclear. In the Clalit registry, 5-year DRFI was significantly higher with chemotherapy than without for people with RS 26 to 100 ($p=0.017$), but not for people with RS 31 to 100. Although data was not reported by menopausal status, some analyses of 5-year OS from the NCDB registry for older people with RS 0 to 25 found a significant benefit of chemotherapy, while others did not. So, the results do not clearly support or refute the RxPONDER findings.

## Decision impact

Decision impact studies assess how recommendations or decisions to use or not to use chemotherapy change before and after the test. The EAG identified 12 studies based in the UK (5 studies) or Europe (7 studies), all examining the effect of Oncotype DX results on chemotherapy recommendations or decisions. No studies were found that looked at the decision impact of other

test results. For full detail please see section 3.6 (page 69) in the external assessment report.

Generally, fewer people had chemotherapy after receiving Oncotype DX results than they would have if no test had been done. In the UK-based studies, the proportion of people with a chemotherapy recommendation or decision after testing reduced by 28% to 75%, while in the European studies the reduction was 12% to 73%. Within studies reporting data by Oncotype DX risk group, there were greater reductions in chemotherapy recommendation in the low-risk and intermediate-risk groups than in the high-risk groups, although the cutoffs used were variable.

## Health-related quality of life outcomes

No studies were identified that assessed health-related quality of life (HRQoL) or anxiety associated with tumour profiling testing in the LN positive population, so the EAG summarised studies in a broader early breast cancer population (LN negative or mixed nodal status) that were identified in the NICE diagnostics guidance 34 assessment. Six studies were included, of which 1 used EndoPredict, 1 used MammaPrint, 2 used Oncotype DX and 2 used Prosigna. Some studies reported a significant improvement in anxiety after testing, while others found no difference in HRQoL or anxiety. Anxiety generally decreased in people whose treatment plan changed from chemotherapy to no chemotherapy after testing, but increased when treatment was upgraded to chemotherapy, or if both clinical assessment and the test result indicated high risk. For full detail please see section 3.7 (page 75) in the external assessment report.

## Prespecified subgroups

A number of subgroups were specified in the scope (see population).

- No studies reported comparisons of the tumour profiling tests against currently-used tools to assess clinical risk such as PREDICT or NPI (MINDACT reported data by mAOL risk but this tool is no longer available).
- There was very little data on men with breast cancer. A subgroup analysis of the SEER database (Massarweh et al. 2018) reported significant prognostic ability of Oncotype DX for BCSS and OS in men with breast cancer.
- In RxPONDER, differences in 5-year invasive DFS within the RS 0 to 25 group were reported according to ethnicity (Abdou et al. 2023: white, 92%; black, 87%; Asian, 94%). But, no prognostic or predictive data were reported by ethnicity. In another subgroup analysis of the SEER database (Petkov et al. 2016), Oncotype DX was only significantly prognostic in white participants, but this was based on small numbers.
- No data was identified on people with comorbidities that may be particularly affected by the side effects of chemotherapy.

# 3 Cost effectiveness evidence

## Systematic review of cost-effectiveness evidence

The external assessment group (EAG) did a systematic review to identify any published economic evaluations of tumour profiling tests to guide treatment decisions in people with ER positive, HER2 negative, LN positive early breast cancer. This included studies that were identified during the development of NICE diagnostics guidance 34 (DG34) as well as those published later. The aim was to explore the methodological choices in these evaluations, rather than to assess the conclusions reached. Seven previously identified studies and 5 new studies were included in the review. These largely applied a similar modelling approach as the EAG's model (see economic analysis below). Some studies compared tumour profiling tests against chemotherapy for all, while others compared the tests against existing decision-making. Only 1 study included all 4 tumour profiling tests specified in the scope (Harnan et al.

2019, reporting on DG34), but as newer evidence was available, the EAG did not consider it appropriate to use this directly in this assessment. Find the full systematic review results in section 4.1 (page 77) of the external assessment report.

## Manufacturer submissions

The EAG also reviewed economic submissions from the manufacturers of MammaPrint and Oncotype DX. The EAG had serious concerns with several assumptions made in the MammaPrint model. In the EAG's exploratory analyses using preferred assumptions and correcting errors, the incremental cost-effectiveness ratio (ICER) for MammaPrint changed from dominating usual care to being dominated by usual care. The Oncotype DX model was considered to be generally well-programmed and free from errors. It concluded that Oncotype DX dominated current practice for post-menopausal women but was dominated in pre-menopausal women. The EAG made several minor adjustments to this model but these did not change the economic conclusions. Neither model was used as a base for the EAG's economic assessment, although the general structures were similar (see model structure). The full review and critique of the models and associated reports can be found in section 4.2 (page 84) of the external assessment report.

## Economic analysis

The EAG constructed an economic model to assess the cost effectiveness of tumour profiling tests versus current decision-making processes, based on the model used in NICE diagnostics guidance 34. The model assessed the health outcomes and costs associated with each testing strategy over a lifetime horizon, from the perspective of the UK NHS and Personal Social Services. Find a full description of the model in section 4.3 (page 108) of the external assessment report.

## Population

The population modelled was women with ER or PR positive, HER2 negative, LN positive (1 to 3 nodes) early breast cancer.

## Interventions

The interventions were all tumour profiling tests specified in the scope. Oncotype DX was assessed as both a 3-level test (RS 0 to 17, 18 to 30 and 31+), and as a 2-level test (RS 0 to 25 and 26+). This was to reflect the different ways that the test can be used.

## Comparator

The comparator was current decision making. This may be informed by tools such as PREDICT or NPI, or through consideration of clinical and pathological features – no specific tool was modelled. No incremental analysis was done comparing the tests against each other because:

- There were different evidence sources used for clinical outcomes between tests
- Evidence on menopausal status differed between the tests
- In TransATAC, overlapping but non-identical samples were used for different tests.

## Base cases

The EAG presented 7 base cases, which are outlined in Table 1:

**Table 1: EAG base case scenarios**

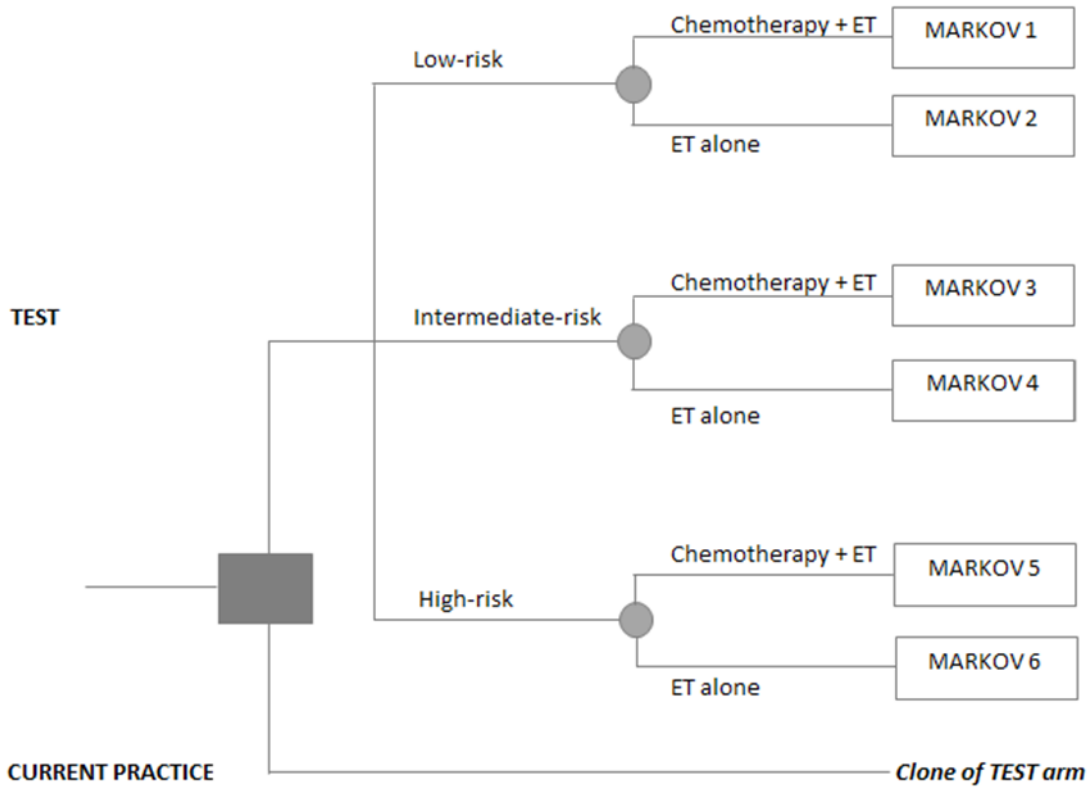
| Base case | Test | Menopausal status | Details |
|---|---|---|---|
| BC1 | Oncotype DX | Pre-menopausal | RxPONDER (2 level test, predictive) |
| BC2 | Oncotype DX | Post-menopausal | RxPONDER (2 level test, predictive) |
| BC3 | Oncotype DX | Post-menopausal | TransATAC (3 level test, predictive) |
| BC4 | Oncotype DX | Post-menopausal | TransATAC (3 level test, non-predictive) |

| | | | |
|---|---|---|---|
| BC5 | Prosigna | Post-menopausal | TransATAC (non-predictive) |
| BC6 | EPclin | Post-menopausal | TransATAC (non-predictive) |
| BC7 | MammaPrint | 33% pre-menopausal | MINDACT clinical high risk (non-predictive) |

## Model structure

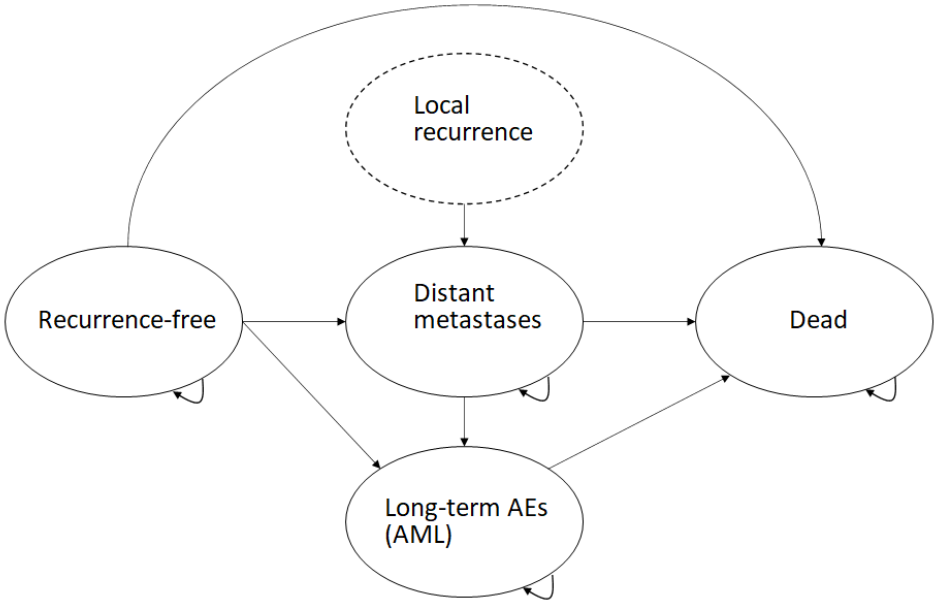
The model consisted of a hybrid decision tree and long-term Markov model. In the decision tree, people were stratified into either 2 or 3 groups according to genomic risk (depending on whether 1 or 2 thresholds are used). Within these groups, they then received either chemo-endocrine therapy or endocrine monotherapy (see Figure 1). People in the current practice arm were assigned to chemotherapy in the same proportion regardless of underlying genomic risk. For people in the test arm the probability of receiving chemotherapy was dependent on the test result (see Model inputs).

**Figure 1: EAG model decision tree**




Each branch then entered a Markov model which predicted lifetime quality-adjusted life years (QALYs) and costs according to risk of distant metastasis, chemotherapy prescription, and the effect of chemotherapy on the rate of distant metastasis. If predictive benefit was included, then the test result influenced the efficacy of chemotherapy. The model included 4 possible health states: recurrence-free, distant metastases, acute myeloid leukaemia (AML) or dead (see Figure 2). People entered the model in the recurrence at age 62 if post-menopausal or 44 if pre-menopausal, and continued until the cohort had reached age 100.

**Figure 2: EAG model Markov structure**




AE, adverse event; AML, acute myeloid leukaemia

The impact of local recurrence was applied as a one-off cost and QALY loss for a proportion of people with distant metastases (see model inputs). Adverse events of chemotherapy were modelled as a short-term QALY loss for the first year for people who had chemotherapy, and through the inclusion of the AML state.

## Model inputs

Find the full list of model parameters in section 4.3.3 (page 115) of the external assessment report.

### Risk classification probabilities

Risk classification probability refers to the probability of a person being assigned to a low-, intermediate- or high-risk group by the test. The probabilities for each base case (BC) are shown in Table 2.

**Table 2: Risk classification probabilities**

| Scenario | Description | Source | P low risk | P intermediate risk | P high risk |
|---|---|---|---|---|---|
| BC1 | Oncotype DX (2-level, pre-menopausal) | RxPONDER Kalinsky 2021 | 0.89 | - | 0.11 |
| BC2 | Oncotype DX (2-level, post-menopausal) | RxPONDER Kalinsky 2021 | 0.89 | - | 0.11 |
| BC3 | Oncotype DX (3-level, predictive) | TransATAC Sestak 2018 | 0.57 | 0.32 | 0.11 |
| BC4 | Oncotype DX (3-level, non-predictive) | TransATAC Sestak 2018 | 0.57 | 0.32 | 0.11 |
| BC5 | Prosigna | TransATAC Sestak 2018 | 0.08 | 0.32 | 0.60 |
| BC6 | EPclin | TransATAC Sestak 2018 | 0.23 | - | 0.77 |
| BC7 | MammaPrint | MINDACT Piccart 2021 | 0.69 | - | 0.31 |

BC, base case; P, probability.

### Distant recurrence-free interval with endocrine monotherapy

Distant recurrence-free interval (DRFI) estimates for people receiving endocrine therapy without chemotherapy were sourced from the RxPONDER, TransATAC and MINDACT trials. The timepoints for reporting DRFI differed between the trials, so the cumulative probability of DRFI was converted to a 6-month probability for the economic model, assuming a constant rate over time (Table 3). In BC7, the DRFI estimate for women who are clinical high risk and genomic high risk but receive endocrine monotherapy was not available from MINDACT data. So, the EAG's value was calculated by applying a hazard ratio from the Early Breast Cancer Trialists' Collaborative Group meta-analysis (EBCTCG 2012) to the clinical high risk, genomic high risk group. For more detail, please see table 32 and accompanying text (page 118) in the external assessment report.

**Table 3: 6-month probabilities of distant recurrence by test risk category**

| Scenario | Description | Source | 6m cumulative DRFI low risk | 6m cumulative DRFI int risk | 6m cumulative DRFI high risk |
|---|---|---|---|---|---|
| BC1 | Oncotype DX (2-level, pre-menopausal) | RxPONDER Kalinsky 2021 TransATAC Sestak 2018 | 0.0063 | - | 0.0236 |
| BC2 | Oncotype DX (2-level, post-menopausal) | RxPONDER Kalinsky 2021 TransATAC Sestak 2018 | 0.0035 | - | 0.0236 |
| BC3 | Oncotype DX (3-level, predictive) | TransATAC Sestak 2018 | 0.0107 | 0.0170 | 0.0236 |
| BC4 | Oncotype DX (3-level, non-predictive) | TransATAC Sestak 2018 | 0.0107 | 0.0170 | 0.0236 |
| BC5 | Prosigna | TransATAC Sestak 2018 | 0.0000 | 0.0115 | 0.0182 |
| BC6 | EPclin | TransATAC Sestak 2018 | 0.0029 | - | 0.0179 |
| BC7 | MammaPrint | MINDACT IPD | 0.0057 | - | 0.0184 |

6mDRFI, 6-month distant recurrence-free interval; BC, base case; int, intermediate; IPD, individual patient data.

### Effect of chemotherapy on distant recurrence

The relative treatment effects for chemo-endocrine therapy versus endocrine monotherapy differed between the base case scenarios depending on source and whether the test was assumed to be predictive of chemotherapy benefit (Table 4).

- For the analysis of Oncotype DX using a single cutoff (BC1 and BC2), the model applied competing risk adjusted hazard ratios (HRs) by menopausal

subgroup (RxPONDER, Kalinsky et al. SABCS 2021). As women with RS 26 to 100 were excluded from RxPONDER, the HR for chemotherapy in this group was based on the HR for women with RS 31 to 100 reported by Albain et al. 2010. This indirectly assumed that Oncotype DX was predictive of chemotherapy benefit.

- For the analysis of Oncotype DX as a 3-level test but assuming predictive benefit (BC3), the model applied different HRs by Oncotype DX risk category (RS 0 to 18; 18 to 30; 31 to 100) based on Albain et al. 2010.
- For the analyses of all tests without predictive benefit (BC4 to BC7), the model applied an HR for DRFI based on the EBCTCG meta-analysis (2012).

**Table 4: Hazard ratios for distant recurrence with chemotherapy versus no chemotherapy (median values)**

| Scenario | Description | HR low risk | HR intermediate risk | HR high risk | Source |
|---|---|---|---|---|---|
| BC1 | Oncotype DX (2-level, pre-menopausal) | 0.64 | - | 0.59 | RS 0 to 25: Kalinsky SABCS 2021<br>RS 26 to 100: Albain 2010 |
| BC2 | Oncotype DX (2-level, post-menopausal) | 1.12 | - | 0.59 | RS 0 to 25: Kalinsky SABCS 2021<br>RS 26 to 100: Albain 2010 |
| BC3 | Oncotype DX (3-level, predictive) | 1.02 | 0.72 | 0.59 | Albain 2010 |
| BC4 to BC7 | Non-predictive scenarios | 0.71 | 0.71 | 0.71 | EBCTCG 2012 |

BC, base case; EBCTCG, Early Breast Cancer Trialists' Collaborative Group; HR, hazard ratio; RS, recurrence score; SABCS, San Antonio Breast Cancer Symposium.

### Probability of receiving chemotherapy

The probabilities of receiving chemotherapy were sourced from data submitted by Holt et al. This currently unpublished trial examined the impact of Oncotype DX on decisions about chemotherapy for people with LN positive early breast cancer in NHS hospitals from 2017 to 2022. Data was available for use of Oncotype DX as a 3-level test (RS 0 to 17, 18 to 30 and 31 to 100),

and as a 2-level test (RS 0 to 25 and 26 to 100). In the absence of decision impact data for other tumour profiling tests, the results from this trial were used to inform the decision impact of all, assuming that tests would be interpreted in the same way as other tests with the same number of risk categories (Table 5).

**Table 5: Pre- and post-test probabilities of receiving chemotherapy**

| Scenario | Description | P chemotherapy low risk | P chemotherapy intermediate risk | P chemotherapy high risk |
|---|---|---|---|---|
| Pre-test | - | 0.80 | 0.80 | 0.80 |
| BC1 | 2-level, pre-menopausal | 0.37 | - | 0.96 |
| BC2, BC6 | 2-level, post-menopausal | 0.11 | - | 0.96 |
| BC3 to BC5 | 3-level, post-menopausal | 0.08 | 0.49 | 0.98 |
| BC7 | MammaPrint, 33% pre-menopausal | 0.19 | - | 0.96 |

BC, base case; P, probability.

### Pathway inputs

Other parameters relating to probabilities of events after the decision to offer chemotherapy are presented in Table 6. People with distant metastases were assumed to be treated with ribociclib and letrozole. People who developed acute myeloid leukaemia were assumed to be treated with liposomal cytarabine/daunorubicin. As in previous models, local recurrence was only considered in people who had distant metastases, and was applied as a one-off cost and QALY loss. For more detail, please see pages 121 to 123 in the external assessment report.

**Table 6: Downstream parameters**

| Parameter | Value | Source |
|---|---|---|
| 6-month probability of death due to distant metastases (assuming treatment with ribociclib+letrozole) | 0.1025 | Estimated from Suri et al. 2017 |
| Probability of previous local recurrence for those with distant metastases | 0.1054 | de Bock et al. 2009 |
| 6-month probability of developing AML after chemotherapy | 0.00025 | Wolff et al. 2015 |
| 6-month probability of death due to AML (assuming treatment with liposomal cytarabine/daunorubicin) | 0.1977 | Estimated from Bewersdorf et al. 2022; Martin et al. 2010; ONS life tables |
| All-cause mortality | Varied by age | ONS life tables |

AML, acute myeloid leukaemia; ONS, Office of National Statistics.

## Costs

Find the full detail of costs used in the EAG's model from page 125 of the external assessment report. Costs were uplifted to current prices where appropriate. A summary of costs is presented in Table 7.

**Table 7: Summary of costs**

| Component | Mean cost | Source |
|---|---|---|
| EPclin | £1,500.00 | Manufacturer |
| MammaPrint | £2,616.00 | Manufacturer |
| Oncotype DX | £2,580.00 | Manufacturer |
| Prosigna | £1,896.00 | Manufacturer |
| Adjuvant chemotherapy (one-off). Includes drug costs, pharmacy costs, outpatient monitoring and tests. | £7,410.48 | Berdunov 2022; eMIT; NHS reference costs; Ward 2013 |
| Endocrine therapy years 1-2 (per cycle) | £66.95 | Clinical input; eMIT; BNF, PSSRU |
| Endocrine therapy years 3-5 (per cycle) | £66.44 | Clinical input; eMIT; BNF, PSSRU |
| Endocrine therapy years 6-10 (per cycle) [80% of cohort] | £53.16 | Clinical input; eMIT; BNF, PSSRU |
| Bisphosphonates (per cycle) [60% post-menopausal women] | £320.84 | Clinical input; NHS reference costs |
| Ovarian suppression (per cycle) [60% pre-menopausal women] | £496.73 | Clinical input; NHS reference costs; PSSRU |
| Adverse events from chemotherapy (once-only) | £1,249.58 | Ellis 2009; NHS reference costs |
| Follow-up, year 1 (per year) | £360.48 | Clinical input; NHS reference costs; Ward 2013 |
| Follow-up, years 2-5 (per year) | £139.00 | Clinical input; NHS reference costs; Ward 2013 |
| Local recurrence (once-only) | £16,494.23 | Karnon 2007 |
| Distant metastases (once-only) | £117,482.09 | Suri 2017 |
| Acute myeloid leukaemia (once-only) | £132,185.91 | Bewersdorf 2022; Zeidan 2016; Lancet 2018; NCPE report |
| End of life care (once-only) | £4,898.17 | Hinde 2019 |

BNF, British National Formulary; eMIT, electronic market information tool; NCPE, National Centre for Pharmacoeconomics; PSSRU, Personal Social Services Research Unit. NHS reference costs were for 2021/2022.

### Tumour profiling test costs

Table 7 reports the list prices of the intervention tests. MammaPrint and Oncotype DX are processed outside the NHS, and the test costs include the costs of shipping and other costs associated with processing the sample and reporting test result. EPclin and Prosigna costs reflect the test being done in

NHS laboratories, and include all reagents and consumables. Confidential price discounts apply to EPclin, Oncotype DX and Prosigna.

### Cost of treating distant metastases and acute myeloid leukaemia (AML)

The lifetime cost of treating distant metastases was based on treatment with ribociclib plus letrozole as reported in Suri et al. 2017. The lifetime cost of treating AML includes the cost of liposomal cytarabine/daunorubicin. Both therapies are used in the NHS with confidential price discounts.

## Health-related quality of life and QALY losses

Utility values and QALY losses are summarised in Table 8. For full details of how these were determined, please see pages 123 to 125 in the external assessment report. QALY losses for chemotherapy adverse events and local recurrence were applied as one-off decrements for 1 year.

**Table 8: Utility values and QALY losses**

| Parameter | Mean value | Source |
|---|---|---|
| Utility, recurrence-free | 0.824 | Lidgren 2007 |
| Utility, distant metastases | 0.685 | Lidgren 2007 |
| Utility, acute myeloid leukaemia | 0.590 | Estimated by dividing the mean QALYs by mean life-years gained in the rebuilt model based on Bewersdorf 2022 |
| QALY loss chemotherapy | −0.038 | Campbell 2011 |
| QALY loss local recurrence | −0.108 | Campbell 2011 |

QALY, quality-adjusted life year.

## Key assumptions

For a full list of assumptions please see page 115 in the external assessment report. Some key assumptions not already described include:

- The risk of distant metastases with endocrine monotherapy was assumed to remain constant over time
- The risk of death for people in the recurrence-free state was assumed to be the same as the age-matched general female population

- The risk of death in women with distant metastases or acute myeloid leukaemia was constrained to be at least equal to that of the general female population.

## Scenario analyses

The EAG examined a number of deterministic scenario analyses to explore alternative assumptions or evidence sources as outlined in Table 9. For full details, please see page 131 in the external assessment report.

**Table 9: Summary of deterministic scenario analyses**

| Scenario | Parameter | Base case | Scenario |
|---|---|---|---|
| DSA1 | Proportion with Oncotype RS 26 to 100 | 11% | 17% |
| DSA2 | Prosigna test category probabilities and DRFI estimates (low/int/high) | Source: TransATAC<br>Probability: 0.08/0.32/0.60<br>10-year DRFI: 0.00/0.21/0.31 | Source: Gnant 2014<br>Probability: 0.04/0.34/0.62<br>10-year DRFI: 0.00/0.06/0.24 |
| DSA3 | EPclin test category probabilities and DRFI estimates (low/high) | Source: TransATAC<br>Probability: 0.23/0.77<br>10-year DRFI: 0.06/0.30 | Source: Filipits 2019<br>Probability: 0.35/0.65<br>15-year DRFI: 0.15/0.25 |
| DSA4, DSA5, DSA6, DSA8 | Post-test chemotherapy probabilities for 3-level tests (low/int/high) | Source: Holt 2023<br>See Table 5 | Source: Llombart-Cussac 2023 (0.09/0.46/1.00); Loncaster 2017 (0.08/0.63/0.83); Zambelli 2020 (0.01/0.33/1.00); Harnan 2019 (0.31/0.72/0.95) |

| | | | |
|---|---|---|---|
| DSA7, DSA9 | Post-test chemotherapy probabilities for 2-level tests (low/high) | Source: Holt 2023<br>See Table 5 | Source: Dieci 2019 (0.22/1.00); Harnan 2019 (0.36/0.96) |
| DSA10 | Risk of distant metastases decreases over time | No tapering | Risk decreases by 50% after 10 years and drops to 0% at 15 years |
| DSA11 to DSA13 | Hazard ratio for chemotherapy versus no chemotherapy | Predictive benefit assumed for BC1 to BC3 (see Table 4); Prognostic benefit only for BC4 to BC7 (HR=0.71 for all risk categories) | HR equal for all tests and risk categories (prognostic benefit only):<br>DSA11: 0.60<br>DSA12: 0.71<br>DSA13: 0.80 |
| DSA14 to DSA16 | QALY loss from chemotherapy | −0.038 | QALY loss was halved (DSA14), doubled (DSA15) or tripled (DSA16) |
| DSA17 | Pre-test probability of receiving chemotherapy | 0.8 | 0.9 |
| DSA18, DSA19 | Starting age of population | 62 (post-menopausal)<br>44 (pre-menopausal) | Starting age increased (DSA18) or decreased (DSA19) by 5 years |
| DSA20 | Utility values for recurrence-free and distant metastases health states | Source: Lidgren 2007<br>Recurrence free: 0.824<br>Distant metastases: 0.685 | Source: Verrill 2020<br>Recurrence free: 0.73<br>Distant metastases: 0.60 |
| DSA21 | Probability of developing AML | 6-month probability after chemotherapy: 0.00025 | 6-month probability after chemotherapy: 0 |
| DSA22, DSA23 | Cost of chemotherapy | £7,410.00 | Cost was halved or doubled |
| DSA24, DSA25 | Lifetime cost of treating distant metastases | £117,482 | Cost was halved or doubled |
| DSA26, DSA27 | Lifetime cost of treating AML | £132,186 | Cost was halved or doubled |

AML, acute myeloid leukaemia; BC, base case; DRFI, distant recurrence-free interval; DSA, deterministic scenario analysis; HR, hazard ratio; QALY, quality-adjusted life year; RS, recurrence score.

## Base case results

The EAG noted that Exact Sciences, the manufacturer of Oncotype DX, has indicated that it intends for the test to be used with a single cutoff (recurrence score 25) for people with LN positive disease. So, BC3 and BC4 are less relevant for decision making. As such, only results from BC1 and BC2 are presented here for Oncotype DX. For full details see section 4.3.6 (page 133) in the external assessment report.

As previously described, confidential price discounts are in place for EPclin, Oncotype DX and Prosigna, as well as for some therapies used downstream of testing (ribociclib and liposomal cytarabine/daunorubicin). The results of the cost effectiveness analysis presented in this overview are based on the list prices for the tests and medicines. The results incorporating the cost reductions are presented in a confidential appendix for the committee. Qualitative descriptions of the effects of these discounts are provided here for transparency.

An overview of the probabilistic base case results is presented in Table 10. Deterministic results were similar (see table 42, page 136 in the external assessment report). The EAG also evaluated further clinical and economic outcomes, which are summarised in Table 11.

**Table 10: Summary of probabilistic base-case results**

| Base case | Source for risk classification probabilities and DRFI | Inc. QALYs | Inc. costs | List price ICER (£/QALY gained) | ICER with confidential price reductions applied (£/QALY) |
|---|---|---|---|---|---|
| BC1: Oncotype DX, 2-level, pre-menopausal, predictive | RxPONDER/TransATAC | −0.18 | £1,810 | Dominated | SWQ ICER <£20,000 saved/QALY lost |
| BC2: Oncotype DX, 2-level, post-menopausal, predictive | RxPONDER/TransATAC | 0.11 | −£4,273 | Dominating | Dominating |
| BC5: Prosigna, post-menopausal, non-predictive | TransATAC | 0.03 | £1,084 | £39,357 | <£30,000/QALY gained |
| BC6: EPclin, post-menopausal, non-predictive | TransATAC | 0.06 | £231 | £4,113 | Dominating |
| BC7: MammaPrint, 33% pre-menopausal, non-predictive | MINDACT | −0.07 | £786 | Dominated | Dominated |

BC, base case; DRFI, distant recurrence-free interval; ICER, incremental cost-effectiveness ratio; Inc, incremental; QALY, quality-adjusted life year; SWQ, south-west quadrant (lower cost and lower QALYs than current practice).

**Table 11: Incremental clinical and economic outcomes per 1,000 women tested (test versus current decision-making)**

| Base case | N receiving chemotherapy | Infusion chair hours | N with distant metastases in lifetime | Cost to NHS and PSS | List price NHB (£20k/QALY) | List price NHB (£30k/QALY) |
|---|---|---|---|---|---|---|
| BC1: Oncotype DX, 2-level, pre-menopausal, predictive | −361 | −1,854 | 41 | £1,786,628 | −267 | −237 |
| BC2: Oncotype DX, 2-level, post-menopausal, predictive | −594 | −3,051 | −13 | −£4,282,569 | 327 | 255 |
| BC5: Prosigna, post-menopausal, non-predictive | −46 | −235 | −3 | £1,107,509 | −28 | −9 |
| BC6: EPclin, post-menopausal, non-predictive | −39 | −203 | −8 | £305,191 | 39 | 45 |
| BC7: MammaPrint, 33% pre-menopausal, non-predictive | −370 | −1,900 | 24 | £791,671 | −105 | −92 |

BC, base case; N, number; NHB, net health benefit; PSS, personal social services; QALY, quality-adjusted life year

**Oncotype DX**

Oncotype DX was found to be dominated by current practice for use in pre-menopausal women (BC1), that is, Oncotype DX was less clinically effective and more expensive than current practice. This was driven by the reduction of chemotherapy prescribing for people who would have benefitted from the treatment, and therefore worse clinical outcomes for these people. This was maintained across all deterministic scenario analyses, except where the cost of chemotherapy was doubled, which moved the incremental cost-effectiveness ratio into the south-west quadrant, that is, Oncotype DX was less clinically effective and cheaper than current practice (£5,007 saved per QALY lost).

In the post-menopausal population (BC2), Oncotype DX dominated current practice, that is, Oncotype DX was more effective and less expensive than current practice. This result was driven by a large reduction in chemotherapy use in people who would not have benefitted from the treatment. This was maintained in all scenarios where predictive benefit was assumed. However, if the test was assumed to be non-predictive (DSA11, 12 and 13), then the ICERs moved into the south-west quadrant, that is use of Oncotype DX resulted in lower costs and fewer QALYs than current practice. This was because more people had distant recurrence with testing than in current practice in these scenarios. The size of the cost saving increased as the efficacy of chemotherapy decreased (Table 12), as did the probability that QALYs would be gained rather than lost (see addendum). However, the scenario analyses all assumed higher efficacy of chemotherapy (HR 0.6 to 0.8) than was observed in the low-risk group in RxPONDER (HR 1.12). With confidential price reductions applied, the ICERs were all above £30,000 saved per QALY lost. Similarly, when Oncotype DX was used as a 3-level test, it was dominating when predictive benefits were included (BC3), but not when the test was assumed to be non-predictive (BC4).

**Table 12: Deterministic scenario analyses for Oncotype DX in post-menopausal women with and without predictive benefit**

| Scenario | HR for distant recurrence CT versus no CT RS 0 to 25 | HR for distant recurrence CT versus no CT RS 26 to 100 | Inc. QALYs | Inc. costs | List price ICER (£ saved/QALY lost) |
|---|---|---|---|---|---|
| Deterministic BC2 | 1.12 | 0.59 | 0.11 | −£4,283 | Dominating |
| BC2 DSA11 | 0.60 | 0.60 | −0.07 | −£638 | £9,772 (SWQ) |
| BC2 DSA12 | 0.71 | 0.71 | −0.03 | −£1,351 | £42,518 (SWQ) |
| BC2 DSA13 | 0.80 | 0.80 | −0.01 | −£1,882 | £279,599 (SWQ) |

BC, base case; CT, chemotherapy; DSA, deterministic scenario analysis; Inc, incremental; QALY, quality-adjusted life year; RS, recurrence score; SWQ, south-west quadrant (lower cost and lower QALYs than current practice).

### Prosigna

The probabilistic ICER for Prosigna versus current decision making in the base case was £39,357 per QALY gained. This was driven by a small decrease in the use of chemotherapy and in the lifetime probability of distant metastases, but with additional costs from testing. With confidential price reductions applied, the base case ICER was reduced to less than £30,000 per QALY gained. Deterministic scenario analyses indicated that the ICER was sensitive to:

- the source of test risk classification probabilities and associated DRFI estimates (the test was more cost-effective when the risk of distant recurrence was reduced)
- the HR for distant recurrence with or without chemotherapy (the test was more cost-effective when chemotherapy was more effective)
- the costs of drugs in the treatment pathway (the test was more cost-effective when the drugs were more expensive)
- the source of decision impact data (although all available data was from studies of Oncotype DX – see DSA4, 5, 6 and 8).

Several scenario analyses still resulted in ICERs over £30,000 per QALY with confidential price reductions applied.

### EPclin

The probabilistic ICER for EPclin versus current decision making in the base case was £4,113 per QALY gained. This was driven by a small decrease in the use of chemotherapy and in the lifetime probability of distant metastases, but with additional costs from testing. With confidential price reductions applied, EPclin dominated current practice. Deterministic scenario analyses indicated that the ICER was sensitive to a number of factors, but no scenarios resulted in an ICER above £30,000 per QALY gained. Only 1 scenario resulted in EPclin being dominated by current practice – if test risk classification probabilities and DRFI estimates were sourced from Filipits et al. 2019 rather than from TransATAC (DSA3, see Table 9).

### MammaPrint

MammaPrint was dominated by current practice when used in a clinical high-risk, 33% pre-menopausal population as reported in MINDACT, that is, MammaPrint was less clinically effective and more expensive than current practice. This was driven by a large decrease in the number of people receiving chemotherapy, an increase in the probability of developing distant metastases, and additional costs of testing. With confidential price reductions for the downstream medications applied, MammaPrint remained dominated by current practice. Across all deterministic scenario analyses, MammaPrint was dominated or had south-west quadrant ICERs.

# 4 Summary

## Clinical effectiveness

There was evidence from reanalyses of randomised trials and cohort studies that all 4 tests were prognostic for distant recurrence. Oncotype DX tended to

assign more people to low-risk categories than the other tumour profiling tests.

Data on predictive ability in the LN positive population was only available for MammaPrint and Oncotype DX. The EAG noted that new studies on predictive ability of tumour profiling tests are unlikely due to ethical concerns, particularly for people who are considered to have clinical or genomic high-risk of distant recurrence.

It was not possible to determine from the MINDACT trial whether MammaPrint was predictive of chemotherapy benefit because data was not reported for the clinical high-risk, MammaPrint high-risk subgroup. The only other study for MammaPrint did not find a significant interaction between MammaPrint score and the effect of chemotherapy.

For Oncotype DX, reanalysis of the SWOG-8814 randomised controlled trial (Albain et al. 2010) reported a significant effect of chemotherapy for people with RS 31 to 100, but not those in intermediate (RS 18 to 30) or low-risk (RS 0 to 17) groups. It found significant interactions between risk group in some analyses, but not in others. The RxPONDER randomised controlled trial reported significant chemotherapy benefit in pre-menopausal women with RS 0 to 25, but not for post-menopausal women. However, a test for interaction between RS (from 0 to 25) and the effect of chemotherapy on invasive disease-free survival was not statistically significant. As people with RS 26 to 100 were excluded from the study, it was not possible to determine whether the test was predictive for chemotherapy benefit between people with RS 0 to 25 and those with RS 26 to 100. Data from registries and databases did not clearly support or refute the findings from RxPONDER.

Data on decision impact of tumour profiling tests was only available for Oncotype DX. In 5 UK-based studies, use of the test reduced the number of people with a chemotherapy recommendation or decision by 28% to 75%. In studies where decision impact by risk group was reported, greater reductions

in chemotherapy use were seen in low- and intermediate-risk groups than in the high-risk groups.

Aside from menopausal status, there was limited data identified on the pre-specified subgroups. No data was found comparing the tumour profiling tests against currently-used tools such as PREDICT or NPI, or on people with comorbidities that may be particularly affected by the side effects of chemotherapy. Only 1 analysis reported on men with breast cancer, finding that Oncotype DX was significantly prognostic for breast cancer-specific survival and overall survival. A subgroup analysis of RxPONDER found minor differences in invasive disease-free survival between ethnic groups, but did not report on prognostic or predictive data by ethnicity. A database analysis found that Oncotype DX was only significantly prognostic in white participants, but this was based on small numbers.

## Cost effectiveness

The EAG updated the economic model used to inform NICE diagnostics guidance 34, and produced base case analyses for each of the tests versus current decision making.

When using Oncotype DX as a 2-level test (as in the RxPONDER trial), it was dominated by current practice in a pre-menopausal population. However, in post-menopausal women Oncotype DX dominated current practice. This finding was dependent on an assumption of predictive benefit. When this assumption was removed, Oncotype DX had a south-west quadrant ICER (it cost less and produced fewer QALYs than current practice). The magnitude of the cost saving was inversely proportionate to the effectiveness of chemotherapy. However, the non-predictive scenario analyses all assumed a higher efficacy of chemotherapy than observed in the post-menopausal low-risk group (RS 0 to 25) in RxPONDER.

The probabilistic ICER for Prosigna versus current decision making in the base case was £39,357 per QALY gained, although with confidential price

reductions applied, this was reduced to less than £30,000 per QALY gained. The ICER was particularly sensitive to the source of test risk classification probabilities and DRFI estimates, and to the effectiveness of chemotherapy. Several scenario analyses resulted in ICERs over £30,000 per QALY with confidential price reductions applied.

EPclin dominated current practice when confidential price discounts were applied, although this was highly dependent on the source of test risk classification probabilities and DRFI estimates. When a different evidence source was used that reported a higher proportion of people classified as low risk, and a higher rate of distant recurrence, EPclin was instead dominated by current practice (or had a south-west quadrant ICER less than £20,000 saved per QALY lost with confidential price discounts applied).

In a population of clinical high risk, 33% pre-menopausal women, MammaPrint was dominated by current practice. This finding was consistent across sensitivity analyses.

The model predicted that use of the tumour profiling tests would reduce the number of people receiving chemotherapy. The size of the reduction was much larger with Oncotype DX (59%) or MammaPrint (37%), while EPclin and Prosigna resulted in more modest reductions (4% and 5% respectively).

# 5 Issues for consideration

## Population

All the tumour profiling tests except Prosigna are indicated for both pre- and post-menopausal women. However, there was not enough evidence on distant recurrence with endocrine monotherapy separated by menopausal status for the EAG to model these populations separately for EPclin and MammaPrint. For EPclin, only a post-menopausal population was modelled. For MammaPrint, a mixed population was modelled (33% pre-menopausal) that reflected the MINDACT LN positive population.

Data for Oncotype DX from RxPONDER was reported by menopausal status, so these populations were modelled separately (base case 1 and 2). RxPONDER (Kalinsky et al. 2021) found that there was a significant benefit to chemotherapy for pre-menopausal women with recurrence score (RS) 0 to 25, but not for post-menopausal women in this RS group. In the modelled pre-menopausal population, Oncotype DX was dominated by current practice because of a loss of quality-adjusted life years (QALYs) – this was consistent across all scenario analyses.

What can be concluded about the efficacy of each of the tests by menopausal status?

## Strength of evidence for predictive benefit

Only Oncotype DX had any evidence showing a significant interaction between risk score and chemotherapy effect. Albain et al. 2010 reported a significant interaction for 5-year disease-free survival (DFS) and overall survival (OS) when adjusted for the number of positive lymph nodes, but the interaction was not significant in years 5 to 10 (see page 58 in the external assessment report). Additionally, the age of this trial means that the chemotherapy regimens used are unlikely to reflect current practice.

The randomised clinical trial RxPONDER (Kalinsky et al. 2021) found that there was a significant benefit to chemotherapy for pre-menopausal women with RS 0 to 25, but not for post-menopausal women in this group. However, since people with RS 26 to 100 were not included in the trial, it could not provide evidence on whether there is a predictive effect between groups with RS 0 to 25 and RS 26 to 100.

The results of the EAG's economic modelling for Oncotype DX were highly dependent on whether or not predictive benefit was assumed. In the base case using Oncotype as a 2-level test in post-menopausal women, predictive benefit was indirectly assumed as the effect of chemotherapy for the RS 0 to 25 population was based on RxPONDER, but for the RS 26 to 100

population the effect was based on Albain et al. 2010. Under this assumption, Oncotype DX dominated current practice. Results using data from Albain et al. for all risk groups (using Oncotype DX as a 3-level test) were similar (see base case 3, Table 42 in the external assessment report, page 136). However, when the predictive benefit was removed, the incremental cost-effectiveness ratio (ICER) for Oncotype DX versus current practice moved to the south-west quadrant (it cost less and produced fewer QALYs than current practice – see Table 12). The efficacy of chemotherapy used in these analyses was higher than observed in the post-menopausal RS 0 to 25 group in RxPONDER. It was uncertain whether the change in QALYs would be positive or negative, particularly in analyses where chemotherapy was less effective.

Is it reasonable to assume predictive benefit of Oncotype DX for chemotherapy effectiveness?

## Source of test risk classification probabilities and distant recurrence estimates

The base case analyses of EPclin and Prosigna, test risk classification probabilities and distant recurrence-free interval (DRFI) values were taken from the TransATAC study (Sestak et al. 2018). The cost-effectiveness estimates were highly dependent on these values.

For EPclin, the base case analysis with TransATAC data produced a deterministic ICER of £5,580 per QALY gained. If values from Filipits et al. 2019 were used instead (which increased the proportion classified as low risk, and increased the chances of distant recurrence, see Table 9), EPclin was dominated by current practice. The populations in the TransATAC and Filipits studies were similar (HR positive, HER2 negative, post-menopausal, 1 to 3 positive lymph nodes). However, the estimate for distant recurrence-free rate (DRFR) from Filipits et al. was taken over 15 years, compared to the 10-year DRFI from TransATAC. At this time point, there were few people left in the

study. If using 10-year DRFR data from Filipits instead, EPclin dominated current practice.

With Prosigna, the base case analysis with TransATAC data produced a deterministic ICER of £40,220 per QALY gained. If instead, Gnant et al. 2014 was used, the ICER became £23,853. Gnant et al. reported on the Austrian ABCSG-8 trial, which was similar to TransATAC in that both compared different endocrine therapy regimens in post-menopausal women with HR positive, HER2 negative, LN positive breast cancer. However, the ABCSG-8 study included 11% people with 4 or more positive lymph nodes, and included more participants (413 versus 183 in TransATAC). Confidential price reductions further increased the cost-effectiveness of Prosigna.

Which are the preferred sources of test risk classification probabilities and distant recurrence estimates for EPclin and Prosigna?

## Source of decision impact data

Decision impact data was only available for studies of Oncotype DX. In the economic model, the EAG chose to use data on pre- and post-test probabilities of receiving chemotherapy from unpublished data submitted by Holt et al. (see Table 5). This was because it was a recent UK-based study that used Oncotype DX as both a 2-level test and as a 3-level test. In the absence of other data, the EAG assumed that other 2- or 3-level tests would be interpreted in the same way.

The effect of the source of decision impact data on cost-effectiveness was investigated in deterministic scenario analyses (see DSA4 to DSA9, Table 43 in the external assessment report). However, the alternative sources were also for Oncotype DX only, did not use multiple sets of thresholds, and many were not UK-based.

Is decision impact data on Oncotype DX generalisable to other tests?

### Impact on chemotherapy services

The EAG's model predicted that use of the tests in predominantly post-menopausal populations would reduce the number of people receiving chemotherapy, and therefore the number of infusion chair hours (see Table 11). This was most pronounced for MammaPrint and Oncotype DX. Tumour profiling tests could therefore free up infusion capacity for other indications, potentially reducing waiting lists and producing benefits outside of the assessment population. However, this benefit may be negated or outweighed if people who forego chemotherapy develop distant metastases at a higher rate than they would if they had received chemotherapy.

How should this potential uncaptured benefit be considered by the committee?

### OPTIMA

The OPTIMA trial is an ongoing randomised controlled trial of Prosigna, comparing test-directed chemotherapy use with standard chemotherapy prescribing. The population includes people with a high clinical risk of recurrence and are mostly LN positive (1 to 9 nodes). No results have yet been published.

## 6 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

All people with cancer are covered under the disability provision of the Equality Act (2010) from the point of diagnosis. Breast cancer is less common in men than women: in 2020, 44,943 women and 348 men were diagnosed with new cases of breast cancer in England (NHS Digital 2022). Breast cancer is underdiagnosed and often undertreated in men. Some tests may not be validated for use in men with breast cancer. Clinical and manufacturer advice

is that the tests can be used for men with breast cancer, but could perform differently. Only 1 study reported on men with breast cancer, finding that Oncotype DX was significantly prognostic for breast cancer-specific survival and overall survival.

Women from South Asian, Black African or Caribbean family backgrounds are more likely to have less favourable breast tumour characteristics at diagnosis (stage, grade, ER or HER2 status) compared with white women, and are more likely to be diagnosed younger (Gathani et al. 2021; Breast Cancer Now 2021). There was limited data on the effects of ethnicity. In 1 publication, data from a small number of participants seemed to indicate that the prognostic ability of Oncotype DX was not significant in non-white populations.

Clinical experts highlighted that some NHS trusts are already offering tumour profiling tests to inform chemotherapy decisions for people with LN positive early breast cancer through early or compassionate access schemes. Many began doing so during the COVID-19 pandemic to help relieve pressure on infusion services. There is therefore a level of geographic inequality in access to tumour profiling tests in the LN positive population.

# 7 Implementation

## Location of testing

Some tumour profiling tests have the option of testing samples in a local laboratory or sending samples away for testing in a centralised laboratory. The location of the testing may impact on factors such as test throughput, processing errors, quality assurance and the level of training required. Personal data management and confidentiality policies may need careful consideration for tests that are processed outside of the UK. Local laboratories may need additional resource to process increased testing volumes.

### Interpreting and acting on results

All of the tumour profiling tests provide a continuous score that is then categorised into risk groups. Some tumour profiling tests categorise results as low risk, intermediate risk or high risk of distant recurrence, whereas others report a binary risk level of either low or high. Some clinical experts noted that intermediate risk results could be problematic as they could introduce uncertainty about optimal treatment planning, although intermediate risk groups could also indicate that the clinical decision should be particularly carefully considered.

When trusts are new to tumour profile testing, agreement would need to be reached on who will take responsibility for acting on the test result. Training on test result interpretation and counselling for the person with breast cancer would also be required to support safe adoption.

### Patient selection

Different oncologists may use different risk assessment tools to decide who should be offered a tumour profiling test, for example NPI or PREDICT (or none at all). The choice of the initial decision-making tool may influence subsequent treatment options. No data was found comparing the tumour profiling tests to PREDICT or NPI.

## 8 Authors

**Jacob Grant**
Topic lead

**Frances Nixon**
Technical adviser

# Glossary

## Acute myeloid leukaemia

Acute myeloid leukaemia (AML) is a cancer of the white blood cells. Previous chemotherapy is a risk factor for AML.

## Adjuvant! Online

Adjuvant! Online was an online tool used to calculate 10-year survival probability for people with breast cancer based on age, comorbidities, tumour size, grade and ER status. In MINDACT, the tool was modified to also include HER2 status. Adjuvant! Online is no longer available.

## Adjuvant therapy

Additional treatment given as well as the primary treatment to improve outcomes.

## Distant recurrence

Cancer that comes back in a different area to the original cancer after initial treatment. Also referred to as distant metastasis.

## Endocrine therapy

Hormones such as oestrogen and progesterone can fuel the growth of some breast cancers. Hormone therapies, such as tamoxifen and aromatase inhibitors, aim to block the availability of hormones such as oestrogen and progesterone and prevent the cancer growing.

## HER2

HER2 is an oncogene which encodes for a cell-surface receptor. Cancer cells may have additional copies of HER2 which leads to an increased number of HER2 receptors and growth of the cancer. HER2 status is assessed using either immunohistochemistry or tests which detect HER2 gene amplification.

HER2 positive cancers may respond to treatment with trastuzumab, a biological treatment which targets HER2 receptors.

## Ki67

Ki67 is a protein associated with cell proliferation.

## Lymph node

A small structure that contains white blood cells, and acts as a filter for foreign particles like cancer cells. There are several lymph nodes in the armpit or near the breastbone where breast cancer cells may be found in people with early breast cancer.

## Micrometastatic disease

Occurs when very small metastatic tumours have formed that are too small to detect on a scan.

## Nottingham Prognostic Index (NPI)

The Nottingham Prognostic Index (NPI) is a validated equation which predicts 5-year survival for operable primary breast cancer. The NPI incorporates tumour grade, size and number of positive nodes.

## Oestrogen receptor

Cancer cells that have oestrogen receptors can use the hormone oestrogen to grow. See also endocrine therapy.

## PREDICT

The PREDICT calculator is an online prognostic and treatment benefit tool that presents 5-, 10- and 15-year survival estimates following surgery both with and without adjuvant therapy. It uses information on age, tumour size, tumour grade, number of positive nodes, menopausal status, ER status, HER2 status, Ki67 status and mode of detection (screening or symptomatic).

## South-west quadrant ICER

A south-west quadrant ICER refers to an ICER in which both costs and QALYs for the intervention are lower than for current practice, that is, the intervention is less clinically effective and cheaper than current practice. Interpretation of south-west quadrant ICERs is reversed from usual considerations (the higher the ICER, the more cost-effective).

# Patient Group Submission Template for Diagnostic Technologies

**NICE**

**Health Technology Assessment (HTA) on** Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer

**Please read the accompanying guide fully before completing this submission template.**

| Information about your organisation | |
|---|---|
| **Organisation name** | Breast Cancer Now |
| **Contact person's name** | [redacted] |
| **Role or job title** | [redacted] |
| **Email** | [redacted] |
| **Telephone** | [redacted] |
| **Postal address** | Breast Cancer Now<br>[redacted] |
| **Organisation type** | Patient/carer organisation (e.g. a registered charity) ☒<br>Informal self-help group ☐<br>Unincorporated organisation ☐<br>Other, please state: |
| **Organisation purpose** (tick all that apply) | Advocacy ☐<br>Education ☒<br>Campaigning ☒<br>Service provider ☒<br>Research ☒<br>Other, please specify: |
| **What is the membership of your organisation (number and type of members, region that your group represents, demographics, etc)?** | |

# Patient Group Submission Template for Diagnostic Technologies

Breast Cancer Now is *a UK charity providing world-class research and life-changing care for people affected by breast cancer*.

## Declarations

**Do you have any conflicts of interest?**

Breast Cancer Now hosts the UK Interdisciplinary Breast Cancer Symposium (UKIBCS) alongside a number of partners including professional bodies and charities. The meeting is held every 2 years and the UKIBCS provides a space to bring together those with an interest in breast cancer research and treatment to advance understanding of the disease. The event is managed by a third party who receive and process sponsorship on behalf of the host and partners. Sponsors have no control over the running of the event and editorial control has been retained by the UKIBCS executive board.

For the 2024 symposium, the UKIBCS has received sponsorship from:

- Exact Sciences - £30,000
- Veracyte - £12,500

In June 2023, Breast Cancer Now received a grant of £15,194 from Exact Sciences towards our helpline.

| | |
|---|---|
| **Did anyone outside your organisation help you prepare this submission?** | Yes ☐ No ☒ |
| **If yes – who helped you and in what way? Please tell us if the people helping you were paid and if they have any conflicts of interest.**<br>**N/A** | |
| **Are you willing for this submission to be shared on our website?** | Yes ☒ No ☐ |
| **We may invite you to a scoping meeting where this technology is to be discussed. Would a member of your group be willing to join such a meeting (this may be in person or virtually)?** | Yes ☒ No ☐ |

# Patient Group Submission Template for Diagnostic Technologies

## Impact of the symptoms, condition or disease on patients

**1. How do symptoms and/or the condition or disease affect patients' lives or experiences?**

A diagnosis of breast cancer can cause considerable anxiety to the patient as well as their family and friends. The initial diagnosis can be extremely shocking and impact on people's emotional wellbeing, whilst in the longer-term, the fear of breast cancer returning or spreading to other parts of the body (such as the bone, liver, lung and brain) which is known as secondary (or metastatic) breast cancer and is incurable can be extremely frightening and distressing for patients.

Breast cancer patients tell us about the impact of the disease on their day-to-day lives, for example the side effects of treatments and visits to hospital taking a significant toll on their general wellbeing, everyday activities, ability to work and their relationships.

Treatment for primary breast cancer is usually a combination of surgery, radiotherapy and chemotherapy. Up to 80% of breast cancers are oestrogen receptor positive (ER+). This would be around 44,000 cases in the UK each year. Patients with hormone-receptor positive breast cancer will receive endocrine (hormone therapy). Research suggests that patients' experience of hormone therapy side effects can affect adherence and in some cases result in treatment discontinuation. Some patients will also have adjuvant chemotherapy which can be associated with additional side effects, such as nausea, neutropenia and hair loss.

This NICE diagnostics assessment is looking at a particular subset of patients with breast cancer which is lymph node positive.

A patient with experience primary breast cancer which was lymph node positive explains:

"Having cancer had a big impact on me, the diagnosis was a shock and as it had spread to my lymph nodes my biggest fear was not surviving."

She goes on to explain the impact of the treatment:

"Chemotherapy was difficult, particularly towards the end. I lost my hair, had temporary neuropathy and developed permanent tinnitus. I had constant nausea for a week following each EC chemotherapy and ended up in A&E three times, staying in overnight, due to pneumonitis, severe anaemia and

elevated troponin. I had to take a lot of time off work, prior to this I was healthy and had never been a hospital inpatient."

The person goes onto the explain the longer-term impact:

"One year post chemotherapy I have low lymphocytes and I have lesions in multiple locations, originally thought to be metastases, which caused a lot of distress. The lesions are now thought to be an autoimmune sarcoid like reaction, possibly to the chemotherapy. I have had biopsies that were inconclusive and seven months after they were identified on a scan I'm still waiting to see a specialist."

Breast cancer can impact every part of an individual's life. A person explains their fears around the impact of treatment on fertility:

"My fertility options were not fully explained to me, the conversation was probably missed because I moved to a new area shortly after diagnosis. I therefore didn't have any fertility preservation other than Zoladex during chemotherapy to try to protect my ovaries. One year post chemotherapy my hormone levels are still in the menopausal range and menstruation has not returned. I am awaiting a fertility referral but I'm worried that the chemotherapy may have caused early menopause."

## Impact of the symptoms, condition or disease on family and carers

### 2. How do symptoms and/or the condition or disease affect carers/unpaid care-givers and family?

A person's diagnosis of breast cancer can also have a significant impact on their family and friends, as they are likely to experience feelings of worry and uncertainty about what the diagnosis of breast cancer means for their loved one. They may require information, advice and support themselves. Supporting a friend or loved one with breast cancer can be upsetting and be emotionally difficult, whilst breast cancer can also impact relationships. For example, couples facing cancer can feel emotional distress.

With the treatment associated with breast cancer, care-givers and loved ones may also need to take on additional tasks and responsibilities, for example, taking people to appointments and household and caring responsibilities.

A person with experience of primary breast cancer which was lymph node positive explains the impact on those around her:

"My partner took time off work to take me to appointments and my mother helped with housework."

# Patient Group Submission Template for Diagnostic Technologies

## Experiences and availability of current diagnostic technologies

**3. What role do currently available diagnostic technologies play in helping patients manage their symptoms and/or the condition or disease?**

Currently a number of tumour profiling tests are available on the NHS for people with ER positive, HER2 negative and lymph node negative (including micrometastatic disease) primary breast cancer. The availability of these tests can be described as having transformed clinical practice for the eligible group of patients. It has been a real step towards personalising individual treatment and being able to safely de-escalate treatment where appropriate. Chemotherapy and the gruelling side effects that can be associated with treatment can be particularly burdensome for patients and a frightening prospect. The availability of these tests have been a welcome step forward in tailoring treatment pathways. For those patients for whom they could be spared chemotherapy, it means they have been able to avoid the potential short and long-term effects of chemotherapy, whether that is the impact on fertility, infection, hair loss or fatigue. It can also potentially mean less time requiring hospital appointments and avoiding the impact that may have on work and/or caring responsibilities.

An individual with experience of a test explains "my theory is that if it gave me the potential chance to avoid harsh and horrible chemotherapy treatment, it was worth having it". The individual also recognised the anxious wait for the result "it added another few weeks to work out what the next steps for my treatment would be, which can be frustrating". Another person explains "the waiting for the results can be the worst bit".

Alongside a number of tests being recommended by NICE and funded by NHS England in the lymph node positive group, it is estimated that about 47% of breast units have access to a tumour profiling test in the 1-3 node positive patients with breast cancer, although some patients will have access via clinical trials such as OPTIMA. [1] This has enabled even more patients to be able to safely avoid chemotherapy and the potentially gruelling side effects that can be associated with it.

Given the significant inequity in access to these tests in the lymph node positive group across England, we are pleased that NICE is now reviewing the evidence in order to make a recommendation on their use for a wider group of patients. We hope national guidance will result in a greater standardisation of the pathway for this group of patients and reduce the variability of care that currently exists across the country. Approval of the tests in this group of patients could save a significant number of patients the hardships of chemotherapy and the extra hospital admissions which can be associated with chemotherapy related toxicity. There is an opportunity to improve the quality of life for this group of patients.

The tests can also be reassuring and provide patients with confidence regarding how their condition is being managed.

[1] Data from a poll of members of the UK Breast Cancer Group in November 2022, a forum of medical and clinical oncologists

# Patient Group Submission Template for Diagnostic Technologies

**4. What unmet information needs do people currently have due to the lack of an available diagnostic technology for their symptom or condition?**

For the group of potentially eligible patients with lymph node positive disease, the availability of further information to guide treatment decisions would be hugely welcome - it could provide important information on the risk of their breast cancer returning and the possible benefits of chemotherapy. This provides an opportunity to provide a more personalised treatment approach - as it imperative to identify people who have a high risk of recurrence and are more likely to benefit from chemotherapy, as well as identifying those who are more low risk and where chemotherapy can be safely omitted. Whether or not chemotherapy is needed can often be one of the top concerns for patients and having more information to inform discussions and provide reassurance would be an important step froward.

# Patient Group Submission Template for Diagnostic Technologies

## About the diagnostic technology being assessed

**5. What are the most important things people would like to gain from the information provided by, and/or the use of, the diagnostic technology being assessed?**

- Reassurance and confidence about their treatment plan - whether it is the need to go ahead with chemotherapy or being able to safely avoid having chemotherapy. It is important that patients have a clear understanding of the role of the test within wider decision-making.

- Timely turnaround of test results as many people explain that the waiting can be the worst component. There should be clear timelines set out for the results so patients know what to expect to help them be as prepared as possible during the wait.

**6. For those people <u>with</u> experience of this diagnostic technology, what difference did the information provided by, and/or the use of, the technology make in their lives or the lives of family and carers?**

As previously outlined, a key difference tumour profiling tests being available in the population being assessed is that it could crucially identify more people who may be able to safely avoid chemotherapy and the difficult side effects it can be associated with. Importantly it may also identify and/or confirm patients who would benefit from chemotherapy. For both of these groups, the tests can provide confidence and reassurance about treatment decisions which can be significant for patients at a time where they may be facing uncertainty and anxiousness about their diagnosis.

However, it is important to recognise that patients can feel nervous and uncertain whilst awaiting the results of the tumour profiling tests and understanding what it means. A patient with experience of the test explains "I found it very hard to get my head around so much information and sometimes you think chemotherapy is the answer. I wanted to understand about the clinical factors which was suggesting that I didn't need chemotherapy". Another patient with experience of the test who went on to have chemotherapy outlined that "the score helped me get through chemotherapy, knowing it could make a difference".

People with experience of the test who safely avoided chemotherapy can still feel nervous and anxious about the decision, so it is important they are provided with information and support. People explain "we can make decisions on the information we have been given - unfortunately that doesn't mean we are not going to worry over whether we have made the right decision or not".

# Patient Group Submission Template for Diagnostic Technologies

**7. For those without experience of this diagnostic technology, but who are aware of studies or other sources of evidence of value, what are the expectations/limitations of having the information provided by the diagnostic technology and/or using the diagnostic technology?**

For those without direct experience of tumour profiling tests but for whom have some awareness of the tests, they explained having further information to discuss as part of their treatment decision making process would have been beneficially and provided some additional reassurance. A patient with lymph node positive disease who did not have access to a test explains:

"I feel having the extra information would have been beneficial. My way of coping was to learn as much as possible about my cancer and the more I learned the more empowered and less afraid I felt. I don't know what my chances of recurrence are, this makes it difficult to move on and make decisions such as choosing which hormone therapy to take and whether to pause it to try to conceive."

The individual goes on to explain:

"Even if the statistics are not favourable, I still find it reassuring to know the facts and if I don't know I imagine the worst, knowing my recurrence risk would reassure me. I have spent a lot of time looking up studies and survival statistics for my type of cancer, and the information would be more accurate if I knew more about the tumour profile. If tumour profiling had taken place and shown that chemotherapy was not required in my case, the treatment would have had much less of an effect on my physical and emotional wellbeing."

## Additional information

**8. Please include any additional information you believe would be helpful in assessing the value of the diagnostic technology (e.g. equality issues, ethical or social issues and/or socio-economic considerations).**

There is currently a significant equality issue regarding access to the tests in the eligible lymph node positive group as outlined earlier in this submission.

Furthermore, at a time when healthcare professionals are telling us that chemotherapy units have limited capacity which is undermining the ability to deliver systemic anti-cancer treatment in a timely manner, increasing the number of patients who could safely avoid chemotherapy would free up space for those that would benefit from chemotherapy and other treatments and help ensure they could receive it in a timely manner.

# Patient Group Submission Template for Diagnostic Technologies

## Key messages

**9. In up to five statements please list the most important points of your submission.**

- The use of tumour profiling tests in the lymph node positive group could help inform decisions about the use of adjuvant chemotherapy and help categorise patients' risk of recurrence and identify those who are most likely to benefit from chemotherapy.
- Chemotherapy can be particularly frightening for patients given it can be associated with both short and long-term side effects and patients fear how it may impact their day-to-day lives. Therefore, identifying those patients who may be able to safely avoid chemotherapy, as well as those may benefit from chemotherapy, could enable a more personalised treatment plan for a wider group of patients which would be a welcome step forward. It can reassure and provide confidence to a patient about their treatment.
- It is important to recognise that having the test does not remove all fear and anxiety for patients as they wait to find out the results of the tests. It is important that there is timely turnaround of the test results and that patients understand how long they may need to wait for the results and what to expect.
- It can importantly equip patients with invaluable information to help them and their clinician make a decision about the best treatment for them, whilst also helping to free up capacity within overstretched breast cancer services.

**Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence. Diagnostics Assessment Report**

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

## Addendum 1: Probabilistic sensitivity analysis of non-predictive scenarios for Oncotype DX in the post-menopausal subgroup

| | |
|---|---|
| **Produced by** | Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health, University of Sheffield |
| **Authors** | Paul Tappenden, Professor of Health Economic Modelling, SCHARR, University of Sheffield, Sheffield, UK |
| **Correspondence author** | Paul Tappenden, Professor of Health Economic Modelling, SCHARR, University of Sheffield, Sheffield, UK |
| **Date completed** | 7th October 2023 |

**Figure 1: BC2, DSA11, HR=0.60, non-predictive, cost-effectiveness plane**


**Figure 2: BC2, DSA11, HR=0.60, non-predictive, cost-effectiveness acceptability curves**

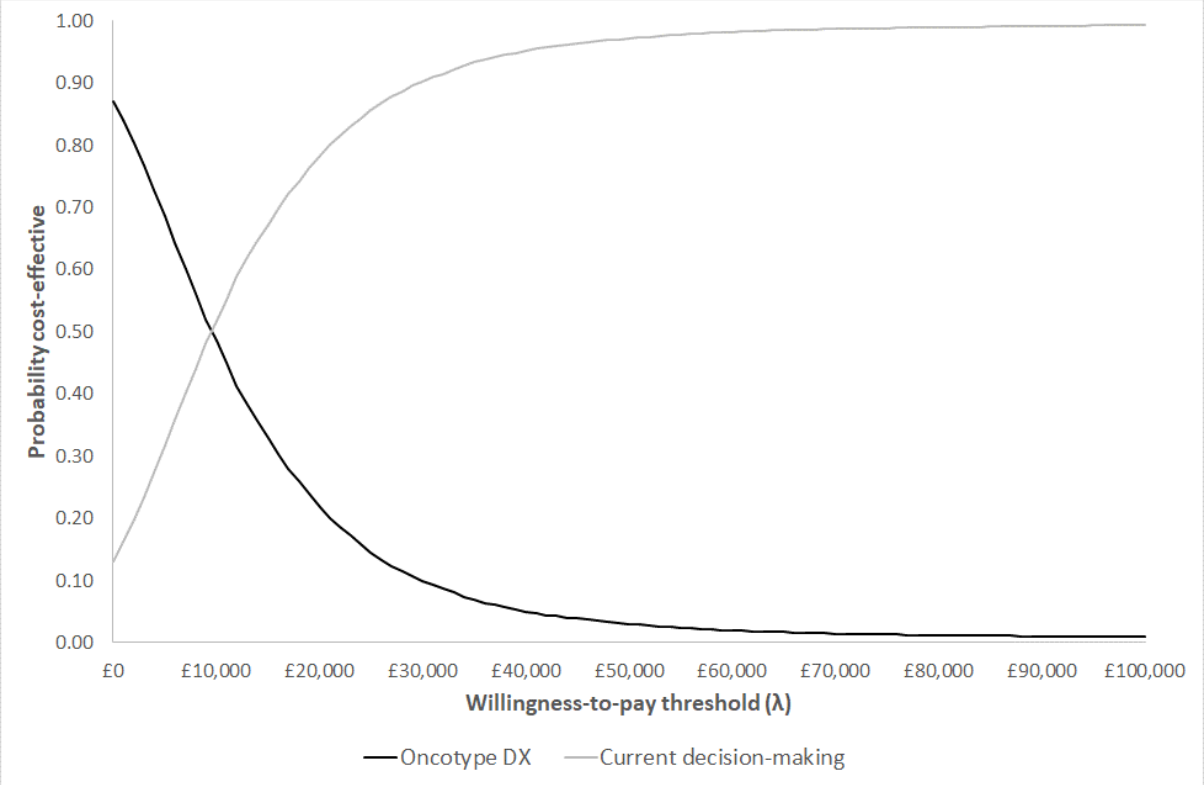


**Figure 3: BC2, DSA12, HR=0.71, non-predictive, cost-effectiveness plane**

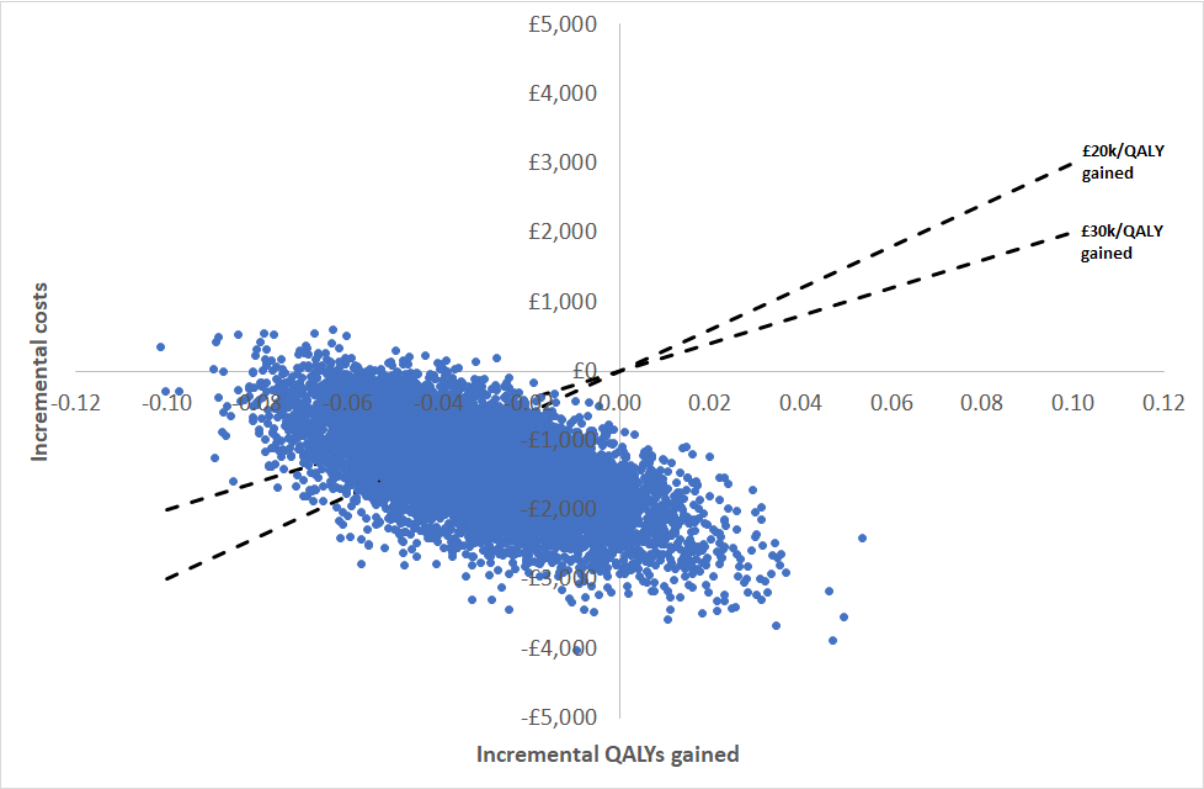



**Figure 4: BC2, DSA12, HR=0.71, non-predictive, cost-effectiveness acceptability curves**

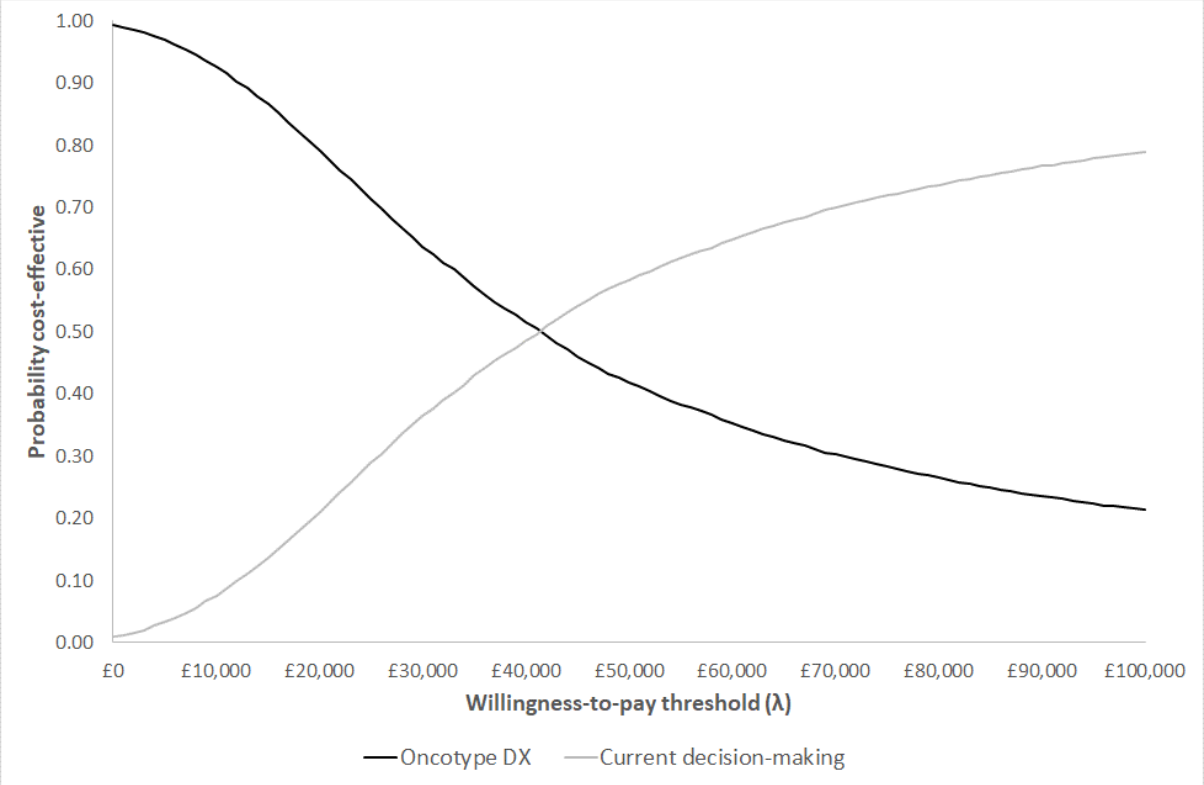


**Figure 5: BC2, DSA13, HR=0.80, non-predictive, cost-effectiveness plane**

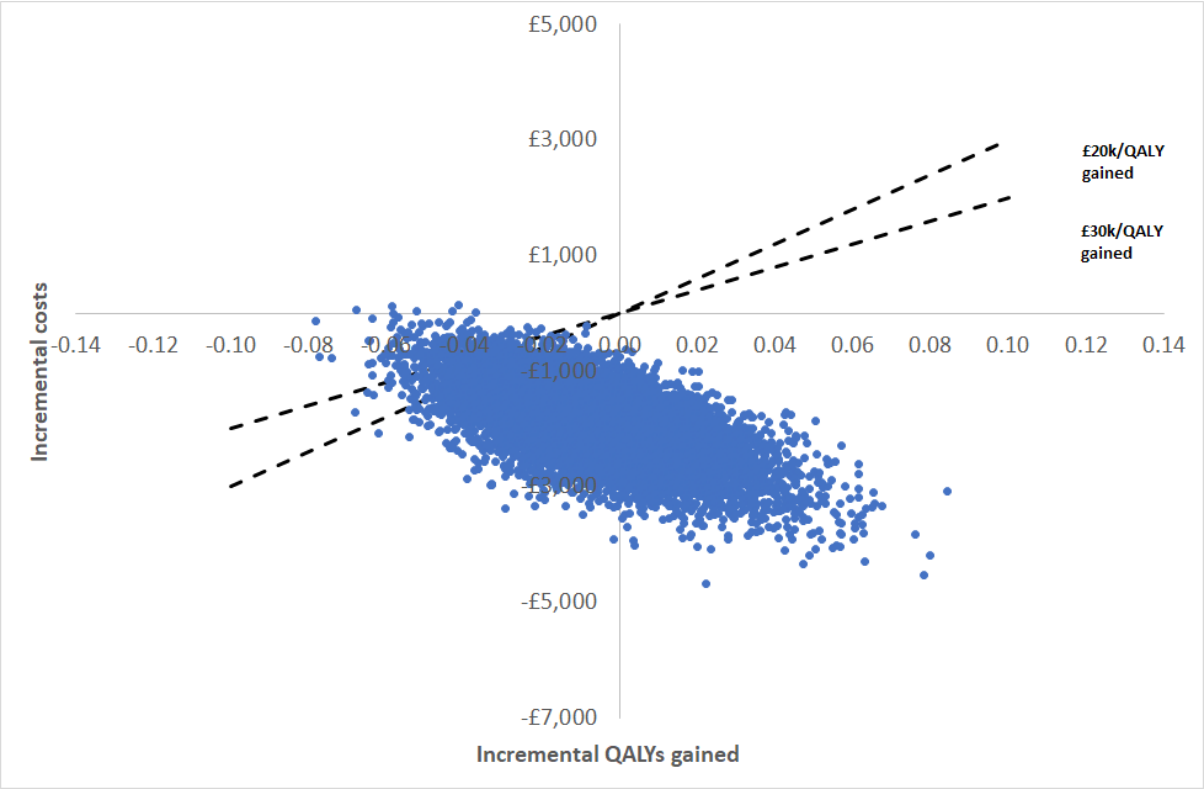
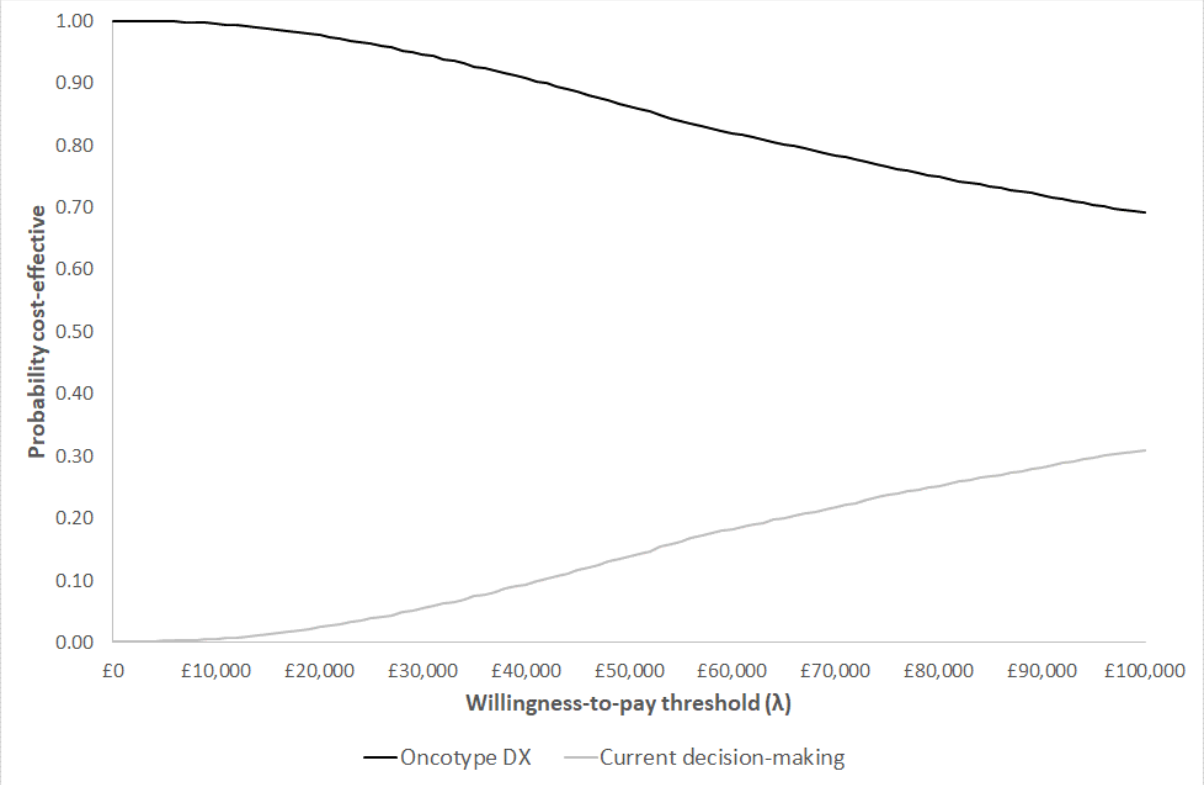



**Figure 6: BC2, DSA13, HR=0.80, non-predictive, cost-effectiveness acceptability curves**



**Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence. Diagnostics Assessment Report**

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

## Addendum 2: Additional economic analysis of Prosigna versus current decision-making using updated list price of £1,488


| | |
|---|---|
| **Produced by** | Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health, University of Sheffield |
| **Authors** | Paul Tappenden, Professor of Health Economic Modelling, SCHARR, University of Sheffield, Sheffield, UK |
| **Correspondence author** | Paul Tappenden, Professor of Health Economic Modelling, SCHARR, University of Sheffield, Sheffield, UK |
| **Date completed** | 18th October 2023 |

**Table 1: Central estimates of cost-effectiveness, Prosigna versus current decision-making (BC5), including updated Prosigna list price of £1,488**

| Option | LYGs* | QALYs | Costs | Inc. LYGs* | Inc. QALYs | Inc. costs | Incremental cost per QALY gained |
|---|---|---|---|---|---|---|---|
| **Probabilistic** | | | | | | | |
| Prosigna | 19.73 | 10.28 | £47,019 | 0.06 | 0.03 | £676 | **£24,547** |
| Current DM | 19.67 | 10.25 | £46,342 | - | - | - | - |
| **Deterministic** | | | | | | | |
| Prosigna | 19.71 | 10.32 | £47,242 | 0.06 | 0.03 | £700 | **£25,403** |
| Current DM | 19.65 | 10.30 | £46,543 | - | - | - | - |

** Undiscounted*

*LYG - life year gained; QALY - quality-adjusted life year; Inc. - incremental; DM - decision-making*

**Table 2: Deterministic sensitivity analysis results for Prosigna versus current decision-making (BC5), including updated Prosigna list price of £1,488**

| DSA | BC5 – Prosigna, TransATAC post-menopausal, non-predictive |
|---|---|
| **Deterministic base case ICER** | **£25,403** |
| DSA1: 17% of women assumed to be in RS >25 group (Oncotype DX only) | N/a |
| DSA2: Prosigna test classification probabilities and DRFI from Gnant *et al.*,[54] | £14,209 |
| DSA3: EPclin test classification probabilities and DRFI from Filipits *et al.*[161] | N/a |
| DSA4: 3-level post-test chemotherapy probabilities - Llombart Cussac *et al*[43] | £23,843 |
| DSA5: 3-level post-test chemotherapy probabilities - Loncaster *et al.*[36] | Dominated |
| DSA6: 3-level post-test chemotherapy probabilities - Zambelli *et al.*[44] | £36,233 |
| DSA7: 2-level post-test chemotherapy probabilities – Dieci *et al.*[41] | N/a |
| DSA8: 3-level post-test chemotherapy probabilities - UKBCG survey (3-level tests)[10] | £26,931 |
| DSA9: 2-level post-test chemotherapy probabilities - UKBCG survey (2-level tests)[10] | N/a |
| DSA10: Risk tapering to 50% at 10 years then 0% at 15 years | £25,757 |
| DSA11: CET vs ET HR = 0.60 in all genomic risk groups (non-predictive) | £13,882 |
| DSA12: CET vs ET HR = 0.71 in all genomic risk groups (non-predictive) | £25,403 |
| DSA13: CET vs ET HR = 0.80 in all genomic risk groups (non-predictive) | £40,152 |
| DSA14: Chemotherapy QALY loss halved | £28,061 |
| DSA15: Chemotherapy QALY loss doubled | £31,339 |
| DSA16: Chemotherapy QALY loss tripled | £33,986 |
| DSA17: Baseline probability of chemotherapy = 0.90 | Dominated |
| DSA18: Start age + 5 years | £33,400 |
| DSA19: Start age – 5 years | £21,195 |
| DSA20: Utility values from Verrill *et al.*[160] | £28,039 |
| DSA21: AML removed from model | £30,980 |
| DSA22: Chemotherapy cost halved | £31,559 |
| DSA23: Chemotherapy cost doubled | £13,092 |
| DSA24: DM lifetime cost halved | £31,458 |
| DSA25: DM lifetime cost doubled | £13,294 |
| DSA26: AML costs halved | £26,358 |
| DSA27: AML costs doubled | £23,494 |

*BC - base case; ICER - incremental cost-effectiveness ratio; DSA - deterministic sensitivity analysis; N/a – not applicable; RS - recurrence score; DRFI - distant recurrence-free interval; UKBCG - UK Breast Cancer Group; CET - chemotherapy plus*

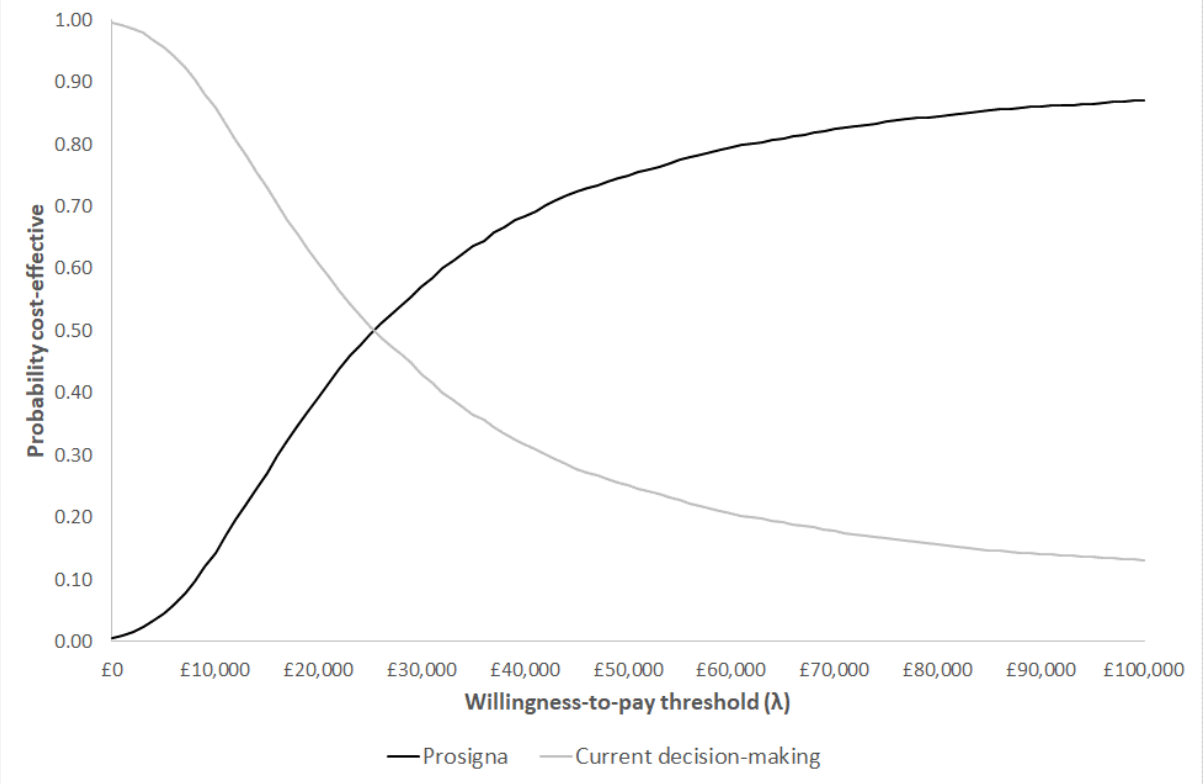
*endocrine therapy; ET - endocrine therapy; HR - hazard ratio; QALY - quality-adjusted life year; AML - acute myeloid leukaemia; DM - distant metastases; SWQ - South-West quadrant*

**Table 3: Model-predicted incremental clinical and economic outcomes per 1,000 women tested – Prosigna versus current decision-making, including updated Prosigna list price of £1,488**

| Incremental model outcome (test versus current decision-making) | BC5 – Prosigna, TransATAC post-menopausal, non-predictive |
|---|---|
| Number of women receiving chemotherapy | -46 |
| Number of infusion chair hours | -235 |
| Number of women experiencing DM during their lifetime | -3 |
| LYGs (undiscounted) | 59 |
| QALYs gained (discounted) | 28 |
| Additional costs to NHS/PSS (discounted) | £699,509 |
| Net health benefit (£20,000 per QALY gained) | -7 |
| Net health benefit (£30,000 per QALY gained) | 4 |

*BC - base case; DM - distant metastases; LYG - life year gained; QALY - quality-adjusted life year; NHS - National Health Service; PSS - Personal Social Services*

**Figure 1: CEACs, BC5 – Prosigna, TransATAC, post-menopausal (non-predictive benefit), including updated list price of £1,488**



# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

## Section A: External Assessment Report - Comments

**Note from the EAG** – comments in which amendments will be made to the EAG report, or where addenda have been provided, are highlighted in **bold red text**.

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| Veracyte Inc. | 1 | 18 | 1.7 & 5.6 | We suggest that the research priorities make note of the ongoing OPTIMA trial, sponsored by the National Institute for Health Research (NIHR). This prospective, randomized non-inferiority trial is specifically assessing GEP guided (Prosigna) therapy versus standard chemo-endocrine therapy in a high-risk HR+, predominantly LN+ ESBC population. As noted in the first bullet point, the prediction of chemotherapy benefit by GEPs is still uncertain. OPTIMA will provide additional data to help answer this question. The study design includes a relatively high cut point (ROR> 60) for defining the high-risk cohort and allows inclusion of patients with N2 disease. This design should allow assessment of GEP score interaction with chemotherapy benefit. As well, completion of trial accrual would allow for release of OPTIMA-preliminary trial outcomes, providing the only trial-based prospective comparative data for Oncotype DX, MammaPrint, Prosigna, and EndoPredict. EAR endorsement of the trial would be helpful for completing trial accrual and seems appropriate for these sections. | Section 5.6 of the EAG report (Suggested research priorities) relates to studies which are not currently ongoing. The ongoing OPTIMA trial is mentioned elsewhere in the EAG report (chapters 1, 3 and 5). We agree that the results of the OPTIMA trial may be useful. However, the EAG does not believe that the report needs amending. |
| Veracyte Inc. | 2 | 33-35 | 2.3.2 | It is noted that Oncotype DX is a Conformité Européene (CE) marked assay. In the following sentence it is noted that “The company claims that the test can also predict the likelihood of chemotheraphy benefit”. If the assay is CE-marked there is also a defined intended use statement in the package insert (see attached Prosigna package insert). We agree with the Technology assessment report authors that the assay is the most important element to have a CE mark for.<br><br>Veracyte is not aware that the assay for e.g. Oncotype DX is CE marked. In recent published reports (hiips://www.tlv.se/download/18.7102c4617a75ed7acf77376/1630506397339/bed210602_Oncotype_dx.pdf ) by other HTA agencies the Swedish TLV it is noted in section 3.2.1 that the company has informed TLV that the Oncotype DX is not a CE marked assay but have a self reported CE mark for the sample collection kit and the software. Veracyte feels that it is very important for safety and decision making that patients, relatives and health care professionals are fully informed about the regulatory status of the different tests and particularly the most important element the assay and thereby regulatory approved claims for intended use. The Prosigna® Breast Cancer Prognostic Gene Signature Assay is CE marked. The Assay consists of two separate IVD medical devices: 1. Prosigna® Breast Cancer Prognostic Gene Signature Assay Kit 2. Prosigna® Breast Cancer Prognostic Gene Signature Assay Software Module Both devices are CE | The request for information provided by Exact Sciences states that Oncotype DX is CE marked. Exact Sciences will be able to provide further information on this point.<br><br>Section 2.3.2 of the EAG already clearly states where the companies have made claims of predictive benefit. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | marked and currently designated as Class C “legacy devices” in compliance with the IVDR (EU 2017/746). They were self-declared under the IVD Directive (see attached Declarations of Conformity) and are currently available under the EU IVDR transition legislation (EU 2022/112). It is planned to submit these devices for Notified Body review to obtain IVDR CE certification prior to the end of the transition period (May 25, 2026) to ensure long-term compliance and availability within the EU. Outside the EU, the Prosigna Assay is available as a fully registered IVD medical device in several countries including 3 MDSAP countries - Australia (TGA, Class 3), Canada (Health Canada, Class 3) and the USA (FDA, Class II, 510(k) cleared). Further it should be noted that Prosigna® is UK GDPR compliant as no patient sensitive information or biological material are shipped outside the country and thereby do not leave the closed loop system and safeguarding of the NHS. Throughout the EAR there is referral to the intended use statement for the Prosigna assay e.g. Table 4 noting that e.g. premenopausal patients are not included for Prosigna (as derived from the CE marking íntended use statement). However for Oncotype DX (p. 33), Endopredict (p. 34) and Mammaprint (p. 35) there is statements “that the company claims” without a single reference to an intended use statement from the package insert as part of the CE marking on an assay and any restrictions in claims such a document would come with. Veracyte therefore ask that it made clear in the text and Table 4 which products are actually CE marked assays and which claims are included in the intended use statement and not just claims from a company. Veracyte also request that it is noted in Table 4 whether the assays are UK GDPR compliant as is the case for Prosigna. | The EAG does not have information on whether EPclin, Oncotype DX and MammaPrint assays are GDPR compliant. NICE may be able to provide further information. |
| Veracyte Inc. | 3 | 33-35 | 2.3.2<br><br>2 v. 3 level cut point | Veracyte ask that it is noted that risk is indeed a continuous variable. The development of risk cut points for gene expression profiling is one of the more challenging aspects of these assays. One of the fundamental questions is how many cut points should be implemented in the development of the test. Underlying this challenge is the clear observation that prognostic estimates of distant recurrence have routinely and consistently been shown to be continuous. No assay, either conventional clinical pathologic assessments immunohistochemistry, or gene expression profiling has been shown to have clear dichotomous cut points wherein patients below the cut point have no recurrences and patients above the cut point all have recurrences. The complexity of the biology suggests that such a cut point is not possible. For gene expression profiling assays, two or three levels of risk have been developed for most assays. A dichotomous low/high cut point structure is inherently simpler and advocates for this approach argue that it makes clinical decision making easier. However, this distillation of risk to a dichotomous structure inevitably oversimplifies assessment of the underlying risk. If the gene expression profiling test alone were definitive for guiding selection of adjuvant therapy, this approach would be more tractable. However, gene expression profiling is additive to overall clinical pathologic assessment. Multi-level risk assessments can provide a more nuanced assessment of risk, especially in the context of also supplying a continuous risk estimate curve. | This is useful information. However, no response or amendment of the EAG report is required. |
| Veracyte Inc. | 4 | 175 | Appendix 3: Domain | In Section 3.5.8, the EAR notes the analyses from NCDB. However, we wanted to highlight the aspects of the RxPONDER design that likely contributed to biased enrolment and the evidence of this bias in comparison to NCDB data. In the RxPONDER design, bias may have been introduced in the modified | The EAG agrees that it is possible that the RxPONDER trial |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | 3 (Outcomes) | PROBAST Domain 3 Outcomes Criterion, "Was chemotherapy decision made before test result known?". The study registration and randomization design show that patients and clinicians were aware of Oncotype results in many cases prior to making the final decision to participate and be randomized. A large proportion of potentially otherwise eligible subjects were not randomized (25%; Kalinsky, NEJM, 2022). An analysis using National Cancer Database (NCDB, American College of Surgeons) (Cao Ref. 80; published in Stabellini, Frontiers in Oncology, doi: 10.3389/fonc.2023.1115208, 2023) suggests conflicting results for the benefit of chemotherapy for women with recurrence scores under 26. Moreover, the distribution of Recurrence Scores in this population-based study suggests that patients with RS 20-25 may be under-represented in the RxPONDER cohort. While not prospective and randomized, the data do come from a large, population-based registry and may be more representative of the patients in question. The results are also more consistent with the findings in the EBCTCG meta-analyses. Taken together, these results suggest that RxPONDER was biased toward a lower risk population and for this reason did not show prediction of chemotherapy benefit. In section 5.6 on suggested research priorities, it is noted in the first bullet that such studies may not be considered ethical. We want to make the EAR aware that the ongoing OPTIMA study will provide further data on GEP chemotherapy benefit prediction without such ethical issues. | eligibility criteria influenced the make-up of the randomised trial population. However, the EAG considers that the PROBAST criterion refers to whether the test result influenced the chemotherapy decision, and that this was not the case in RxPONDER as chemotherapy use was randomised.<br><br>In terms of representation of RS ranges in RxPONDER, the predictive benefit broken down by RS range is presented in the EAG report Section 3.5 Table 10.<br><br>The ongoing OPTIMA study is described in the EAG report. |
| Veracyte Inc. | 5 | General | | Veracyte: Supplementary information to the Request for Information process, shared with NICE project team on 11 October 2023.<br><br>List price for Prosigna (incl request to EAR and the NICE project team to update the base case Prosigna list price): In the Request for Information (RFI) process only the list price for the 1 Kit option was included (see Table 1). Veracyte wants to highlight that the Prosigna list price used in the model is the most expensive and least used option namely the 1 kit version. Prosigna is delivered an used in 1, 2, 3, 4 and 10 kit versions with an average list price of 1.296 GBP per test (simple average price across the 5 kit configurations). 4 and 10 test kits are mostly used in the UK (83 % of total sales). Including other expenses to run the Prosigna assay (nCounter DX Flex, nCounter servicing, High pure RNA isolation kit and total laboratory staff costs the list price is therefore 1.488 GBP per test (see Table 2) since the other costs sums | **The EAG has re-run the analysis using the updated list price. The updated results have been presented in a separate addendum to the NICE report. This analysis will also be included in the HTA monograph.** |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>to 192 GBP. Veracyte therefore ask the EAR and NICE project team to use this Prosigna price in the health economic modelling as the base case. Also to update e.g. Table 37 and Table 38 in the EAR report on the Prosigna list price so the total for Prosigna is 1.488 GBP instead of 1.896 GBP. Further Prosigna is already offered in the UK for the node negative patients at a further reduced disccounted price and this confidential price will also be offered for the node positive patients as has been confirmed to NICE<br><br>Table 1. Prosigna list price across kit configurations.
<table>
<tr><th>Prosigna Kit</th><th>List price per test (GBP) (assay only VAT excluded)</th><th>Percentage of total sales in England</th></tr>
<tr><td>Prosigna 1 test kit</td><td>1.580</td><td>>1 %</td></tr>
<tr><td>Prosigna 2 test kit</td><td>1.344</td><td>9 %</td></tr>
<tr><td>Prosigna 3 test kit</td><td>1.264</td><td>8 %</td></tr>
<tr><td>Prosigna 4 test kit</td><td>1.187</td><td>68 %</td></tr>
<tr><td>Prosigna 10 test kit</td><td>1.106</td><td>15 %</td></tr>
<tr><td>Prosigna simple average price per test across kits configurations</td><td>1.296</td><td>~ 100 %</td></tr>
</table>
Table 2. Total costs of running a Prosigna test in the UK based on average kit configurations</td><td></td></tr>
</table>

## Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
<table>
<tr><th>Resources used for test</th><th>Description</th><th>Mode</th><th>Cost (VAT excluded)</th><th>Daily maximum test</th><th>Total tests performed in life span (7 years)</th><th>Unit cost (VAT excl.)</th></tr>
<tr><td>nCounter DX Analysis - FLEX</td><td>Includes the Prep Station and the Digital Analyzer with the FLEX Configuration</td><td>NCT-SYST-FLEX</td><td>£230,850</td><td>30</td><td>76,650</td><td>£3.01</td></tr>
<tr><td>nCounter Servicing</td><td>Annual service over 7 years</td><td></td><td>£85.050</td><td>30</td><td>76,650</td><td>£1,11</td></tr>
<tr><td>Prosigna Gene Signature Assay – list price</td><td>Complete kit for running Prosigna tests. Includes all CodeSet and Master Kit components; does not include RNA Isolation Kit.</td><td>PROSIGNA (Averaged across kit configurations 1, 2, 3, 4 and 10 Test kits)</td><td>£1,296</td><td></td><td></td><td>£1,296.00</td></tr>
<tr><td>High Pure RNA</td><td>For 25 reactions</td><td>Roche product</td><td>£405</td><td></td><td></td><td>£16.20</td></tr>
</table>
</td><td></td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | **Isolation Kit**<br><br>**Total laboratory staff costs per test** £171.68<br><br>**Unit cost of performing Prosigna test in house in NHS** Unit cost of nCounter DX Analysis + unit cost of Prosigna Gene Signature Assay + unit cost of RNA isolation Kit + lab costs per test **£1,488.00** | |
| Myriad International GmbH | 6 | 13, 48, 49 | 1.4.1 (3.4) | - **Myriad strongly agrees with the EAG's positive assessment of the prognostic performance of EndoPredict® in node-positive patients.**<br><br>Before commenting on the list of studies selected by the EAG, Myriad Genetics can only agree with the authors who have chosen to depart from the initial exclusion criteria of the systematic review to properly assess the prognostic performance of the signatures in node-positive patients: **inclusion of studies conducted on mixed nodal status population with more than 20% of pN0.**<br><br>Indeed, EndoPredict® was developed to predict the risk of recurrence of early-stage invasive breast cancer, RO+, HER2, at 10 years, **regardless of the nodal status.** The clinical validation studies were conducted on mixed nodal status populations, with a consistent proportion of patients with node-positive disease (around one third), which led to a CE-mark in both nodal populations. | The NICE scope is restricted to patients with 1 to 3 positive nodes (LN0 populations were addressed in NICE DG34). In line with this, the EAG report was restricted to studies or subgroup analyses in which at least 80% of patients were LN+. |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
EndoPredict® was clinically validated in several prospective-retrospective studies on independent series of a total of 3,570 patients, 3,015 patients treated with hormone therapy alone for 5 years and 555 treated with chemotherapy followed by hormone therapy. Patients were included in 4 randomized controlled phase III studies and one cohort study:<br><br>
Table 1. Clinical validation of the prognostic performance of Endopredict®.
<table>
<tr><th>Study</th><th>Publications</th><th>Patient characteristics</th></tr>
<tr><td>ABCSG-6 (n=378) and ABCSG-8 (n=1,324) phase III studies</td><td>- Filipits et al. (2011)(1)<br>- Dubsky et al. (2013)(2)<br>- Filipits et al. (2019)(3)</td><td>- Postmenopausal<br>- pN0 and pN+ (536)<br><br>Filipits et al. (2019), pN+: 536 (31.5%), pN1-3 : 453 (26.6%)</td></tr>
<tr><td>TransATAC phase III study (n=928)</td><td>- Buus et al. (2016)(4)<br>- Sestak et al. (2018) (5)<br>- Sestak et al. (2020) (Lobular) (6)25.10.2023 07:39:00</td><td>- Postmenopausal<br>- pN0 and pN+<br>Sestak et al. (2018), pN+: 183 (24%))</td></tr>
<tr><td>GEICAM/9906 phase III study (n=555)</td><td>- Martin et al. (2014) (7)<br>- Martin et al. (2016) (8)</td><td>- All menopausal statuses<br>- All pN+<br>Martin et al. (2016), pN1-3: 344 (64.2%))</td></tr>
<tr><td>Cohort study (n=385)</td><td>- Constantinidou et al. (2022) (9)</td><td>- Premenopausal<br>- pN0 and pN+<br>pN+: 62 (16.1%)</td></tr>
</table>
In green, are highlighted the 5 publications selected by the EAG to determine the prognostic performance of EndoPredict® in node-positive patients. We take not of the EAG's willingness to avoid double-counting by not selecting all publications related to one study (as explained in the DAR comments table and EAG response, in 2018 (DG 34)(10)). But, since all publications have been selected for the GEICAM/9906 phase III study, and two out of three for the TransATAC study, the rationale behind the exclusion of the other publications, Filipits et al. (2011)(1), Dubsky et al. (2013)(2) and Buus et al. (2016)(4), appears rather unclear.
</td><td>
Where there were multiple publications per study, the EAG included the publication with the most recent and comprehensive data. If earlier publications reported additional data meeting the review inclusion criteria then these were also included (as for GEICAM/9906).<br><br>
For ABCSG-6/8, Filipits 2011 and Dubsky 2013 were excluded as they did not report any additional relevant LN+ data beyond that presented in Filipits 2019.<br><br>
For TransATAC, Buus 2016 was excluded as it did not report any additional relevant LN+ data beyond that reported in Sestak 2018 and Sestak 2020.<br><br>
Sestak 2019 (indirect comparison of trial analyses) was excluded because only 35% of patients were LN+ (and only 26% had 1-3 positive nodes) and there
</td></tr>
</table>

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>
Furthermore, the results of these prospective-retrospective clinical validation studies were subsequently confirmed in a retrospective indirect comparative analysis of several randomized controlled phase III trials (Sestak et al. (2019))(11) and four prospective studies (Ettl et al. (2020)(12) / Klein et al. (2022)(13), Lehmann-Che et al. (2021)(14), Penault-Llorca et al. (2021)(15) and Schmitt et al. (2022)(16)).
<br><br>
Table 2. Prospective validation of the clinical performance of Endopredict®.
<table>
<tr><th>Study type</th><th>Publication / poster</th><th>Patient characteristics</th><th>Publication status</th></tr>
<tr><td>German single-centre prospective registry (Munich)</td><td>Ettl et al. (2020): 3-year results(12)<br>Klein et al. (2022)(13) SABCS 2022: 5-year results</td><td>- All menopausal statuses<br>- pN0 and pN+ (88 (23.9%))</td><td>Manuscript ongoing with submission in November 2023</td></tr>
<tr><td>French prospective multicentre registry SiMoSein</td><td>Lehmann-Che et al. (2021)(14)</td><td>- All menopausal statuses<br>- pN0, pNmic and pN+ (1,442 (30.3%))</td><td>Data disclosure ASCO 2024 (Poster)</td></tr>
<tr><td>International academic phase III RCT (France, Belgium, UK) UNIRAD</td><td>Penault-Llorca et al. (2021) (15) SABCS 2021</td><td>- All menoposal statuses<br>- pN+<br>- Very high risk patients</td><td>Publication expected for Q4 2023/Q1 2024</td></tr>
<tr><td>German single-centre prospective registry (Berlin)</td><td>Schmitt et al. (2022)(16) SABCS 2022: 5-year results</td><td>- All menopausal statuses<br>- pN0, pNmic (55 (6.5%) et pN+ (pN1-3 : 252 (29.9%))</td><td>Publication expected for Q1/Q2 2024</td></tr>
</table>
</td><td>were no subgroup analyses in a LN+ population.<br><br>The prospective studies cited here were all identified by the EAG and were excluded due to not reporting relevant data in a population or subgroup that was at least 80% LN+, or other non-relevant population e.g. post-neoadjuvant therapy.<br><br>In terms of unpublished data, no such data in a relevant population were provided to the EAG by the company during the assessment.</td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | For data whose publication is imminent, a bespoke analysis could be sought, or unpublished data be shared under confidence allowing an exhaustive evaluation of the prognostic performance of Endopredict® in node-positive patients. **The use of unpublished data has proved crucial to properly asses tumor profiling in node-negative patients (TransATAC data, 2017(17)) and in node-positive patients (Holt *et al.* (2023)(18)).** | |
| Myriad International GmbH | 7 | 14, 15 | | - Myriad agrees that chemopredictive data have not been collected for a strictly LN+ population but not with the statement that chemopredictive data have not been collected for LN+ patients.<br>- Myriad deplores that the data collected on a mixed nodal status population have not been developed, as it is the case for the prognostic ability (page 48) and the HRQoL/anxiety dimension (page 75).<br>- Myriad disagrees with EAG's conclusion on the predictive benefit of EndoPredict®.<br><br>EP's predictive benefit was demonstrated and validated using a cross comparison between 5 RCTs: TransATAC, ABCSG-6 and 8, GEICAM 9906 and 2003/02 (Sestak *et al.* (2019)(19)). The authors compared the benefit of the EPClin score in the cohort of patients treated with chemotherapy and hormone therapy (ACT-HT) for 5 years in the GEICAM/9906 (n=500) and GEICAM/2003-02 (n=616) trials with that of patients treated with hormone therapy alone for 5 years in the ABCSG-6 (n=378), ABCSG-8 (n=1,324) and TransATAC (n=928) trials.<br><br>Patients treated with hormone therapy alone and presenting a high EPclin score, in this case 5, had a risk of distant recurrence at 10 years of 46.1% (40.2-51.4%) compared with 25.8% (22.0-29.5%) for patients treated with ACT-HT (increase in absolute risk of 20.3%). Patients with a high EPclin score treated with ACT benefited from a statistically significant increase in distant recurrence-free survival at 10 years, compared with those who did not receive ACT. **This benefit in 10-year distant recurrence-free survival was not found in patients with a low EPclin score who received ACT.**<br><br>**The interaction test between the EPclin score (continuous variable) and the treatment received (HT or ACT-HT) was statistically significant (p=0.022), thus attesting to the predictive nature of the EPclin score (benefit expressed in terms of distant recurrence-free survival at 10 years).** | The Sestak (2019) study was excluded because most of the patients included in the analysis had node-negative disease and so this study was out of scope. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Univariate analysis showed that the EPclin score was highly predictive of breast cancer recurrence in women treated with Ht alone and CTA-HT :<br>- 10-year distant recurrence-free survival in women treated with HT alone (HR=2.79 (2.49-3.13), p<0.0001) and in women treated with HT-ACT (HR=2.27 (1.99-2.59), p<0.0001);<br>- Survival without distant recurrence between 0 and 5 years, in women treated with HT alone (HR=2.76 (2.39-3.18)) and in women treated with ACT-HT (HR=2.49 (2.13-2.92));<br>- Survival without distant recurrence between 5 and 10 years, in women treated with HT alone (HR=2.85 (2.37-3.43)) and in women treated with ACT-HT (HR=1.86 (1.46-2.36)).<br>- The EPclin score remained an independent prognostic factor in multivariate analysis, after adjustment for clinical criteria.<br><br>This indirect comparative analysis:<br>- Confirms the independent prognostic character of the EPclin score for the prediction of the risk of distant recurrence at 10 years, between 0 and 5 years and between 5 and 10 years, as well as for the prediction of the risk of breast cancer recurrence, in patients treated with HT alone or with the combination CTA-HT;<br>- Demonstrates a statistically significant benefit of CTA in patients whose tumours have a high EPclin score, and no benefit if the EPclin score is low;<br>- Attests to the predictive nature of the EPclin score, with a statistically significant formal interaction test, whether the benefit is expressed in terms of distant recurrence-free survival at 10 years or of recurrence-free survival (locoregional and distant recurrences).<br><br>The registry conducted in Munich (Ettl *et al*. (2020)(12)) prospectively confirms the results of this retrospective comparative analysis. To assess the potential benefit of adjuvant chemotherapy (ACT) in high EPclin patients, disease-free survival was compared between patients with a high EPclin score who had or had not received the recommended ACT, based on their EPclin score. The 3-year disease-free survival rate was 96.3% (95% CI: 92.2-100) for high EPclin patients treated with adjuvant chemotherapy *versus* 91.5% (95% CI: 82.7-100%) for high EPclin patients not treated with chemotherapy: HR=0.32; 95% CI: 0.10-1.05; p=0.061. ACT treatment of patients with a high EPclin score was associated with a 68% reduction in events (3-year disease-free survival).<br><br>The 5-year results of the German registries (Klein *et al. (202*2)(13), Schmitt *et al.* (2022)(16)) provide further confirmation of the predictive benefit of Endopredict®. A bespoke analysis could be sought, or unpublished data be shared under confidence allowing a relevant evaluation of the predictive performance of Endopredict® in patients with node-positive disease. **The use of unpublished data has proved crucial to** | |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | **properly asses tumor profiling in node-negative patients (TransATAC data, 2017(17)) and in node-positive patients (Holt *et al.* (2023)(18)).** | |
| Myriad International GmbH | 8 | 15 (120)<br>69 | 1.4.1<br>3.6.1 | Myriad understands from methodological standpoint to use the unpublished decision impact study of Oncotype DX (Holt *et al.* (2013)(18), with regard to its strictly N+ target population but deplores that the data collected on a mixed nodal status population have not been developed in this section, as it is the case for the prognostic ability (page 48) and the HRQoL/anxiety dimension (page 75).<br>Decision impact studies have been conducted in Europe and other reference countries for mixed nodal status populations, with consistent proportion of patients with node-positive disease in different countries: (UK 2, Germany 2, Australia 1, France 1, Mexico 1).<br><br>- Decision Impact Study Charité Berlin (Germany), N=167, **38% node positive** (1-3 PLN) Change in treatment recommendation: 37.7%, Net change in chemotherapy: - 13.1%, Müller *et al. (*2013)(20)<br>- Decision Impact Study TU Munich (Germany), N=395, **23% node positive** (1-3 PLN) Change in treatment recommendation: 43%, Net change in chemotherapy: - 33%, Note: For a substantial proportion of these patients also 5-years survival analysis is available from a local registry, Ettl *et al. (*2017)(21)<br>- Decision Impact Study Brighton (UK) N=149, **33% node positive** (1-3 PLN), Change in treatment recommendation: 36.9%, Net change in chemotherapy: +0.7%, Fallowfield *et al. (*2018)(22)<br>- Decision Impact Study London (UK) N=120, **node positive patients included**, Change in treatment recommendation: 31%, Net change in chemotherapy: -16%, In patients with NPI>3.4 who were candidates for chemotherapy (N=79) 35% were classified as low risk by EPclin and were recommended not to receive chemotherapy. Mokbel *et al. (*2018)(23)<br>- Decision Impact Study ADENDOM (France) N=201, **9% pN1mic**, Change in treatment recommendation: 35.8%, Net change in chemotherapy: -20.9%, Penault-Llorca *et al.* (2020)(24)<br>- Decision Impact Study Monterrey (Mexico), N=91, premenopausal, **28% node positive** (1-3 PLN), Change in treatment recommendation: 17%, Net change in chemotherapy: - 10.5%, Villareal-Garza *et al. (*2020)(25)<br>- Decision Impact Study PROSPER (Australia), N=233, **30.5% node positive** (1-3 PLN), Full cohort: Change in treatment recommendation: 23%, Net change in chemotherapy: - 4.3%, Selectively recruited cohort (N=110): Change in treatment recommendation: 39.1%, Net change in chemotherapy: - 11.9%, Dinh *et al. (*2021)(26). | As noted in the EAG's protocol, where studies include patients who are non-early stage or who are otherwise out of scope, and no subgroup data are available, the following rule was applied: if the percentage of patients out of scope is ≤20% then the study will be included (and its contribution to outcome heterogeneity considered), whilst if >20% are out of scope the study will be excluded. In each of the studies identified by Myriad, a minority of patients had node-positive disease; hence, the studies were not eligible for inclusion in the review. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Myriad considers acceptable to include these data, as the figures of the studies not included (see above) confirm the Holt et al. data on Oncotype. | |
| Myriad International GmbH | 9 | 126 | Table 38 | The list price of 1500, - GBP does not correspond to the discounted final price agreed with the NHS in 2018. | The list price of £1,500 was provided by Myriad in the request for information submitted to NICE in June 2023. This is also the list price used in NICE DG34. The confidential price discount for EPclin has been included in a confidential appendix to the report. This appendix has already been provided for NICE. |
| Myriad International GmbH | 10 | 150 | 5.5 | Myriad strongly agrees with final statement that EAG model suggest all the tumour profiling tests are expected to result in fewer woman receiving adjuvant chemotherapy and reducing costs. | This isn't the conclusion drawn by the EAG for all tests. The model suggests that all tests will result in a reduction in the use of adjuvant chemotherapy, but some of the tests will result in a net increase in costs to the NHS and PSS. |
| Myriad International GmbH | 11 | | | **Appendix: Bibliography**<br>1. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res Off J Am Assoc Cancer Res. 15 sept 2011;17(18):6012-20. | - |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | 2. Dubsky P, Filipits M, Jakesz R, Rudas M, Singer CF, Greil R, et al. EndoPredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. Ann Oncol. mars 2013;24(3):640-7.<br><br>3. Filipits M, Dubsky P, Rudas M, Greil R, Balic M, Bago-Horvath Z, et al. Prediction of Distant Recurrence Using EndoPredict Among Women with ER $^{+}$, HER2 $^{-}$ Node-Positive and Node-Negative Breast Cancer Treated with Endocrine Therapy Only. Clin Cancer Res. 1 juill 2019;25(13):3865-72.<br><br>4. Buus R, Sestak I, Kronenwett R, Denkert C, Dubsky P, Krappmann K, et al. Comparison of EndoPredict and EPclin With Oncotype DX Recurrence Score for Prediction of Risk of Distant Recurrence After Endocrine Therapy. J Natl Cancer Inst. nov 2016;108(11):djw149.<br><br>5. Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C, et al. Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor–Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol. 1 avr 2018;4(4):545.<br><br>6. Sestak I, Filipits M, Buus R, Rudas M, Balic M, Knauer M, et al. Prognostic Value of EndoPredict in Women with Hormone Receptor–Positive, HER2-Negative Invasive Lobular Breast Cancer. Clin Cancer Res. 1 sept 2020;26(17):4682-7.<br><br>7. Martin M, Brase JC, Calvo L, Krappmann K, Ruiz-Borrego M, Fisch K, et al. Clinical validation of the EndoPredict test in node-positive, chemotherapy-treated ER+/HER2− breast cancer patients: results from the GEICAM 9906 trial. Breast Cancer Res. avr 2014;16(2):R38.<br><br>8. Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L, et al. Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study. Breast Cancer Res Treat. févr 2016;156(1):81-9.<br><br>9. Constantinidou A, Marcou Y, Toss MS, Simmons T, Bernhisel R, Hughes E, et al. Clinical Validation of EndoPredict in Pre-Menopausal Women with ER-positive, HER2-negative Primary Breast Cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 31 août 2022;CCR-22-0619.<br><br>10. NICE. Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer. Diagnostic guidance DG34. 2018; | |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | 11. Sestak I, Martín M, Dubsky P, Kronenwett R, Rojo F, Cuzick J, et al. Prediction of chemotherapy benefit by EndoPredict in patients with breast cancer who received adjuvant endocrine therapy plus chemotherapy or endocrine therapy alone. Breast Cancer Res Treat. juill 2019;176(2):377-86.<br><br>12. Ettl J, Anders SI, Hapfelmeier A, Paepke S, Noske A, Weichert W, et al. First prospective outcome data for the second-generation multigene test Endopredict in ER-positive/HER2-negative breast cancer. Arch Gynecol Obstet. déc 2020;302(6):1461-7.<br><br>13. Klein E, Josipovic A, Anders S, Noske A, Mogler C, Weichert W, Hapfelmeier A, Kiechle M, Ettl J et al. Long-term outcome data using Endopredict as risk stratification and chemotherapy decision biomarker in hormone receptor positive, HER2-negative early breast cancer. SABCS. 2022;<br><br>14. Lehmann-Che J, Wong J, Cayre A, Lcaroix L, Lacroix-triki M, Arnould L, et al. SiMoSein, a real-life prospective evaluation of EndoPredict (EPclin) use in early ER-positive, HER2-negative breast cancers. Poster. ESMO. 2021.<br><br>15. Penault-Llorca F, Dalenc F, Chabaud S, Cottu P, Allouache D, Cameron D, et al. Prognostic value of EndoPredict test in patients screened for UNIRAD, a UCBG randomized, double blind, phase III international trial evaluating the addition of everolimus (EVE) to adjuvant hormone therapy (HT) in women with high-risk HR+ HER2- early breast cancer (eBC). Poster. SABCS. 2021.<br><br>16. Schmitt W, Jank P, Hoffman I, Pfitzner B, Lenz L, Clegg W, Keil E, Ratzel S, Cogan E, Blohmer J, Wimberger P, Kronenwett R, Horst D, Carsten D et al. Retrospective evaluation of outcomes in a real-world, prospective cohort using EndoPredict : results from the Charité registry. 2022;<br><br>17. Sestak I, Dowsett M, Cuzick J. NICE request - TransATAC data analysis. 2017.<br><br>18. Holt SD, Sai-Giridhar P, Verrill M, Pettit L, Rigg A, Hickish T, et al. Abstract P6-01-11: A UK prospective multicentre decision impact, decision conflict and economic evaluation of the use of Oncotype DX® to guide chemotherapy in 680 women with hormone receptor positive, HER2 negative breast cancer and 1 to 3 nodes involved. Cancer Res. 1 mars 2023;83(5_Supplement):P6-01-11-P6-01-11.<br><br>19. Sestak I, Martín M, Dubsky P, Kronenwett R, Rojo F, Cuzick J, et al. Prediction of chemotherapy benefit by EndoPredict in patients with breast cancer who received adjuvant endocrine therapy plus chemotherapy or endocrine therapy alone. Breast Cancer Res Treat. juill 2019;176(2):377-86. | |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | 20. Müller BM, Keil E, Lehmann A, Winzer KJ, Richter-Ehrenstein C, Prinzler J, et al. The EndoPredict Gene-Expression Assay in Clinical Practice - Performance and Impact on Clinical Decisions. Lo AWI, éditeur. PLoS ONE. 27 juin 2013;8(6):e68252.<br><br>21. Ettl J, Klein E, Hapfelmeier A, Grosse Lackmann K, Paepke S, Petry C, et al. Decision impact and feasibility of different ASCO-recommended biomarkers in early breast cancer: Prospective comparison of molecular marker EndoPredict and protein marker uPA/PAI-1. Coleman WB, éditeur. PLOS ONE. 6 sept 2017;12(9):e0183917.<br><br>22. Fallowfield L, Matthews L, May S, Jenkins V, Bloomfield D. Enhancing decision-making about adjuvant chemotherapy in early breast cancer following EndoPredict testing. Psychooncology. avr 2018;27(4):1264-9.<br><br>23. Mokbel K, Wazir U, Wazir A, Kasem A, Mokbel K. The Impact of EndoPredict Clinical Score on Chemotherapy Recommendations in Women with Invasive ER+/HER2- Breast Cancer Stratified as Having Moderate or Poor Prognosis by Nottingham Prognostic Index. Anticancer Res. août 2018;38(8):4747-52.<br><br>24. Penault-Llorca F, Kwiatkowski F, Arnaud A, Levy C, Leheurteur M, Uwer L, et al. Decision of adjuvant chemotherapy in intermediate risk luminal breast cancer patients: A prospective multicenter trial assessing the clinical and psychological impact of EndoPredict® (EpClin) use (UCBG 2–14). The Breast. févr 2020;49:132-40.<br><br>25. Villarreal-Garza C, Lopez-Martinez EA, Deneken-Hernandez Z, Maffuz-Aziz A, Muñoz-Lozano JF, Barragan-Carrillo R, et al. Change in therapeutic management after the EndoPredict assay in a prospective decision impact study of Mexican premenopausal breast cancer patients. Buyukhatipoglu H, éditeur. PLOS ONE. 11 mars 2020;15(3):e0228884.<br><br>26. Dinh P, Graham JD, Elder EN, Kabir M, Doan TB, French J, et al. Impact of the EndoPredict genomic assay on treatment decisions for oestrogen receptor-positive early breast cancer patients: benefits of physician selective testing. Breast Cancer Res Treat. févr 2022;191(3):501-11. | |
| Exact Sciences UK Ltd. | 12 | 14, 15, 16, 17, 18, 91, | 1.4.1, 1.4.2, 1.5, 1.6, 1.7, | The report states that the assumption of predictive benefit for scenarios 2 & 3 for the post-menopausal patient group remains subject to some uncertainty. Current comparators (PREDICT/NPI) as well as three of the four interventions are prognostic only and not predictive of chemotherapy benefit. The Oncotype DX | The EAG disagrees with some of the points made by Exact Sciences. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | 143, 148, 149, 151 | 3.5.8, 4.2.1.2, 4.3.6, 4.4, 5.1.2, 5.6 | test has been proven in an independently conducted prospective Phase III RCT to predict that postmenopausal LN+ patients with RS 0-25 do not benefit from chemotherapy and can be safely spared treatment.<br><br>**Prognosis** gives the risk of recurrence for a patient if treated with endocrine therapy alone but does not inform whether adding chemotherapy will reduce this risk (not in scope).<br>**Prediction** informs the reduction in risk of recurrence by adding chemotherapy to patients' treatment plan (in scope).<br><br>The decision problem for DAP71 is to decide which LN+ patients benefit from chemotherapy. Prognostic information alone does not directly inform who will benefit from chemotherapy.<br><br>• **Prediction of a lack of CT benefit in postmenopausal LN+ patients has only been proven for the Oncotype DX test.**<br>• **RxPONDER very clearly shows that with the Oncotype DX test an RS 0-25 is predictive of a lack of chemotherapy benefit (postmenopausal women).**<br>• This directly addresses the very clear and urgent need to reduce the significant over-treatment with chemotherapy for post-menopausal LN+ patients, as the large-scale meta-analysis of 100,000 breast cancer patients in chemotherapy studies conducted by the Oxford Overview of breast cancer found that **less than 10% of post-menopausal LN+ patients benefit from chemotherapy** (EBCTCG. Lancet. 2012;379(9814):432-4.).<br>• **Prognostic-only tools/tests (PREDICT, NPI, Prosigna, EPClin) categorize the majority of LN+ patients (e.g., 60-80%) as high risk, but do not inform CT benefit (which is the scope of this appraisal).**<br>• In contrast, the Oncotype DX test identifies only ~15% with a high Recurrence Score and ~85% who will not benefit from chemotherapy and can safely avoid unnecessary side-effects and the associated costs to the NHS.<br>• **As stated by the EAG, the key economic driver of the decision problem should be the reduction in the number of patients receiving chemotherapy who will not benefit.**<br>• The ability of the Oncotype DX test to predict a lack of chemotherapy benefit among post-menopausal LN+ patients is highly conclusive and widely acknowledged by UK clinicians and in international clinical consensus practice guidelines.<br><br>A crucial point is that the dominant results for Oncotype DX in these scenarios are driven by a reduction in chemotherapy in patients who would not have benefitted from the treatment. Therefore, the key question is whether the test can reliably identify patients who would not have benefitted from chemotherapy treatment. | These are detailed below.<br><br>The scope of this NICE appraisal includes outcomes relating to both the prognostic and predictive ability of the four tumour profiling tests. It is not the case that a non-predictive test has no value. It is important that both of these factors are considered in the assessment of each of the tests.<br><br>As discussed in the EAG report, RxPONDER does not show a predictive benefit for Oncotype DX, as the tests for interaction on IDFS were not statistically significant within the range RS 0-25. The argument that Oncotype DX is predictive of chemotherapy benefit in post-menopausal women is reliant on the consideration of external data (specifically SWOG-8814) in women who were not eligible for RxPONDER (those with |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Phase III Level 1A prospective RCT evidence from the RxPONDER study addressed this question directly and definitively for Oncotype DX. Post-menopausal LN+ patients with a RS 0-25 by Oncotype DX (~85%) were found not to derive any benefit from chemotherapy, as per the study protocol (in fact chemotherapy was numerically associated with poorer outcomes for these patients). For the first time, these patients who can be safely spared toxicity from chemotherapy treatment, without a detrimental impact on cancer outcomes. This has only been demonstrated for the Oncotype DX test and none of the other tests in DAP71.<br><br>A UK clinician-led decision impact study in 680 patients across 14 centres that included only the Oncotype DX test in LN+ patients (Holt et al.) showed that clinicians forego chemotherapy for the vast majority of patients with a low RS result, which would indicate a high degree of confidence and low level of uncertainty that no benefit of chemotherapy is expected for these patients.<br><br>We strongly question the relevance of scenario 4 and its inclusion in the report, when more recent evidence from a phase III Level 1A prospective RCT for Oncotype DX has become available since DG34. Scenario 4 relies on the now disproven assumption that post-menopausal LN+ patients benefit equally from chemotherapy, regardless of whether they have a RS 0-25 or 26-100. This assumption can no longer be supported.<br><br>The conclusions for the Oncotype DX test (dominant strategy) are not reliant on the assumed hazard ratio for chemotherapy benefit in the high RS group. We conducted exploratory analysis in the EAG model, using the highly conservative assumption whereby the benefit of chemotherapy in patients with an RS of >25 was assumed to be in line with the average benefit across all patients (HR from EBCTCG, as per scenario 4, applied to the RS >25 arm), the Oncotype DX test still dominates current decision making.<br><br>The lack of benefit for patients with Oncotype DX RS 0-25 is what drives the cost-effectiveness in the EAG and Exact Sciences models and is supported by a high level of evidence. | higher RS), and from other from evidence which suggests that chemotherapy is effective in unselected post-menopausal women (e.g., the EBCTCG meta-analysis). As noted in the DG34 document, the previous EAG report by Harnan *et al.* (2019) and the current EAG report, the findings of SWOG-8814 remain subject to some uncertainty, with some analyses demonstrating a significant interaction, and others not. Other concerns with this study have been highlighted by Agendia in this consultation response (see comments below). As such, there is still some uncertainty around the predictive benefit of Oncotype DX.<br><br>As noted in the EAG report, all of the Exact Sciences base case economic analyses indirectly assume that Oncotype DX is predictive of benefit. EAG |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | | Scenario BC4 (Oncotype DX using TransATAC and assuming no predictive benefit) was presented in the current EAG report because it reflects the non-predictive analysis presented in DG34. A similar analysis assuming no predictive benefit was also presented in the Berdunov *et al.* JME publication, but was not included in the Exact Sciences submission. As acknowledged in the EAG report, BC4 uses the older 3-level cut-offs and may be less relevant for decision-making.<br><br>The EAG highlights that the additional sensitivity analysis described in the company's response is still assuming a predictive benefit for Oncotype DX because different HRs are being used in the genomic low- and high-risk groups. |
| Exact Sciences UK Ltd. | 13 | 143 | 4.4 | The following sentence (with the use of the word 'but') seems to imply that the non-statistically significant test for interaction for the post-menopausal LN+ patient subgroup in RxPONDER was an unfavourable result in terms of supporting the clinical utility of the Oncotype DX test. This is not the case, as we describe below. | **The EAG takes the point that if there is no benefit of chemotherapy with lower RS scores then** |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | *"RxPONDER indicates that chemotherapy is not beneficial to post-menopausal women who have an RS of 0-25, **~~but~~** the test for interaction between the treatment group and the continuous RS, when adjusted for the continuous RS, menopausal status, and treatment group, was not statistically significant (p=0.35)."*<br><br>The RxPONDER study proved the ability of the Oncotype DX test to identify a large proportion of post-menopausal LN+ patients who would not benefit from chemotherapy, therefore directly addressing the main clinical and economic need to reduce significant over-treatment.<br><br>Hence, we request that the above sentence should be rephrased so as not to imply a negative result from the RxPONDER study;<br><br>*"RxPONDER indicates that chemotherapy is not beneficial to post-menopausal women who have an RS of 0-25, **as** the test for interaction between the treatment group and the continuous RS, when adjusted for the continuous RS, menopausal status, and treatment group, was not statistically significant (p=0.35)."*<br><br>In line with the EAG's assessment:<br>• The RxPONDER study was not designed to investigate the predictive ability of the Oncotype DX test in general, as this had already been indicated in the SWOG-8814 study.<br>• Hence patients with RS>25, for whom a significant benefit from chemotherapy had been found, were excluded in RxPONDER for ethical reasons (to avoid undertreating patients randomised to no chemotherapy).<br>• The study was originally designed with a specific step-wise analysis plan to determine whether a clinically and statistically significant benefit of chemotherapy exists for LN+ patients within the RS 0-25 range, to inform with greater precision the RS cut-off that should be used for recommending chemotherapy.<br>• The study found no significant chemotherapy benefit exists for post-menopausal LN+ patients within the RS 0-25 range, supporting the conclusion that ~85% of patients can, for the first time with precision, be safely spared unnecessary chemotherapy and the associated risks of side-effects.<br><br>The negative test for interaction in the primary analysis that RxPONDER could be misinterpreted to mean that it was a 'negative study' or disproved the ability of Oncotype DX to predict chemotherapy benefit, however this is not the case. This is not a simple drug vs. placebo study. The finding that post-menopausal LN+ patients with a low RS result do not benefit from chemotherapy further supports the ability of the Oncotype DX test to predict a lack of chemotherapy benefit in a large proportion of patients. International | **one may not expect to see a benefit in any of the range RS 0-25. Conversely, if there is a benefit for patients with RS>25 then one might expect to see a trend towards benefit in higher RS scores within the RS 0-25 group. Given this uncertainty in interpretation, the EAG considers it would be most appropriate to word this sentence neutrally as detailed below. We have also clarified that this finding relates only to the RS range 0-25. This change will be made in the EAG report Section 4.4:**<br><br>***"RxPONDER indicates that chemotherapy is not beneficial to post-menopausal women who have an RS of 0-25. The test for interaction in RxPONDER between the treatment group and the continuous RS, when adjusted for the continuous RS, menopausal status, and*** |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | clinical practice guidelines were subsequently updated to reflect the practice-changing findings from the RxPONDER study. | ***treatment group, was not statistically significant within the range RS 0-25 (p=0.35)."*** |
| Exact Sciences UK Ltd. | 14 | 142, 143 | 4.4 | The EAG states the following:<br><br>*"Unlike the analyses presented to inform DG34, the current EAG model applies pre- and post-test chemotherapy probabilities for all tests based on analyses of the same UK decision impact study of Oncotype DX evaluated using both the older 3-level and newer 2-level RS cut-offs (Holt et al.)."*<br><br>The EAG acknowledges that *"This absence of relevant evidence (for the Prosigna, EndoPredict and MammaPrint tests) means that the results of the analyses presented for each of these tests is highly uncertain and should be interpreted with some caution."*<br><br>We agree that this approach is highly uncertain for three of the interventions other than the Oncotype DX test, it is therefore not appropriate to apply chemotherapy allocation data for the Oncotype DX® test to the other tests, given that the tests are so different and not at all interchangeable.<br><br>• For example, EPClin is expected to classify 77% of patients as high risk whereas the Oncotype DX test classified 14% as high RS (Holt et al.). There is no reason to believe that clinicians would interpret the tests in the same way and recommend chemotherapy to the same proportion of patients classified into the high-test score categories by the different tests.<br>• The current approach is even more questionable considering that the other tests do not provide the same type of information to clinicians to help guide chemotherapy treatment decisions. Unlike Oncotype DX, the other tests do NOT inform whether a patient with a high score is likely to benefit from CT or that those with a low score will not benefit.<br>• Furthermore, the Oncotype DX test has been clinically developed and validated with specific relevant cancer genes and the other tests measure almost completely different genes.<br><br>The assumption made in the base case analyses that all tests would be interpreted the same way in clinical practice is not supported by evidence and does not reflect clinical reality or the scientific basis for these tests. Each test should be evaluated separately based on its respective evidence.<br><br>Due to the fact that the tests are entirely different, they classify patients into test score groups very differently, as reflected by the large difference in the proportion of patients assigned to the high-test score (11% for Oncotype, compared to 60% for Prosigna, 77% for EPClin and 31% to MammaPrint). | A similar pragmatic assumption was also necessary in DG34. The EAG report clearly states this is a modelling assumption (Section 4.3.2), highlights the absence of decision impact studies in LN+ women for other tests (Section 3.6), discusses the limitations and uncertainties associated with the absence of evidence for other tests (Section 5.3) and highlights the need for further evidence on decision impact for these tests (Section 5.6). The EAG believes that this is sufficient and as such no amendment has been made to the EAG report. The EAG believes it is up to the Appraisal Committee to determine whether this assumption is appropriate, and if not, what the best alternative approach would be. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | In the study by Holt et al., in the post-RxPONDER time period, 100% of high RS (>25) patients received CT and only 4% for low RS (0-25) patients (Academic In Confidence). This is due to the high-quality large Phase III study evidence for Oncotype DX. The other tests do not have the same level of evidence to support decision-making.<br><br>In certain scenario analyses conducted by the EAG, the ICER for non-Oncotype DX tests are approximately double the base case analysis.<br><br>If any scenarios were to be included in the report based on evidence imputed from Oncotype DX, they should be much more clearly highlighted as '**based on Oncotype DX evidence**' throughout the report and a more extensive and explicit explanation provided about their lack of available evidence and, crucially, the potential consequences for patient outcomes and the cost-effectiveness estimate of the assumptions proving not to be accurate. | |
| Exact Sciences UK Ltd. | 15 | 119 | 4.3.3 | We would like to highlight an important dynamic resulting from the economic modelling for tests assumed to be prognostic only.<br><br>• Based on the modelling approach, it is unlikely that any test that significantly reduces chemotherapy and therefore addresses the unmet need, would be able to demonstrate cost-effectiveness.<br>• Prognostic-only tests which lead to high rates of chemotherapy are most likely to be found to be cost-effective, which runs counter to the unmet clinical need.<br><br>The assumption that all patients assigned to chemotherapy will derive a benefit, irrespective of their genomic risk profile, inadvertently favours tests which increase chemotherapy use, and 'penalises' any tests that decrease chemotherapy use. This is particularly pertinent for the comparatively higher risk LN+ patient group.<br><br>By this logic, a strategy which involves assigning all patients to chemotherapy would be considered the most cost-effective. However, this of course runs counter to the intended purpose and impact of gene expression profiling tests, especially for LN+ patients, which is to reduce the high level of chemotherapy over-treatment.<br><br>Based on the modelling approach, it is unlikely that any test, when assumed to be prognostic only, that significantly reduces chemotherapy, would be able to demonstrate cost-effectiveness. This means there is | The EAG is unsure exactly what argument the company is making on this point. If it is the case that chemotherapy is, on average, beneficial in reducing the absolute risk of distant recurrence, regardless of genomic risk, then a prognostic test which substantially reduces use of chemotherapy in that population is unlikely to be either clinically effective or cost-effective. That is not to say however that a prognostic test cannot be cost-effective. The benefit of these tests |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | no possibility of both addressing the clinical unmet need and demonstrating cost-effectiveness, due to the inherent bias in the modelling approach.<br><br>A further issue with this assumption of equal chemotherapy benefit across test risk groups is the large difference in the proportion of patients assigned to chemotherapy across tests: Oncotype DX test: 31% (scenario 4), Prosigna: 75% (scenario 5), EPClin: 76% (scenario 6), MammaPrint: 43% (scenario 7). It is questionable whether it makes logical sense for the same chemotherapy benefit to exist across all tests, in light of such a large variation in chemo allocation.<br><br>We are not aware of any obvious solution to this issue in the modelling approach and indeed this approach was used in exploratory analysis in the Exact Sciences model.<br><br>However, we nonetheless believe it is an important issue to highlight to the NICE committee, as a concern could be that prognostic-only tests which classify a large proportion of patients as high risk and are assumed to lead to high chemotherapy treatment rates are 'rewarded' from a health economic perspective, despite this running contrary to the widely acknowledged clinical unmet need.<br><br>This issue was briefly acknowledged in DG34 in section 5.11 *"The committee considered the modelled impact of these data on chemotherapy use, and noted that although clinical and patient experts thought that the main benefit of the tests was in avoiding unnecessary chemotherapy, most tests were estimated to increase chemotherapy use at least in some subgroups (see section 4.49). The committee concluded that there was much more uncertainty around chemotherapy decision making for the 2-level tests, and for the subgroups who were not included in the original NICE recommendation on tumour profiling tests (LN-negative disease and a NPI of 3.4 or less, and LN-positive disease)."* It was not clear how such a clear and important mismatch between clinical unmet need and expected impact of certain tests was taken into consideration when deciding on the final recommendations.<br><br>We hope this issue will be a central consideration for the NICE Committee, given the very clear need to reduce the significant over-treatment among LN+ post-menopausal patients.<br><br>The described dynamic within the model is not relevant to tests assumed to be predictive of chemotherapy effect i.e., able to spare chemotherapy for LN+ patients who will not benefit. As such, scenarios 2/3 found the Oncotype DX test to dominate current decision-making, based on the RxPONDER study which demonstrated that post-menopausal LN+ patients spared chemotherapy based on a RS 0-25 would not have benefitted from chemotherapy treatment. | depends on which patients receive chemotherapy and their underlying risk of recurrence. The economic models developed by Exact Sciences and the EAG both illustrate that assumptions of predictive benefit are key drivers of cost-effectiveness for Oncotype DX. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| Exact Sciences UK Ltd. | 16 | 92 | 4.2.1.2 | The EAG identified uncertainty around the probability of being in the Oncotype DX RS >25 group, due to the assumption being calculated based on the number of patients who were screened for eligibility for entry into RxPONDER who had an RS >25 as the numerator (N=1,035) and the overall number of patients who were registered for screening in RxPONDER as the denominator (N=9,383).<br><br>We agree with the EAGs suggestion that "…within the RxPONDER trial, the proportion of women who had an RS of >25 lies somewhere between a minimum value of 0.11 (assuming that all other excluded patients have an RS of <25 [1035/9383]) and a maximum value of 0.17 (including only patients with a known RS in the calculation [1035/6118])."<br><br>In the Holt et al. study, the proportion of patients assigned to the high-risk category (RS>25) was 14%, which is consistent with the EAG's suggestion.<br><br>Applying this assumption in the model does not change the conclusions that the Oncotype DX test dominates current decision-making. | No amendment required. |
| Exact Sciences UK Ltd. | 17 | | | The EAG identified uncertainty around relevant cut-offs for NICE decision-making.<br><br>We do not believe there is significant uncertainty.<br><br>In the UK decision-impact study by Holt et al., the post-test chemotherapy allocation reported for patients tested after the results of the RxPONDER study were available, show that clinicians follow the Oncotype DX RS 25 cut-off for post-menopausal patients, in line with the results of the RxPONDER study.<br><br>International clinical practice guidelines have also been updated to reflect the results of the RxPONDER study.<br><br>The IFU for the Oncotype DX test will be updated with the relevant cut-offs as part of the transition to IVDR. | The EAG report explains that the current IFU reflects the older RS cut-offs. The company has clarified that a single cut-off of RS 25 is how they intend the test to be used. However, it should be noted that the external evidence supporting the predictive benefit of Oncotype DX, which is used in both the EAG and Exact Sciences models, reflects the older cut-offs (0-17, 18-30, >30). |
| Exact Sciences UK Ltd. | 18 | 11 | 1.1 | We refer to the following quote from the Executive Summary: | The EAG believes that the company is confusing absolute and relative |

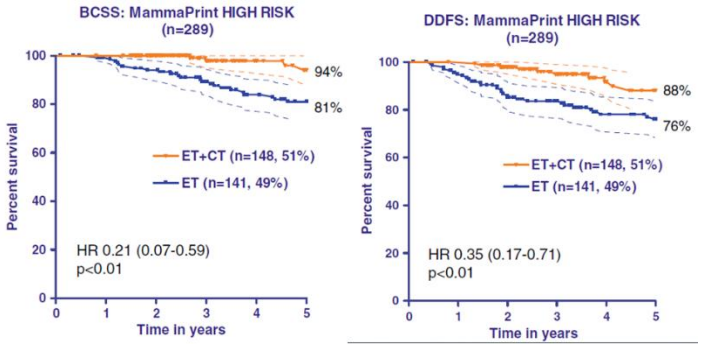
| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | *"Avoiding chemotherapy in patients at low risk of recurrence, who would therefore obtain limited benefit, avoids the unpleasant side effects of chemotherapy and may reduce expenditure on both the chemotherapy itself and the treatment of adverse effects."*<br><br>It is not necessarily correct to assume that patients with a lower risk of recurrence would not obtain a clinically meaningful benefit from chemotherapy and safely avoid treatment. This would only be true for patients with such a low risk that chemotherapy could not offer a clinically meaningful benefit (based on the true relative risk reduction and not only the average relative benefit artificially assumed across all patients).<br><br>It is also not necessarily correct to assume that patients with a higher risk of recurrence will benefit from chemotherapy (as demonstrated by the categorization of most LN+ patients as high risk, but only <10% actually benefit).<br><br>Studies randomising patients to endocrine therapy alone vs. chemo-endocrine therapy are needed to determine the interaction between treatment and test score.<br><br>In the interest of advancing a precision medicine approach to cancer treatment, it is important not to use the terms 'risk' (or 'prognosis') and 'benefit' (or 'treatment effect') interchangeably as this can perpetuate misunderstanding. 'Risk' or 'prognosis', in the context of DAP71, refers to risk of distant recurrence with endocrine therapy alone, whereas treatment benefit refers to the magnitude of reduction in risk from adding chemotherapy to endocrine therapy.<br><br>We suggest amending the wording in this section of the report and in the final guidance. | benefit. Regardless of any consideration of predictive benefits, if patients are already at a low risk of recurrence, the absolute benefit of adjuvant chemotherapy will also be small. No amendment is required. |
| Exact Sciences UK Ltd. | 19 | 18 | 1.7 | A suggestion is made that further studies demonstrating a statistical interaction between Oncotype DX RS and long-term chemotherapy benefit across the full range of RS would help to address uncertainty about whether Oncotype DX is predictive of chemotherapy benefit.<br><br>Please see our earlier comment that the RxPONDER study definitively proves a lack of chemotherapy benefit in post-menopausal LN+ patients with a RS 0-25, therefore directly addressing the clinical unmet need to reduce the significant chemotherapy over-treatment issue among the LN+ patient subgroup.<br><br>We note that the EAG did not make any suggestion for other tests to develop evidence of the relationship between test score and treatment effect (prediction of treatment benefit). Considering that this is the gold standard for any diagnostic test, and that prognostic only information is of more limited clinical utility, we would suggest that it would be appropriate for the EAG to include this as a suggestion in this section. | Please refer to the EAG's conclusions around the evidence supporting predictive benefits of Oncotype DX.<br><br>The EAG's suggested research priority focussed on Oncotype DX because this is the only test for which there is any evidence of predictive benefit on |

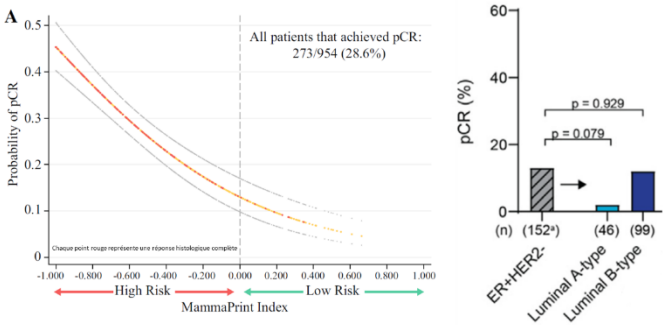
| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | | recurrence risk in LN+ women, although this remains subject to uncertainty. The report has not been amended. |
| Exact Sciences UK Ltd. | 20 | 149 | 5.2.1 | *"A major limitation is that it is difficult to collect new data on predictive ability because it is not considered ethical to randomise patients who are high-risk on any of the tests to chemotherapy vs. no chemotherapy. Therefore, although there are prospective RCTs for the effect of chemotherapy in low- to intermediate-risk patients, data for high-risk patients are limited to retrospective reanalyses of trials, plus observational data in which test results may have influenced treatment."*<br><br>We believe that the term 'treatment benefit' should be used in place of the term 'risk' to ensure accuracy of the statement.<br><br>We are not aware of any clinical consensus where it is considered unethical to randomised patients who are high-risk to chemotherapy vs. no chemotherapy. Indeed, several studies have done so / continue to do so.<br><br>Ethical concerns about randomising patients to no treatment arise when it is considered to have been sufficiently demonstrated that a certain patient group is likely to benefit from treatment, as was the case with RxPONDER, due to the evidence from SWOG-8814.<br><br>The fact that it is considered unethical to conduct an RCT across the full range of RS provides solid evidence that chemotherapy should not be withheld from patients with high RS results and reinforces the predictive capability of the Oncotype DX test.<br><br>We would suggest amending the wording in this paragraph to focus on treatment effect rather than risk to avoid any misinterpretation. Tests which have demonstrated prognostic ability could seek to study the relationship between test score and treatment effect, without ethical concerns that we are aware of. | The EAG believes that this point remains appropriate and clear, and no amendment has been made to the report. |
| Agendia NV | 21 | General comment | General comment | **General comment about how the predictive value of Oncotype DX and MammaPrint seem to be assessed with partiality.**<br><br>Agendia believes that some one-sidedness is being applied by the EAG in assessing the predictiveness of tumour profiling tests, particularly concerning the inclusion and interpretation of clinical evidence. We believe that while the EAG appears very lenient towards the methodological flaws of the SWOG 8814 analysis described below, it applies very strict criteria to MammaPrint by not considering the evidence | The EAG disagrees. The evidence in support of a predictive benefit for Oncotype DX in post-menopausal LN+ women is based on RxPONDER, as well as the re-analysis |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | provided by Agendia in the company's response to EAG on May 30th about the predictive value of MammaPrint. Details and comparative elements are provided below.<br><br>The EAG considered studies eligible when the study population is constituted by at least 80% of patients with lymph node positive (LN+) disease. This has consequently directly led the EAG to exclude results on mixed cohorts, i.e., those with N0 and N1 disease, in which >20% of the population was N0. On the other hand, studies that consisted of patients with 4 or more positive lymph nodes were not excluded from the evidence review, even if >20% of patients had 4 or more positive lymph nodes (N2 or N3), while the present assessment focuses on the use of tumour profiling tests (GEP-test) in patients with 1-3 positive lymph nodes (N1). For reasons explained in forthcoming sections, it would have made more sense to either reverse this exception or allow both.<br><br>Literature currently considers N0 tumours and N1 tumours to be biologically more similar to each other, as compared with N1 and N2/N3 tumours. This is also suggested by the TNM stages, where patients generally considered for GEP-tests with N0 and N1 disease are both classified as Stage II breast cancer, while N2/N3 disease is regarded as Stage 3 breast cancer.<br>Furthermore, in any analysis including a considerable number of patients with N2/N3 disease, the chemotherapy predictiveness results in patients with genomic high risk tumours are likely to be highly driven by the feeble prognosis of patients with tumours characterized as both very high clinical risk and high genomic risk. Indeed, as displayed in Vliek and colleagues (2017), patients with 4 or more positive nodes with a C-high/G-high tumour had a 64.2% DRFI at 10-years, while patients with 1-3 positive nodes with a C-high/G-high tumour had a 79.5% DRFI at 10-years, a 15.3% difference in prognosis. In line with the previous paragraph, MINDACT investigators (Piccart et al., 2021), have shown that patients with a HR+/HER2-/N0 C-high/G-high tumour had an 8-year DMFI of 89.1%, and HR+/HER2-/N1 C-high/G-high had an 8-year DMFI of 80.9%, with a -8.2% difference in prognosis. Taken together, this demonstrates that support the fact that including >20% N2/N3 disease is likely to overestimate chemotherapy benefit because of poor prognosis, while a study that includes >20% N0 disease is likely to underestimate chemotherapy benefit because of a better prognosis. Yet, the methodology of the current assessment only allows to include studies that are in principle more likely to overestimate chemotherapy benefit for patients with N1 disease and exclude those that are in principle more likely to underestimate chemotherapy benefit for patients with N1 disease.<br><br>For the aforementioned reasons, we hope that the EAG/NICE can reconsidered its stance on some of the study inclusion and interpretation and therefore explore the positive predictive value of chemotherapy benefit for MammaPrint High Risk in the model. We understand that the evidence base for this might not be as ideal as wanted for this evaluation. As mentioned in the EAG report itself, gathering the proper evidence for a positive predictive benefit is now unethical and will for that reason never happen. Besides, | of the SWOG-8814 RCT for women with RS>26. The evidence is not ideal and uncertainty around the predictive benefit of Oncotype DX in LN+ women is discussed in DG34 as well as in the current EAG report.<br><br>For MammaPrint, the EAG's review did not find sufficient evidence to support an assumption of predictive benefit on recurrence-based outcomes in LN+ women. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | the reanalysis of SWOG-8814 with Oncotype DX has severe methodological flaws, some of those, but not all are mentioned by the EAG. Yet these severe methodological flaws are only described as 'some uncertainty', resulting in inclusion of the predictive value in RS>25 being part of the base case scenario.<br><br>**Limitations of re-analysis of SWOG-8814 with Oncotype DX**<br>• 38.1% of the cohort has 4 or more positive nodes (Albain 2010, Table 1), outside of the population of interest.<br>• 4.6% of the cohort had a tumour larger than 5 cm (Albain 2010, Table 1) outside of the population of interest.<br>• 11.7% of the cohort likely has HER2+ disease (Albain 2010, Table 1), outside of the population of interest.<br>• Hazard Ratio is presented for Disease Free Survival, not for Distant Recurrence Free Interval. Transferability of a DFS HR to DRFI is highly questionable. In MINDACT we observed that only 33% of DFS events were distant metastasis (Piccart 2021, Table S4). This means that the HR used to model predictiveness for DRFI events for Oncotype DX in the model is likely based on a HR that mostly does not relate to distant recurrence events.<br>• The Hazard Ratio from SWOG-8814 is based on the old ODx high risk cut-off, RS≥30, instead of the cut-off RS>25 used in clinical practice. This creates bias as even if the HR is true, it will likely overestimate chemotherapy benefit for patients with RS25-30. Results from the TAILORx RS 26 to 100 reveal that in Node-Negative patients 43% of patients is RS 26-30 (Sparano, 2020, Table 1 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6777230/). This means that if a similar proportion of patients with Node-Positive disease fall within this same group, that the chemotherapy benefit is likely overestimated in 43% of patients falling in the RS>25 group.<br>• The multivariate analysis used to calculate the interaction p value did not include all known clinical and pathological variables. The analysis only corrected for 4 or more lymph nodes, but not for tumour grade, tumour size, ER- PgR- and HER2-status by RT-PCR assay, all of which were available in the analysis. If one or more of these factors were considered in the analysis, it is likely that the results of the interaction test would no longer be statistically significant.<br><br>Lastly, the most significant methodological flaw from the Albain 2010 analysis (SWOG-8814), not considered in the current assessment, is that in the P interaction analyses Oncotype DX **was not calculated as a chemotherapy effect and risk group.** The hazard ratio was calculated using patients with an increment of 50 units difference in recurrence score, i.e., only including the lowest and the highest recurrence scores (Albain 2010, Table 2). This is different from how the hazard ratio is applied in the model, where the Hazard ratio is used as if the hazard ratio has a significant p interaction in ODx "Low Risk group" versus ODx "High Risk group". From TAILORx we know that only 116 patients were classified as RS 51-100, that is 1.2% of the full TAILORx population (Sparano, 2020, Table 1 | |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6777230/). This means that the interaction test for chemotherapy effect for risk score with an increment of 50, already rightly described with uncertainty by EAG due to the borderline significance, barely applicable to clinical practice, considering that only 1% of patients are classified as RS>50.<br><br>Based on the severe methodological flaws presented above in SWOG-8814, there seem to be two set of rules applied to the assessment of the predictive value of Oncotype DX and MammaPrint. We hope that the present response will bring more elements to light for reconsideration of the inclusion of MammaPrint studies to this assessment.<br><br>Comment #3 summarizes the evidence that supports the assumption of predictive benefit of chemotherapy in MammaPrint High risk. The table below compares MammaPrint’s Knauer study to Oncotype DX’s SWOG-8814 study. In our view this suggest for a model to choose between two options:<br>1. to explore the predictiveness of both MammaPrint and Oncotype DX in the model;<br>2. or to explore the predictiveness of neither test in the model.<br><table><tr><th></th><th>MP – Knauer</th><th>ODx – SWOG-8814</th></tr><tr><td>Significant HR on relevant endpoint (Distant Recurrence)</td><td>Yes</td><td>No</td></tr><tr><td>Proportion N0/N1/N2</td><td>49%/51%/0%</td><td>0%/61.9%/38.1%</td></tr><tr><td>Proportion HER2+</td><td>11%</td><td>11.7%</td></tr><tr><td>Proportion T-size >5cm</td><td>1%</td><td>4.7%</td></tr><tr><td>Randomized CT vs no CT</td><td>No</td><td>Yes</td></tr><tr><td>Risk groups in accordance with use in clinical practice</td><td>Yes</td><td>No</td></tr><tr><td>Significant P-interaction Group Low vs High</td><td>No</td><td>No</td></tr><tr><td>Significant P-interaction on continuous score</td><td>No</td><td>Yes, with a 50-point RS difference</td></tr><tr><td>Group size in Genomic high risk</td><td>Total: 289<br>CT: 148<br>No CT: 141</td><td>Total: 118<br>CT: 71<br>No CT: 47</td></tr></table></td><td></td></tr>
<tr><td>Agendia NV</td><td>22</td><td>13/14</td><td>1.4</td><td>The report states “… however, no HRs or significance tests were reported for prognostic ability.”<br><br>The prospective, randomised, MINDACT trial proved MammaPrint’s independent prognostic ability in a multivariate analysis for Distant Metastasis Free Survival after correcting for clinical factors, including lymph node status, age, and chemotherapy. Results are presented in Table S15 of Piccart et al. 2021 and are applicable to MammaPrint’s prognostic ability in HR+/HER2-/N1 breast cancer, regardless of age.<br><br>MammaPrint High Risk vs Low Risk – Hazard Ratio: 2.13 (P-value < 0.0001)<br><br>Piccart M, van ’t Veer LJ, Poncet C et al. 70-gene signature as an aid to treatment decision in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. Lancet Oncology 2021; 22(4), 476-88</td><td>This analysis was not included in the EAG report as it is not restricted to LN+ patients, as specified in the NICE scope.</td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | MINDACT reported on the prognostic ability of MammaPrint in a manner where results are applicable to patients with HR+/HER2-/N1 disease. As such, the aforementioned and current statement is unfounded and would warrant a correction to recognize that MammaPrint was significantly prognostic for Distant Metastasis Free Survival after adjusting for clinical factors. | |
| Agendia NV | 23 | 14 | 1.4 | The section *Prediction of chemotherapy benefit: MammaPrint* does not include any reference to other studies, listed below, that suggest MammaPrint has positive predictive value for the benefit of chemotherapy.<br><br>**Knauer et al cohort:** In this pooled cohort (90% HR+, 51% N1), the addition of adjuvant CT to ET provided significant and clinically important benefits of 13% in 5-year breast cancer-free survival (HR 0.21 [0.07-0.59], p < 0.01) and 12% in 5-year distant disease-free survival (HR 0.35 [0.17-0.71], p <0.01).<br><br>*Knauer M et al. The predictive value of the 70-gene signature for adjuvant chemotherapy in early stage breast cancer. Breast Cancer Res Treat 2010; 120: 655-661.*<br><br>  | Knauer and PATH were excluded from the review because they were mixed cohorts which were not >80% LN+. Patients with node-negative disease were outside of the NICE scope for this appraisal, as highlighted in the EAG protocol.<br><br>The NBRST and NBRST2 Cohorts reflect neoadjuvant chemotherapy. As noted in the EAG protocol (Section 3.8), neoadjuvant therapy is outside of the scope of the evaluation.<br><br>The multivariate analysis of MINDACT was presented as part of the company's response to clarification questions. The EAG notes that very few CHGH patients in this analysis received no chemotherapy (N=41 and most of the patients |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | **NBRST and NBRST2 Cohorts:** In the neoadjuvant NBRST trial with approximately 1000 patients, MammaPrint and BluePrint were performed on biopsies. The investigators showed an increasing response to chemotherapy with increasing metastatic risk by MammaPrint (i.e., the higher risk the MammaPrint, the higher the pathologic Complete Response [pCR] rate - left graph). In NBRST2, a strictly European cohort, investigators observed an increase in pCR rate from 2% to 12% for patients with Low vs High Risk MammaPrint luminal tumours (right graph, Luminal A vs Luminal B). Evidence of clear association between chemotherapy sensitivity and the MammaPrint risk groups is further supported by these trials.<br><br>*Whitworth P et al. Chemosensitivity and Endocrine Sensitivity in Clinical Luminal Breast Cancer Patients in the Prospective Neoadjuvant Breast Registry Symphony Trial (NBRST) Predicted by Molecular Subtyping. Ann Surg Oncol 2016. DOI 10.1245/s10434-016-5600-x*<br><br>*Göker E et al. Treatment response and 5-yaer distant metastasis-free survival outcome in breast cancer patients after the use of MammaPrint and BluePrint to guide preoperative systemic treatment decisions. European Journal of Cancer 2022;167:92-102.*<br><br>**PATH Cohort**: A ten-year analysis of a German cohort of patients (n = 117) showed that patients with MammaPrint High Risk (n = 50, 36% N+) treated with hormonal therapy alone had an unfavourable<br>  <br>prognosis and that chemotherapy in these patients had a benefit in terms of overall survival at 10 years. This was presented in 2022 at ESMO.<br>• MammaPrint High Risk ET only: 61.5%<br>• MammaPrint High Risk CT +/- ET: 90.9%<br>• Delta of 29.4% (p = 0.052) -> Relative Risk of Chemotherapy 0.24 | included in the analysis population were LN0). |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | *Jakisch C et al. MammaPrint 10-year follow up results from a German breast cancer cohort study. ESMO Congress 2022; 162P.*<br><br>In MINDACT, CT vs no CT randomisation was not performed in C-high/G-high, because this was considered unethical, as also noted by the EAG. However, what is available is data from an analysis based on the MINDACT 2016 data cut-off that looked at the event rate of C-high/G-high HR+/HER2- patients based on actual treatment received. Of 831 patients with a C-high/G-high risk, 41 patients in MINDACT did not receive chemotherapy. A multivariate analysis of distant recurrence free interval controlling for clinical risk features was performed and showed that chemotherapy was significantly associated with a 79% risk reduction (HR 0.21 [0.10-0.46], p < 0.001).The formal test of interaction between chemotherapy and genomic risk score (low versus high) was statistically significant (P interaction = 0.013).<br><br>We believe the analyses summarized above should be considered in the *Prediction of chemotherapy benefit: MammaPrint*, even though the cohorts are mixed node-negative and node-positive cohorts. After all, there is evidence from MINDACT that interaction between chemotherapy and genomic risk score (low versus high) was statistically significant (P interaction = 0.013). Besides, the EAG did consider the results of the Oncotype DX analysis of SWOG-8814, that misleadingly showed a significant p interaction with a 50 RS point difference (e.g., RS 0 versus RS 50), instead of Low-risk group (RS<18) versus High-Risk (RS≥31) group and had various methodological flaws (See Comment #1).<br><br>If the results of SWOG-8814 are considered as sufficient evidence to consider a 'predictive for the benefit of chemotherapy' scenario for Oncotype Dx, the same thing should be done for MammaPrint. Based on the evidence provided above, it would be inconsistent to assume predictiveness for Oncotype DX if this is not done for MammaPrint. | |
| Agendia NV | 24 | 15/16 | 1.4.2 | The EAG states that there is 'some uncertainty' in the assumption of predictive benefit for Oncotype DX and indicates at the same time that the assumption strongly influences the conclusions of the economic analysis.<br><br>As explained in Comment #1, the degree of uncertainty is considerably higher than currently reported for this assumption. This directly results in an unjustifiable advantage for Oncotype DX in this assessment as opposed to other GEP-tests. In particular unjustifiable versus MammaPrint, as MammaPrint also has evidence that would support to explore a scenario where MammaPrint is predictive for chemotherapy benefit in Genomic High risk (Comment #3). | It is unclear what is meant by "the degree of uncertainty is considerably higher than currently reported." The EAG's analysis includes scenarios which include assumptions of no predictive benefit for every test and the implications of this are |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>clearly discussed in the EAG report.</td></tr>
<tr><td>Agendia NV</td><td>25</td><td>16</td><td>1.4.2</td><td>“MammaPrint is dominated by current decision-making. These results are driven by a large reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment…”<br><br>The reason why the results indicate that MammaPrint is dominated and that women would have benefited from adjuvant chemotherapy, is because evidence from the MINDACT trial were disregarded, showing lack of chemotherapy benefit, and instead modelled a chemotherapy benefit from an indirect source. For patients with MammaPrint Low Risk, it is currently assumed a chemotherapy benefit of HR: 0.71, a 29% risk reduction, while this Hazard Ratio should have been set at 1.00 based on MINDACT results, displayed in the table below.<br><br>
<table>
<tr><th colspan="6">C-High/G-low HR+/HER2-/LN+ population</th></tr>
<tr><th colspan="3">5-year DMFI</th><th colspan="2">8-year DMFI</th><th>DMFI</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>96.3%</td><td rowspan="2">0.1%</td><td>92.3%</td><td rowspan="2">1.4%</td><td>0.85 (0.50-1.44)</td></tr>
<tr><td>no ACT</td><td>96.2%</td><td>90.9%</td><td>Reference</td></tr>
<tr><th colspan="3">5-year Overall Survival</th><th colspan="2">8-year Overall Survival</th><th>Overall Survival</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>98.4%</td><td rowspan="2">-0.4%</td><td>95.5%</td><td rowspan="2">0.6%</td><td>0.95 (0.50-1.79)</td></tr>
<tr><td>no ACT</td><td>98.8%</td><td>94.9%</td><td>Reference</td></tr>
</table>
<br>
<table>
<tr><th colspan="6">C-High/G-low HR+/HER2-/LN+ >50 population</th></tr>
<tr><th colspan="3">5-year DMFI</th><th colspan="2">8-year DMFI</th><th>DMFI</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>96.0%</td><td rowspan="2">-0.7%</td><td>91.4%</td><td rowspan="2">0.2%</td><td>0.88 (0.46-1.68)</td></tr>
<tr><td>no ACT</td><td>96.7%</td><td>91.2%</td><td>Reference</td></tr>
<tr><th colspan="3">5-year Overall Survival</th><th colspan="2">8-year Overall Survival</th><th>Overall Survival</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>97.5%</td><td rowspan="2">-0.6%</td><td>94.8%</td><td rowspan="2">-1.1%</td><td>0.99 (0.45-2.18)</td></tr>
<tr><td>no ACT</td><td>98.1%</td><td>95.9%</td><td>Reference</td></tr>
</table>
<br>The data displayed above from the MINDACT trial show that there is no benefit of chemotherapy in the C-High/G-low HR+/HER2-LN+ population, with even clearer results for the C-High/G-low HR+/HER2-/LN+ >50 population.<br>For the group not stratified by age, the EAG had access to the 5-year and 8-year results for DMFI and OS, as this data was published in Table S12 of Piccart et al. (2021). The Adjusted Hazard Ratio for DMFI, that shows a non-significant HR of 0.85 was provided by Agendia in Company’s response to the EAG on June 26 and therefore available for the EAG for modelling purposes. The observed not clinically relevant benefit of chemotherapy (0.1% DMFI difference at 5-years) that does not translate to a clinically relevant improved Overall Survival (0.6%), together with a non-significant HR for both DMFI and OS, indicate that there is no</td><td>The EAG does not believe that there is sufficient evidence to support an assumption of predictive benefit for MammaPrint. As such, a constant HR for CET vs ET is applied to both the MammaPrint low and MammaPrint high groups in the model, based on the relative effect estimated from the EBCTCG meta-analysis. DSAs have been conducted looking at higher and lower HRs (HRs ranging from 0.60 to 0.80) and these do not affect the conclusions of the economic analysis for MammaPrint. We have also run an additional analysis in which HRs of 0.85 are applied in both genomic risk states – this results in a South-West quadrant ICER for MammaPrint vs current decision-making of around £3,000 saved per QALY lost.<br><br>The EAG does not consider the company’s</td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | benefit of chemotherapy for this group, and therefore justifies modelling the HR for CT in the G-low group at 1.00.<br><br>The data of the C-High/G-low HR+/HER2-/LN+ >50 population has not been published in the MINDACT article but could have been provided by Agendia if requested by the EAG. Since later in the report, it is mentioned that it would be the desire to have access to this data to explore such a scenario, the results for the C-High/G-low HR+/HER2-/LN+ >50 population are shared and supports the modelling assumption of a HR for CT of 1.00 in the G-low group.<br><br>**Lastly, with a Relative Risk Reduction of chemotherapy of 0.97 and 1.21 in the all ages and >50 population, respectively, for DMFI in at 5-years (when chemotherapy is supposed to have the most benefit), we believe the data adequately supports the lack of chemotherapy benefit in this group and therefore, the modelling assumption of 1.00 in the G-low group.**<br><br>Adapting this model input completely changes results in a manner where MammaPrint dominates the current decision-making, regardless of if an assumption of positive predictive benefit in G-high holds. Agendia has explored the model outcomes when correcting for this Hazard Ratio assumption in 32 model alterations. See Section B. | assumption that a non-significant HR is equivalent to "no effect" to be reasonable. This issue is discussed in Section 4.2.2.2 of the report and for brevity this is not repeated here. |
| Agendia NV | 26 | 17 | 1.5 | "for each individual test, risk classification probabilities and DRFI estimates have been taken from same source, which avoids the potential for spectrum bias…"<br><br>The report writes that the DRFI estimates within each scenario have been taken from the same source, albeit not the case for BC1 and BC2. In these cases, the Oncotype DX RS>25 had to be informed by data from TransATAC, as written on page 118 of the report. For BC1 and BC2 there is spectrum bias. In addition, Oncotype DX high risk from TransATAC is modelled as RS>25, even though the prognostic results from TransATAC were based on the old RS cut-offs, where high risk was defined as RS≥31. With the change in cut-off from RS≥31 to RS>25, an improved prognosis for the new high-risk group should be expected. This means that with the prognostic input parameter from TransATAC, any chemotherapy benefit modelled using a Hazard Ratio in the G-high group of Oncotype DX, with 0.59 from Albain or 0.71 from EBCTCG will always overestimate chemotherapy benefit in this group, likely introducing an additional source of bias for BC1 and BC2. | Strictly, the Agendia is correct in that the probability of being RS 26+ in BC1 and BC2 is based on RxPONDER, whereas the DRFI outcomes for ET monotherapy within this same group are drawn from TransATAC. **The words "where data permit" will be added to this sentence in the updated EAG report.** The limitation regarding the differences in cut-offs between the studies is already explained in the EAG report. |

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td>Agendia NV</td><td>27</td><td>17</td><td>1.5</td><td>The report marks that it is a weakness that the analyses of MammaPrint reflect a clinical high-risk population identified using AOL, that no longer exists. Al though AOL indeed no longer exists, this is no weakness in our view, as the clinical risk algorithm is 100% concordant with the Nottingham Prognostic Index for patients with N1 disease. As can be seen in the tables below, the exact same clinical risk criteria are necessary to end up in either the C-High by AOL or >3.4 by NPI group and would subsequently be considered for chemotherapy. Since there is 100% concordance with NPI, that is used in the UK, we are of the opinion that this weakness should be stricken from the report, as the weakness is non-existent.<br><br>
NPI Algorithm:<br>NPI = 0.2 x tumor size in cm + nodal stage + tumor grade<br><br>
Nodal stage:<br>Stage A, Nodes absent = 1 point<br>Stage B, 1-3 nodes = 2 points<br>Stage C, >3 nodes = 3 points<br><br>
Tumor grade:<br>Grade 1 = 1 point<br>Grade 2 = 2 points<br>Grade 3 = 3 points<br><br>
Threshold to be NPI >3.4 for patients with Node Positive disease
<table>
<tr><th>For a patient with:</th><th>Tumour size</th><th>Nodal stage</th><th>Grade</th><th>NPI Score</th></tr>
<tr><td>Node positive & Grade 1</td><td>0.2 x 2.1 = 0.42 points</td><td>B = 2 points</td><td>1 = 1 point</td><td>3.42</td></tr>
<tr><td>Node positive & Grade 2</td><td>any size</td><td>B = 2 points</td><td>2 = 2 points</td><td>>4</td></tr>
<tr><td>Node positive & Grade 3</td><td>any size</td><td>B = 2 points</td><td>3 = 3 points</td><td>>5</td></tr>
</table>
Modified Adjuvant Online (mAOL)
<table>
<tr><th>Clinical subtype</th><th>Grade</th><th>Node status</th><th>Tumour size</th><th>mAOL clinical risk</th></tr>
<tr><td rowspan="4">HR+/HER2-</td><td rowspan="2">Grade 1</td><td rowspan="2">1-3 positive nodes</td><td>≤2 cm</td><td>C-Low</td></tr>
<tr><td>2.1cm – 5 cm</td><td>C-High</td></tr>
<tr><td>Grade 2</td><td>1-3 positive nodes</td><td>any size</td><td>C-High</td></tr>
<tr><td>Grade 3</td><td>1-3 positive nodes</td><td>any size</td><td>C-High</td></tr>
</table>
HR = Hormone Receptor, HER2 = Human Epidermal growth factor Receptor 2</td><td>The EAG agrees. This point will be removed from the EAG report.</td></tr>
<tr><td>Agendia NV</td><td>28</td><td>18</td><td>1.7</td><td>The EAG states that there is ‘some uncertainty’ in the assumption of predictive benefit for Oncotype DX and indicates at the same time that the assumption strongly influences the conclusions of the economic analysis.<br><br>As explained in Comment #1, the degree of uncertainty is considerably higher than currently reported for this assumption. This directly results in an unjustifiable advantage for Oncotype DX in this assessment as opposed to other GEP-tests. In particular unjustifiable versus MammaPrint, as MammaPrint also has evidence that would support to explore a scenario where MammaPrint is predictive for chemotherapy benefit in Genomic High risk (Comment #3).</td><td>This has been discussed in responses above.</td></tr>
</table>

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| Agendia NV | 29 | 41 | 3.1.2.1 | "studies included LN+ patients but >20% had >3 positive nodes; these studies were included to ensure inclusion of sufficient relevant evidence, but these limitations were noted."<br><br>As explained in Comment #1, literature currently considers N0 tumours and N1 tumours to be biologically more similar to each other, as compared with N1 and N2/N3 tumours. This is also suggested by the TNM stages, where patients generally considered for GEP-tests with N0 and N1 disease are both classified as Stage II breast cancer, while N2/N3 disease is regarded as Stage 3 breast cancer. For any analysis that is inclusive of a considerable number of patients with N2/N3 disease, the chemotherapy predictiveness results in genomic high risk are likely to be highly biased by the very poor prognosis of combined very high clinical risk and genomic risk. Conversely, any analysis that is inclusive of a considerable number of patients with N0 may actually underestimate such an effect.<br><br>As done for the mixed N1/N2/N3 cohorts, justification could be made toward the inclusion of mixed N0/N1 cohort to ensure inclusion of sufficient evidence, particularly considering the evolution in the breast cancer landscape toward associating N1 disease more with N0 rather than N2. As described in Comment #1, the decision to include N2/N3 likely create a bias and could be detrimental to patients and healthcare system in overestimating the benefits of chemotherapy in a modelled N1 only population.<br><br>It would have been better justifiable to have made exceptions towards cohorts that were a N0 and N1 mixed cohort to ensure inclusion of sufficient evidence, and note this limitation, instead of making an exception for N2/N3 disease that almost without a doubt biases a full lymph node positive cohort. | The EAG report excluded LN0 populations in line with the NICE scope. This appraisal is specifically addressing LN+ patients. LN0 patients are already covered in DG34.<br><br>As noted in the quote, evidence in patients with >3 positive nodes was included in the report to ensure inclusion of sufficient relevant evidence. However, studies in the correct population (1 to 3 positive nodes) were prioritised for inclusion in the model. |
| Agendia NV | 30 | 45 | 3.2.2 | The EAG report marks study of Mook et al. 2009 as a 'reanalyses of trials of chemotherapy versus no chemotherapy.<br><br>This is a misinterpretation of the work performed by Mook and colleagues. They performed an analysis of two selected cohorts from two institutions; NKI-AVL, consecutive series and EIO, consecutive series. Patients in these consecutive series were not randomised to chemotherapy versus no chemotherapy, as now is suggested in the report. | The EAG report Section 3.2.2 states that Mook et al. 2009 is a reanalysis of a cohort. Randomisation is not mentioned here.<br><br>**The EAG report will be amended so that the header for this bullet will read "Reanalysis of studies" rather than "reanalysis of trials". The description of Mook 2009 in the same section will be amended to "reanalysis** |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | | **of two cohorts" rather than "reanalysis of a cohort".** |
| Agendia NV | 31 | 46 | 3.2.2 | It is stated that no HRs or significance tests were reported for the difference in outcomes between test risk groups for MammaPrint.<br><br>The prospective, randomised, MINDACT trial proved MammaPrint's independent prognostic ability in a multivariate analysis for Distant Metastasis Free Survival after correcting for clinical factors including lymph node status, age, and chemotherapy use. Results are presented in Table S15 of Piccart et al. 2021 and are applicable to MammaPrint's prognostic ability in HR+/HER2-/N1 breast cancer, regardless.<br><br>**MammaPrint High Risk vs Low Risk – Hazard Ratio: 2.13 (P-value < 0.0001)**<br><br>*Piccart M, van 't Veer LJ, Poncet C et al. 70-gene signature as an aid to treatment decision in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. Lancet Oncology 2021; 22(4), 476-88*<br><br>MINDACT reported on the prognostic ability of MammaPrint in a manner where results are applicable to patients with HR+/HER2-/N1 disease. As such, the aforementioned and current statement is unfounded and would warrant a correction to recognize that MammaPrint was significantly prognostic for Distant Metastasis Free Survival after adjusting for clinical factors. | This comment is a repetition of comment #22. No further response from the EAG is required. |
| Agendia NV | 32 | 46 | 3.2.2 | The EAG notes that no data were presented for the MammaPrint high-risk group, and no interaction tests were conducted, and that it was therefore not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit.<br><br>This has indeed not been published in the Piccart et al 2021 article, because a treatment randomisation of CT vs no CT was not performed in C-high/G-high, because this would be considered unethical as also noted by the EAG.<br><br>However, what is available is data from an analysis based on the MINDACT 2016 data cut-off that looked at the event rate of C-high/G-high HR+/HER2- patients based on actual treatment received. Of 831 patients with a C-high/G-high risk, 41 patients in MINDACT did not receive chemotherapy. A multivariate analysis of distant recurrence free interval controlling for clinical risk features was performed and showed that chemotherapy was significantly associated with a 79% risk reduction (HR 0.21 [0.10-0.46], p < 0.001). The formal test of interaction between chemotherapy and genomic risk score was statistically significant (P interaction = 0.013). | Please refer to response to comment 23. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | In the Company's response of May 30th, Agendia had provided the EAG with evidence this evidence to further support the evidence of positive predictive value of the benefit of chemotherapy in MammaPrint high risk (also summarized in Comment #3). However, the EAG did not include any of the submitted data in their report. | |
| Agendia NV | 33 | 51 | 3.4.4 | "Within the range RS 0-25, Oncotype DX was significantly prognostic for 5-year IDFS after adjusting for clinical factors, both in the overall population (HR per unit-RS 1.05; 95% confidence interval [CI] 1.04 to 1.07; *p*<0.001) and in the premenopausal and post-menopausal subgroups…"<br><br>The EAG recognizes the multivariate prognostic analysis of Oncotype DX from RxPONDER in this segment, that serves a common goal as the multivariate prognostic analysis of MammaPrint in MINDACT as pointed out in Comments #2 and #11. | No response required. |
| Agendia NV | 34 | 52 | 3.4.5 | The difficulty to assess the prognostic ability of MammaPrint in MINDACT is mentioned again.<br><br>The prospective, randomised, MINDACT trial proved MammaPrint's independent prognostic ability in a multivariate analysis for Distant Metastasis Free Survival after correcting for clinical factors, including lymph node status, age, and chemotherapy. Results are presented in Table S15 of Piccart et al. 2021 and are applicable to MammaPrint's prognostic ability in HR+/HER2-/N1 breast cancer, regardless of age.<br><br>**MammaPrint High Risk vs Low Risk – Hazard Ratio: 2.13 (P-value < 0.0001)**<br><br>*Piccart M, van 't Veer LJ, Poncet C et al. 70-gene signature as an aid to treatment decision in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. Lancet Oncology 2021; 22(4), 476-88*<br><br>MINDACT reported on the prognostic ability of MammaPrint in a manner where results are applicable to patients with HR+/HER2-/N1 disease. As such, the aforementioned and current statement is unfounded and would warrant a correction to recognize that MammaPrint was significantly prognostic for Distant Metastasis Free Survival after adjusting for clinical factors.<br><br>**Especially since the multivariate analysis of Oncotype DX from RxPONDER is recognized in the section prior to 3.4.4, and the multivariate analysis of MammaPrint MINDACT is currently being neglected, see Comment #13.** | This comment is a repetition of comments #22 and #31. No further response from the EAG is required. |
| Agendia NV | 35 | 55 | 3.4.8 | The conclusion about MammaPrint is non-inclusive of the data provided in Table S15 of the MINDACT trial, as also commented in Comments #2, #11, #14. Acknowledgement of this data should also lead to a different conclusion in this segment. | This comment is referring to the same HR as comment #22, #31 and |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | **MammaPrint High Risk vs Low Risk – Hazard Ratio: 2.13 (P-value < 0.0001) for DMFS** | #34. No further response from the EAG is required. |
| Agendia NV | 36 | 57 | 3.5.1 | The predictive value of GEP-tests is two-sided; negative predictive value and positive predictive value. MINDACT provided data for the negative predictive value of MammaPrint in the mAOL high-risk / MP low-risk group that was randomised to CT or no CT.<br><br>We believe that a clear distinction between Negative Predictive Value and Positive Predictive Value should be indicated in the report rather than stating that the predictive value of MammaPrint could not be determined from MINDACT. This is important because the Negative Predictive value of MammaPrint is certainly determined and demonstrated for MammaPrint in MINDACT (see Comment #5).<br><br>The positive predictive value for mAOL high-risk / MP high risk was indeed not assessed in a randomised setting, as this has always been considered to be unethical. However, what is available is data from an analysis based on the MINDACT 2016 data cut-off that looked at the event rate of C-high/G-high HR+/HER2- patients based on actual treatment received. A multivariate analysis of distant recurrence free interval controlling for clinical risk features was performed and showed that chemotherapy was significantly associated with a 79% risk reduction (HR 0.21 [0.10-0.46], p < 0.001).The formal test of interaction between chemotherapy and genomic risk score was statistically significant (P interaction = 0.013). Data made available to EAG in Company's response on May 30th. Also, see Comment #3. | Discussions of predictive ability in the EAG report relate to the ability of the test to predict differences in chemotherapy effect for different test risk groups. They do not relate to negative and positive predictive value.<br><br>Regarding the analysis of chemotherapy vs. no chemotherapy in C-high/G-high patients (company response 30 May), this analysis was not restricted to LN+ patients (over half were LN0). In addition, only 41 patients had no chemotherapy and this likely represented a biased selection, since most patients in this group were indicated for chemotherapy. Therefore the EAG does not believe this is a robust analysis for this appraisal. |
| Agendia NV | 37 | 99 | 4.2.2.1 | Throughout the report the efficacy input for MammaPrint is being referred to as DRFI (Distant Recurrence Free Interval), while in fact the efficacy input for MammaPrint is Distant Metastasis Free Interval (DMFI). DMFI is a more conservative endpoint than DRFI.<br><br>DRFI events are Distant Recurrence and Death due to breast cancer. | The EAG believes it is unlikely that the difference between endpoints would alter the |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | DMFI events are Distant Recurrence, Death due to breast cancer and Death due to unknown cause.<br><br>Al though this is a subtle difference, it should be noted that DMFI is more conservative and in principle has more event rates than DRFI, as it counts an additional type of event. As a result, MammaPrint has a slight disadvantage in the model, versus scenarios of other tests that do in fact report DRFI. | conclusions of the economic analysis. |
| Agendia NV | 38 | 102 | 4.2.2.2 | "(1) Relevance of the model population to the decision problem set out in the final NICE scope".<br><br>To understand the background:<br><br>In Agendia's opinion, the NICE scope for the update of the DG34 should not have been a partial one for lymph node positive alone, but a full one that would also consider new evidence for node negative disease.<br><br>Significant new evidence had become available since DG34 for lymph node negative disease, amongst them the long-term follow-up from MINDACT. These data have a high probability of changing the conclusions reached by NICE in DG34 about MammaPrint. For this reason, Agendia had been preparing for the DG34 update, hence the current model models the full C-high MINDACT population.<br><br>The scenario analysis 5, specific for lymph node positive disease, had been later added once NICE made clear that they did not intend to change the scope of the diagnostic guidance update, to make available a lymph node positive scenario to the EAG. | No EAG response required regarding the appropriateness of the scope. Scenario 5 is the most relevant analysis presented in the Cytel CEA report. |
| Agendia NV | 39 | 102 | 4.2.2.2 | "problematic as they rest on two strong assumptions: (i) that NPI>3.4 is exactly equivalent to mAOL high-risk and (ii) that the characteristics of the patient populations enrolled in TransATAC19 and MINDACT29 are identical with respect to prognostic factors and treatment effect modifiers."<br>With regards to the first, as described in Comment #7, there is a 100% concordance between AOL and NPI. This has been inaccurately labelled as a 'problematic strong assumption.'.<br><br>With regards to the second, we indeed recognized this as a key limitation in the model, as reported in the company submission:<br><br>*"The key limitation of this analysis was the lack of head-to-head comparison of MammaPrint versus other gene profiling tests. As such indirect trial comparison (ITC) was necessary but was subject to bias associated with non-randomised treatment groups. To mitigate this risk, a series of five key scenario analyses additional to the base case to characterize the variability and investigate trends."*<br><br>The results are indeed modelled best when MammaPrint is compared to "usual care', treatment only using clinical factors. These results can be interpreted without the caveats needed for the comparisons to the other test from TransATAC. However, what should be noted is that MINDACT N0/N1 is overall a higher | **The EAG agrees that the comment about mAOL can be removed. The company has agreed that the second assumption is subject to uncertainty. The EAG will amend the wording of the report to read *"problematic as they rest on a strong assumption that the characteristics of the patient populations enrolled in TransATAC[19] and MINDACT[29] are*** |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>clinical risk cohort than TransATAC N0/N1, making any differential prognostic performance in the cohorts be in the disadvantage of MammaPrint. Yet, the MammaPrint arm, modelled with higher clinical risk has shown to be the most cost-effective strategy in all scenarios.</td><td>identical with respect to prognostic factors and treatment effect modifiers."</td></tr>
<tr><td>Agendia NV</td><td>40</td><td>103</td><td>4.2.2.2</td><td>"(3) Relevance of the usual care comparator (mAOL)"<br><br>The evidence submission of Agendia was inclusive of a table showing high concordance (>96%) of mAOL and NPI in the full HR+/HER2-/LN-&LN+ cohort. Specific to the HR+/HER2-/LN+ group, as described in Comment #7, there is a 100% concordance between AOL and NPI.<br><br>Table 4. Agreement between modified Adjuvant Online! (2-level) and Nottingham Prognostic Index (2-level) clinical risk classification tests.<br><table><tr><th>Test</th><th>Clinical risk</th><th colspan="2">mAOL (2-Level)</th><th>Total</th></tr><tr><td></td><td></td><td>mAOL low</td><td>mAOL high</td><td></td></tr><tr><td rowspan="2">NPI (2-level)</td><td>NPI ≤ 3.4</td><td>1976 (97%/96%)</td><td>71 (3%/4%)</td><td>2047 (52%)</td></tr><tr><td>NPI > 3.4</td><td>74 (4%/4%)</td><td>1846 (96%/96%)</td><td>1920 (48%)</td></tr><tr><td>Total</td><td></td><td>2050 (52%)</td><td>1917 (48%)</td><td>3967 (100%)</td></tr></table>Abbreviations: mAOL, modified Adjuvant Online!; NPI, Nottingham Prognostic Index<br>Notes: Cell percentages are ordered Row / Column; NPI (2-level) stratification includes patients LN+ (1-3, nodes).<br><br>Modified adjuvant online (mAOL) is not a web-based tool, rather it is a classification algorithm that can still be used today and is published in Cardoso et al. 2016 in Table S13 and Piccart et al. 2021 in Table S1. Alternatively, the NPI score calculation can also be used to identify the same patients as C-high or C-high in the HR+/HER2-/LN+ cohort.<br><br>The concerns raised about uncertainty in interpreting the model, difficulty in identifying MINDACT risk groups, and the mention of DRFI (which should be DMFI) in this section of the report, as well as other parts of the report, have all been addressed with supporting evidence, providing a more well-rounded perspective on the presented data and assumption, together constituting ground for reconsideration of these concerns.</td><td>The EAG will amend this point in the report to state that mAOL is concordant with NPI in the HR+/HER2 LN+ group, and that whilst mAOL itself is not used in practice, it is equivalent to NPI within this subgroup.</td></tr>
</table>

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>3. Clinical risk assessment according to modified Adjuvant!Online<br>Table S 1: Classification of patients according to clinical risk assessment by the modified version of Adjuvant! Online
<table>
<tr><th>ER status</th><th>HER2 status</th><th>Grade</th><th>Nodal status</th><th>Tumour Size</th><th>Clinical risk in MINDACT</th></tr>
<tr><td rowspan="18">ER positive</td><td rowspan="12">HER 2 negative</td><td rowspan="4">well differentiated</td><td rowspan="2">N-</td><td>≤ 3 cm</td><td>C-Low</td></tr>
<tr><td>3.1-5 cm</td><td>C-High</td></tr>
<tr><td rowspan="2">1-3 positive nodes</td><td>≤ 2 cm</td><td>C-Low</td></tr>
<tr><td>2.1-5 cm</td><td>C-High</td></tr>
<tr><td rowspan="3">moderately differentiated</td><td rowspan="2">N-</td><td>≤ 2 cm</td><td>C-Low</td></tr>
<tr><td>2.1-5 cm</td><td>C-High</td></tr>
<tr><td>1-3 positive nodes</td><td>Any size</td><td>C-High</td></tr>
<tr><td rowspan="3">poorly differentiated or undifferentiated</td><td rowspan="2">N-</td><td>≤ 1 cm</td><td>C-Low</td></tr>
<tr><td>1.1-5 cm</td><td>C-High</td></tr>
<tr><td>1-3 positive nodes</td><td>Any size</td><td>C-High</td></tr>
<tr><td rowspan="6">HER2 positive</td><td rowspan="3">well differentiated OR moderately differentiated</td><td rowspan="2">N-</td><td>≤ 2 cm</td><td>C-Low</td></tr>
<tr><td>2.1-5 cm</td><td>C-High</td></tr>
<tr><td>1-3 positive nodes</td><td>Any size</td><td>C-High</td></tr>
<tr><td rowspan="3">poorly differentiated or undifferentiated</td><td rowspan="2">N-</td><td>≤ 1 cm</td><td>C-Low</td></tr>
<tr><td>1.1-5 cm</td><td>C-High</td></tr>
<tr><td>1-3 positive nodes</td><td>Any size</td><td>C-High</td></tr>
</table></td><td></td></tr>
<tr><td>Agendia NV</td><td>41</td><td>103</td><td>4.2.2.2</td><td>“(4) Inappropriate assumption that some women in the MammaPrint group will not receive the test”<br><br>To allow comparison between MammaPrint and the other tumour profiling tests, this assumption was necessary to make, and is as such not inappropriate.<br><br>The node-positive cohort from TransATAC was in our understanding not limited to NPI > 3.4, with no data on proportions of NPI ≤3.4 and NPI >3.4 available. This means that the group modelled for other tumour profiling tests had a lower aggregate clinical risk than the population modelled for MammaPrint based on MINDACT. To account for this, the N+ clin low population from MINDACT had to be incorporated to have a balanced clinical risk group in each comparative arms, assuming a comparable clinical risk distribution among MINDACT and TransATAC.<br><br>No further bias can be attributed to this methodological assumption, as the efficacy input (DMFI transition probability) for the usual care arm accounts for the prognosis for the C-low/G-low group. The transition probability for usual care is informed with the weighted C-low and C-high prognosis without chemo, 16% and 84% proportions from MINDACT, respectively.</td><td>The EAG disagrees with this approach and considers it more reasonable to:<br>- Focus on the direct comparison of MammaPrint versus usual care, based on MINDACT in clinical high-risk patients<br>- Assume that patients in the test group receive the test and patients in the usual care group do not receive the test.</td></tr>
</table>

# Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node positive early breast cancer

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td>Agendia NV</td><td>42</td><td>105</td><td>4.2.2.2</td><td>“As such, the EAG believes that the company’s interpretation of the results of their IPD analysis are flawed and that it is more reasonable to apply the same treatment effect estimate for chemotherapy to both the MammaPrint low-risk and MammaPrint high-risk groups.”<br><br>While we understand why questions are raised about applying a differential Hazard Ratio for chemotherapy in the genomic high-risk population, we want to repeat the importance to acknowledge the results of MINDACT, which demonstrated the ability to identify patients at genomic low risk who could safely forgo chemotherapy. As noted in Comment #5, it's essential to consider the findings of MINDACT, which indicated a very marginal chemotherapy benefit in the G-low group (Relative Risk Reduction of 0.97 and 1.21, in the all ages and >50 population, respectively, for DMFI in the first five-years, when chemotherapy is supposed to have the most benefit), before deciding to opt to apply a chemotherapy hazard ratio of 0.71 uniformly. We hope that the table provided below, displaying the MINDACT hazard ratio, provides enough evidence for NICE/EAG to reconsider this assumption.<br><br>
<table>
<tr><th colspan="6">C-High/G-low HR+/HER2-/LN+ population</th></tr>
<tr><th></th><th colspan="2">5-year DMFI</th><th colspan="2">8-year DMFI</th><th>DMFI</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>96.3%</td><td rowspan="2">0.1%</td><td>92.3%</td><td rowspan="2">1.4%</td><td>0.85 (0.50-1.44)</td></tr>
<tr><td>no ACT</td><td>96.2%</td><td>90.9%</td><td>Reference</td></tr>
<tr><th></th><th colspan="2">5-year Overall Survival</th><th colspan="2">8-year Overall Survival</th><th>Overall Survival</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>98.4%</td><td rowspan="2">-0.4%</td><td>95.5%</td><td rowspan="2">0.6%</td><td>0.95 (0.50-1.79)</td></tr>
<tr><td>no ACT</td><td>98.8%</td><td>94.9%</td><td>Reference</td></tr>
</table>
<br>
<table>
<tr><th colspan="6">C-High/G-low HR+/HER2-/LN+ >50 population</th></tr>
<tr><th></th><th colspan="2">5-year DMFI</th><th colspan="2">8-year DMFI</th><th>DMFI</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>96.0%</td><td rowspan="2">-0.7%</td><td>91.4%</td><td rowspan="2">0.2%</td><td>0.88 (0.46-1.68)</td></tr>
<tr><td>no ACT</td><td>96.7%</td><td>91.2%</td><td>Reference</td></tr>
<tr><th></th><th colspan="2">5-year Overall Survival</th><th colspan="2">8-year Overall Survival</th><th>Overall Survival</th></tr>
<tr><th></th><th>Survival estimates</th><th>Absolute diff.</th><th>Survival estimates</th><th>Absolute diff.</th><th>Adjusted Hazard Ratio</th></tr>
<tr><td>ACT</td><td>97.5%</td><td rowspan="2">-0.6%</td><td>94.8%</td><td rowspan="2">-1.1%</td><td>0.99 (0.45-2.18)</td></tr>
<tr><td>no ACT</td><td>98.1%</td><td>95.9%</td><td>Reference</td></tr>
</table>
<br>Adapting this model input completely changes results in a manner where MammaPrint dominates the current decision-making, regardless of if an assumption of positive predictive benefit in G-high holds. Agendia has explored the model outcomes when correcting for this Hazard Ratio assumption in 32 model alterations. See Section B.</td><td>As discussed in the EAG report, the EAG does not consider there to be sufficient evidence to support this assumption, and the EAG disagrees with the company’s interpretation of a non-significant HR as being equal to “no effect.”<br><br>The EAG believes it is more appropriate to consider longer-term data where available and notes that the HR takes account of the difference in risk over the whole observed period, rather than at a specific timepoint.</td></tr>
<tr><td>Agendia NV</td><td>43</td><td>105</td><td>4.2.2.2</td><td>“QALY losses associated with adjuvant chemotherapy appear to be based on assumptions rather than empirical evidence.”<br><br>QALY losses associated with adjuvant chemotherapy are not based on assumptions, they are based on the results of Lidgren et al. 2007. This study reported on utilities for “first year after primary breast cancer” and for “second and following years after primary breast cancer”. The study had found that patients on</td><td>This quote is referring to the utility values used in 2 other published models (Wong and Vanderlaan), not the values used in the Agendia model.</td></tr>
</table>

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>chemotherapy had a lower associated utility in the first year, and that this increases again after one year, but still at differential level as patients only treated with endocrine therapy.<br><br>Swedish researchers have studied the Health-Related Quality of Life (HRQoL), scaled 0 to 1, in different disease states of 361 consecutive breast cancer patients that were treated at Karolinska University hospital Solna. The study showed that patients that were on chemotherapy had a worse HRQoL in both “State P” (first year after primary breast cancer) and “State S” (second year and following years after diagnosis) than patients that did not receive chemotherapy. The difference in QoL was 0.124 in “State P” and 0.081 in “State S”. Patients receiving chemotherapy in the first year after primary breast cancer even showed a worse HRQoL than those that experienced metastatic disease. In the table below the health utilities of the health states are displayed.<br><br>Different breast cancer health states and associated utilities<br><table><tr><th>Breast cancer health state</th><th>Associated utility</th></tr><tr><td>State P – 1st year after primary breast cancer, receiving endocrine therapy (No CT)</td><td>0.744</td></tr><tr><td>State P – 1st year after primary breast cancer, receiving chemotherapy</td><td>0.620</td></tr><tr><td>State S – 2nd year and following after diagnosis, receiving endocrine therapy (No CT)</td><td>0.824</td></tr><tr><td>State S – 2nd year and following after diagnosis, receiving chemotherapy</td><td>0.743</td></tr><tr><td>State M – Metastatic disease</td><td>0.685</td></tr></table><br>Although the paper did not report State S for the CT group directly, the paper did report on the utility value for the full group and the group that received endocrine therapy alone within State S. As a result, by having two components of the equation, it is possible to calculate the utility value for State S on chemotherapy. Lidgren et al. reported that the mean EQ-5D value for State S was 0.779 (N = 177) and that patients in State S receiving adjuvant hormone therapy had an EQ-5D value of 0.824 (N = 79). Under the assumption that the remainder of the group received chemotherapy, one can derive the utility value for patients in State S receiving chemotherapy.<br><br>State S on CT = (177 * 0.779 – 79 * 0.824) / (177 – 79)<br>State S on CT = 72.787 / 98<br>State S on CT = 0.743<br><br>Although we agree it would have been better if “State S on CT” was reported in the Lidgren et al. paper, the absence of it leaves this calculation to be the best source at hand and which is plausible considering the published difference in the utility values in the first year after primary breast cancer (0.744 receiving endocrine therapy versus 0.620 receiving CT).<br><br>While the EAG might believe that Lidgren could overestimate the negative effect of chemotherapy on quality of life, we believe that the assumption made by EAG is likely underestimating the negative effect of</td><td>The EAG report comments that the assumed QALY loss in the Agendia model is much higher than most of the other models included in the review of existing models. This is partly driven by the 3-year duration over which the disutility is applied.<br><br>The EAG’s model applies the same QALY loss as that used in DG34 and as that used in the Exact Sciences model for the current appraisal. In the absence of alternative evidence, there is little justification for applying a different value. However, to explore uncertainty around this parameter, DSAs have been conducted whereby the QALY loss is doubled (DSA15) and tripled (DSA16). Results are presented in Table 43 of the EAG report and are not repeated here.</td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | chemotherapy on quality of life. Comment #32 provides more detail to why we believe the EAG assumption underestimates the negative impact. In summary physicians (who see patients treated with chemotherapy) and other literature seem to agree with the fact that a 14-day loss (derived from the Campbell assumption used in DG34 and current EAG model) is vastly underestimating the 1-year QALY decrement of chemotherapy. | |
| Agendia NV | 44 | 108/109 | 4.2.2.3 | "The EAG undertook an additional analysis which attempts to address six of the issues identified in the EAG's critique… MammaPrint is more expensive and less effective than usual care; hence, it is dominated."<br><br>Agendia has looked into the six points of critique and implemented the same changes to mimic the additional analysis performed by EAG. The results presented in the report are not representative of the true model results when one wants to make the six changes proposed by the EAG. Unfortunately, this has been a consequence of a less user-friendly model built than anticipated, but the Hazard Ratio's cannot be simply modified in the model without changing the efficacy (DMFI) inputs as well, as now has been done when attempting to correct for issue (ii).<br><br>The "Probability of distant recurrence free interval per cycle" for genomic high risk is displaying an untreated (no CT) DMFI for CHGH patients, estimating by dividing the DMFI probability by the estimated CT relative risk, which was set at 0.28 in the base case of the LN+ scenario. When a user changes the Hazard Ratio for G-High for the "Adjuvant chemotherapy treatment effect on distant recurrence" from 0.28 to 0.76, the untreated DMFI in the "Probability of distant recurrence free interval per cycle" is not automatically changed with the newly applied Hazard Ratio. As a result, changing the HR to 0.76 for this group results in modelling an extremely poor prognosis for patients with G-high risk which can no longer be 'adequately improved' with the HR of 0.28, but remains low with the newly applied HR 0.76.<br><br>The correct way to make this adjustment, is to not only change the HR to 0.76, but to change on top of that the "Probability of distant recurrence free interval per cycle" for G-high from 0.920 to 0.9822 (indicated in blue), and the comparative arms from 0.981 to 0.9908 (indicated in orange).<br><br>Probability of distant recurrence free interval per cycle<br>Tumor profiling test result<br>Clinical risk \| Tumor profile test \| Clinical risk \| Low \| Intermediate \| High \| Notes \| Source<br>MammaPrint<br>mAOL Low \| MammaPrint \| 0.000 \| MINDACT data<br>mAOL High \| MammaPrint \| 0.991 \| 0.994 \| 0.982 \| ‡ \| MINDACT data; ECBTCG (2012)<br>Placeholder \| MammaPrint \| 0.000 \| MINDACT data<br>Comparators<br>mAOL Low \| Usual care \| 0.000 \| MINDACT data<br>LN+ (1-3 nodes) \| Usual care \| 0.991 \| MINDACT data; ECBTCG (2012)<br>0.000<br>0.000<br>0.000<br>Placeholder \| Usual care \| 0.000 \| MINDACT data; Harnan, S. (2017) | The EAG agrees that the Agendia model is not user-friendly. It is unclear why it should be necessary within the model to amend both the baseline DRFI risk and the relative risk in order to estimate the DRFI risk with chemotherapy. However, the EAG agrees that there is a problem in the EAG's re-analysis and notes that the company's corrected model now produces a similar model trace to the EAG's model for patients at low/high genomic risk, with/without chemotherapy.<br><br>The EAG has scrutinised the version of the company's model used to generate the results on the left. As indicated in the text, the updated Agendia results do not include corrections to issue (iv) in the EAG's critique of the original |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>When following the additional analysis undertaken by EAG in the model, correcting for the DMFI as explained above, with the exception of issue (iv), the results of the EAG additional analysis show that MammaPrint remains to be the dominating treatment strategy. As displayed in the table below, MammaPrint is cost-saving and QALY gaining.<br><br>
<table>
<tr><th>Option</th><th>LYGs</th><th>QALYs</th><th>Costs</th><th>Inc. LYGs</th><th>Inc. QALYs</th><th>Inc. Costs</th><th>ICER</th></tr>
<tr><td colspan="8">Company's Scenario Analysis 5 (LN+), other tests excluded</td></tr>
<tr><td>MammaPrint</td><td>23.45</td><td>11.57</td><td>23327</td><td>1.43</td><td>0.7</td><td>-£4,676</td><td>MammaPrint</td></tr>
<tr><td>Usual care</td><td>22.02</td><td>10.87</td><td>28003</td><td>-</td><td>-</td><td>-</td><td>dominating</td></tr>
<tr><td colspan="8">Company's Scenario Analysis 5 (LN+), other tests excluded, including EAG amendments (corrected)</td></tr>
<tr><th>Option</th><th>LYGs</th><th>QALYs</th><th>Costs</th><th>Inc. LYGs</th><th>Inc. QALYs</th><th>Inc. Costs</th><th>ICER</th></tr>
<tr><td>MammaPrint</td><td>15.56</td><td>11.78</td><td>23045</td><td>-0.04</td><td>0.03</td><td>-£1,307</td><td>MammaPrint</td></tr>
<tr><td>Usual care</td><td>15.61</td><td>11.75</td><td>24351</td><td>-</td><td>-</td><td>-</td><td>dominating</td></tr>
</table>
<br>Al though Agendia does not agree with all of the six issues identified by EAG and the proposed resolution, the results of EAG’s additional analysis suggest that MammaPrint is less expensive and more effective than usual care; hence, it is dominating.<br><br>Agendia has explored the model outcomes of the EAG de novo model in 32 model alterations that show highly consistent results that MammaPrint is a dominating treatment strategy. See Section B.</td><td>model. These errors are unequivocal can be seen in the trace worksheets, whereby several of the discounted costs are higher than their undiscounted counterparts – this is because costs and QALYs in cycle 1 are counted 1.5 cycles times in the same formulae as the discounting.. Correcting these errors alone in the company’s updated model produces the following results:<br>• Incremental LYGs remain similar at -0.04<br>• Incremental QALYs are approximately halved from 0.03 to 0.016<br>• Incremental cost savings are reduced from -£1,307 to -£26<br>• MammaPrint remains dominant, but with very small differences in costs and QALYs. Expected survival is worse in the MammaPrint group.<br><br>The EAG will amend the table in the updated NICE report to reflect</td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | | **the EAG's corrected re-analysis described in the bulletpoints above.**<br><br>The EAG notes that the corrected Agendia model includes several assumptions which differ substantially from the EAG's model, including a >3x higher disutility for chemotherapy, a lower pre-test chemotherapy probability, post-test probabilities based on a study undertaken in women without axillary lymph node involvement, and multiple differences in assumptions and evidence used to estimate the costs and health outcomes associated with downstream treatments. Overall, the EAG considers that the EAG's model represents a more suitable basis for decision-making.<br><br>The EAG disagrees with the company's 32 new analyses using the EAG's model because these all assume a predictive benefit for |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | | MammaPrint. As discussed above, the EAG did not find sufficient evidence to support this assumption. |
| Agendia NV | 45 | 110 | 4.3.1 | Unavailability of AOL is listed as a limitation to interpretation.<br><br>As described in Comment #7, there is a 100% concordance between AOL and NPI.<br>Interpretation of the economic analysis of MammaPrint is not limited but should just be interpreted as if NPI is used. | Please refer to earlier responses above. |
| Agendia NV | 46 | 118 | 4.3.3 | "Across all seven base case scenarios, the source used to inform DRFI was consistent with the source used to inform test risk classification probabilities in the previous section."<br><br>The report writes that the DRFI estimates within each scenario have been taken from the same source, albeit not the case for BC1 and BC2. In these cases, the Oncotype DX RS>25 had to be informed by data from TransATAC, as written on page 118 of the report. For BC1 and BC2 there is spectrum bias. In addition, Oncotype DX high risk from TransATAC is modelled as RS>25, even though the prognostic results from TransATAC were based on the old RS cut-offs, where high risk was defined as RS≥31. With the change in cut-off from RS≥31 to RS>25, an improved prognosis for the new high-risk group should be expected. This means that with the prognostic input parameter from TransATAC, any chemotherapy benefit modelled using a Hazard Ratio in the G-high group of Oncotype DX, with 0.59 from Albain or 0.71 from EBCTCG will always overestimate chemotherapy benefit in this group, likely introducing an additional source of bias for BC1 and BC2. | This point is repetition of comment #26. No further response required. |
| Agendia NV | 47 | 118 | 4.3.3 | Throughout the report the efficacy input for MammaPrint is being referred to as DRFI (Distant Recurrence Free Interval), while in fact the efficacy input for MammaPrint is Distant Metastasis Free Interval (DMFI). DMFI is a more conservative endpoint than DRFI.<br><br>DRFI events are Distant Recurrence and Death due to breast cancer.<br>DMFI events are Distant Recurrence, Death due to breast cancer and Death due to unknown cause.<br><br>Al though this is a subtle difference, it should be noted that DMFI is more conservative and in principle has more event rates than DRFI, as it counts an additional type of event. As a result, MammaPrint has a slight disadvantage in the model, versus scenarios of other tests that do in fact report DRFI. | This comment is repetition of comment #37. |
| Agendia NV | 48 | 119 | 4.3.3 | Table 33: BC7 – MammaPrint – Test risk "Low" – HR for DM for CT versus no CT "0.71"<br><br>This model input for MammaPrint should not have been set at 0.71, as this does not do justice to the available data from MINDACT. As explained in Comment #5, this model input should have been informed | The same HR of 0.71 was applied to all non-predictive test scenarios, based on the EBCTCG |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | with data from the MINDACT trial. The data was made available to the EAG on June 26th by Agendia, and we now provided additional data to explore a scenario that limits the population to post-menopausal patients.<br><br>**C-High/G-low HR+/HER2-/LN+ population**<br>**5-year DMFI** \| **8-year DMFI** \| **DMFI**<br>Survival estimates \| Absolute diff. \| Survival estimates \| Absolute diff. \| Adjusted Hazard Ratio<br>ACT 96.3% 0.1% 92.3% 1.4% 0.85 (0.50-1.44)<br>no ACT 96.2% 90.9% Reference<br>**5-year Overall Survival** \| **8-year Overall Survival** \| **Overall Survival**<br>Survival estimates \| Absolute diff. \| Survival estimates \| Absolute diff. \| Adjusted Hazard Ratio<br>ACT 98.4% -0.4% 95.5% 0.6% 0.95 (0.50-1.79)<br>no ACT 98.8% 94.9% Reference<br><br>**C-High/G-low HR+/HER2-/LN+ >50 population**<br>**5-year DMFI** \| **8-year DMFI** \| **DMFI**<br>Survival estimates \| Absolute diff. \| Survival estimates \| Absolute diff. \| Adjusted Hazard Ratio<br>ACT 96.0% -0.7% 91.4% 0.2% 0.88 (0.46-1.68)<br>no ACT 96.7% 91.2% Reference<br>**5-year Overall Survival** \| **8-year Overall Survival** \| **Overall Survival**<br>Survival estimates \| Absolute diff. \| Survival estimates \| Absolute diff. \| Adjusted Hazard Ratio<br>ACT 97.5% -0.6% 94.8% -1.1% 0.99 (0.45-2.18)<br>no ACT 98.1% 95.9% Reference<br><br>The data displayed above from the MINDACT trial show that there is no benefit of chemotherapy in the C-High/G-low HR+/HER2-LN+ population, with even clearer results for the C-High/G-low HR+/HER2-/LN+ >50 population.<br><br>For the group not stratified by age, the EAG had access to the 5-year and 8-year results for DMFI and OS, as this data was published in Table S12 of Piccart et al. (2021). The Adjusted Hazard Ratio for DMFI, that shows a non-significant HR of 0.85 was provided by Agendia in Company's response to the EAG on June 26 and therefore available for the EAG for modelling purposes. The observed not clinically relevant benefit of chemotherapy (0.1% DMFI difference at 5-years) that does not translate to a clinically relevant improved Overall Survival (0.6%), together with a non-significant HR for both DMFI and OS, indicate that there is no benefit of chemotherapy for this group, and therefore justifies modelling the HR for CT in the G-low group at 1.00.<br>The data of the C-High/G-low HR+/HER2-/LN+ >50 population has not been published in the MINDACT article but could have been provided by Agendia if requested by the EAG. Since later in the report, it is mentioned that it would be the desire to have access to this data to explore such a scenario, the results for the C-High/G-low HR+/HER2-/LN+ >50 population are shared and supports the modelling assumption of a HR for CT of 1.00 in the G-low group. | meta-analysis. Alternative HRs of 0.60 and 0.80 are applied in DSAs.<br><br>The table on the left has already been presented in Comment #42. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Agendia has explored the model outcomes when correcting for this Hazard Ratio assumption in 32 model alterations. See Section B. | |
| Agendia NV | 49 | 119 | 4.3.3 | Table 33: BC2 – Oncotype DX – Test risk "Low" – HR for DM for CT versus no CT "1.12"<br><br>Agendia believes one cannot model the Hazard Ratio for Oncotype DX Low risk with a Hazard Ratio of 1.12. By doing so, one is attributing a 12% relative benefit to patients not receiving chemotherapy. In this case omitting chemotherapy would yield 12% improved survival.<br><br>There is no justification for this assumption. The only thing that RxPONDER has shown is that there is no statistically significant difference in the CT vs no CT arm. One can only model the Hazard Ratio of 1.12 if the Hazard Ratio would be statistically significant, but it is not (p = 0.49). The model assumption of 1.12 unfairly biases the Oncotype DX analysis as compared to the other tests, by assuming that patients will have a 12% lower hazard for distant metastasis when ODx low risk.<br>We believe the legitimate way of modelling the observation of RxPONDER of no chemotherapy benefit, is to impute a Hazard Ratio of 1.00 in the model for Oncotype DX Low Risk in BC2.<br>This follows the same rationale as proposed in Comment #5, #22 and #28 for BC7 of MammaPrint, where the statistically insignificant Hazard ratio of the CT vs no CT arm in the C-High/G-low HR+/HER2-/LN+ population would justify modelling a Hazard Ratio of 1.00 as well. | The EAG disagrees with Agendia's view about the interpretation of non-significant HRs. Please refer to response to comment #25. |
| Agendia NV | 50 | 119 | 4.3.3 | Table 33: BC7 – MammaPrint – Test risk "High" – HR for DM for CT versus no CT "0.71"<br><br>As previously explained in Comment #3, given EAG's exploration of the possibility that Oncotype DX may be predictive, it is reasonable to suggest a similar examination for MammaPrint's predictiveness. Perhaps, it could be considered for inclusion in a new base case "BC8", with the Hazard Ratio for MammaPrint High Risk informed by the Knauer paper, at 0.35. This balanced approach could provide a more comprehensive evaluation.<br>Agendia has explored this scenario in Section B – Model Alteration 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, and 32. | The EAG does not agree with this proposed analysis. There is insufficient evidence to demonstrate a predictive benefit for MammaPrint in the LN+ population. |
| Agendia NV | 51 | 121 | 4.3.3 | The EAG report states that they removed the risk tapering assumption based on findings of a study by Pan et al. that suggests that the risk of Distant Metastasis remains generally flat out to 20 years.<br><br>We are under the impression that removing the risk tapering assumption results in overestimating the chemotherapy benefit in the model. In DG34, implementing the risk tapering was intended to prevent an overestimation of chemotherapy benefit. We are therefore puzzled why the EAR has chosen to change the risk tapering assumption and believe it should be reinstated as was done in DG34.<br><br>Harnan et al. (2019) page 110 | The EAG report already includes a DSA in which risk tapering is reintroduced (DSA10). This has limited impact on the model results. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | *"Although there is some evidence that suggests that for some patients with particular disease subtypes, recurrence rates remain approximately constant between 5 and 20 years, there is also uncertainty surrounding the duration over which the benefit of chemotherapy is sustained; hence, constraining recurrence at 15 years reduces the likelihood of overestimating this benefit of chemotherapy. The impact of removing this assumption of recurrence risk tapering is explored within the sensitivity analyses."*<br><br>As can be read in the quote from the previous tumour profiling test assessment, it was never denied that there remains a risk of distant metastasis for ER+/HER2- patients for a full lifetime horizon. The goal of the risk tapering used in DG10 and DG34 was not to overestimate the benefit of chemotherapy, as there is uncertainty surrounding the duration over which chemotherapy benefit is sustained.<br><br>With the removal of the risk tapering assumption, it is our understanding that the current model assumes that patients benefit from chemotherapy for 38 to 56 years (depending on the initial age in the model). Nevertheless, EBCTCG meta-analyses have demonstrated that the reduction in the recurrence rate with chemotherapy is primarily observed within the first 5 years. This raises concerns about the validity of the base case, given its assumption of a prolonged benefit from chemotherapy spanning 38 to 56 years. As a result, it's worth considering reinstating the risk tapering assumption within the base case of the EAG model. Perhaps it would be more appropriate to explore the option of removing the risk tapering in the context of a sensitivity analysis (DSA) rather than making it the primary approach at this point.<br><br>As can be read in the article of Pan et al., referenced to as the source to remove the risk tapering assumption, it is even mentioned that the results cannot be used to assess the relevance of chemotherapy after year 5. Rather, Pan and colleagues refer to the EBCTCG meta-analyses, that were used in DG34 to choose to adapt the risk tapering assumption.<br><br>Pan et al. (2017)<br>*"Furthermore, we could not reliably assess the relevance of chemotherapy to prognosis after year 5, since the women who received chemotherapy and those who did not receive chemotherapy differed in the extent of nodal involvement, tumor size, tumor grade, and perhaps unrecorded selection factors. The relevance of chemotherapy to prognosis after year 5 is best assessed in meta-analyses of the trials of chemotherapy, since randomization balances known and unknown risk factors between treatment groups. Previous EBCTCG meta-analyses have shown that most of the reduction in the recurrence rate with chemotherapy occurred in the first 5 years, with similar proportional reductions among women with ER-positive and ER-negative disease, but also indicated further benefit in years 5 to 9 with more effective regimens. Thus, the risk of recurrence after 5 years may well be somewhat lower in women who receive contemporary chemotherapy than among those in our study."* | |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Agendia has explored reinstating the risk tapering assumption in the Base Case in Section B – Model alteration 9 to 16 and 25 to 32. | |
| Agendia NV | 52 | 123 | 4.3.3 | Table 36 – QALY loss of chemotherapy -0.038<br><br>Literature thoroughly reports that chemotherapy is associated with toxicities and that the chemotherapy burden is heavy for patients. Therefore, Agendia has always believed that the input used in DG34 (-0.038) is underestimating the negative effects of chemotherapy. In the model submitted by Agendia, we have used Lidgren that assumes a QALY loss of chemotherapy of -0.124, as we believe that Lidgren is better estimating the real chemotherapy disutility. In our view, it would also be logical to use the disutility from Lidgren in the EAG model, as the recurrence free utility without chemotherapy is also based on that study.<br><br>In the Cytel model (submitted by Agendia), the QALY loss of chemotherapy input has been part of the clinical expert survey that was conducted. The uniform feedback from physicians was that the Campbell source (that was used in DG34) is likely underestimating the disutility of chemotherapy and embraced the Lidgren study as a better input parameter.<br>To illustrate with one quote from the clinical expert surveys:<br><br>"*Absolutely yeah. And I mean, I find that very hard to understand why that would be. You know, why they would select such a low impact. I mean, I'm using fairly strong language here and know it's being recorded, but the idea that you get a 3% disutility from chemo over the first year is terribly wrong."*<br><br>Breaking down the Campbell assumption to a number of days, it only accounts for a 14-day loss in a patient's quality of life. This estimate is meant to encompass all the hardship that patients have to endure with actually receiving the treatment, experiencing the side effects (including the impact chemo has on body image), receiving additional treatment prophylactically because of or as a result of the chemotherapy, hospitalization for some etc. It's worth noting that certain chemotherapy regimens necessitate more cycles and, consequently, more hospital visits (16) than the number of QoL days lost represented in the EAG's baseline disutility in the model. This suggests that the actual patient burden are likely much higher than initially estimated in the Campbell assumption and constitute the basis of our argumentation for using the Lidgren assumption instead.<br><br>Physicians (who see patients treated with chemotherapy) and reports in the literature seem to agree with the fact that a 14-day loss is vastly underestimating the 1-year QALY decrement of chemotherapy. Rather, it seems more likely that the 1-year impact of chemotherapy is better modelled using the Lidgren study (45-day loss). | Please refer to response to comment #43. |

<table>
<tr><th>Stakeholder</th><th>Comment no.</th><th>Page no.</th><th>Section no.</th><th>Comment</th><th>EAG Response</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Agendia has explored using Lidgren’s QALY decrement assumption in the Base Case in Section B – Model alteration 17 to 32.<br><br>Even by adjusting to Lidgren for the 1-year impact of chemotherapy, the impact might still be underestimated, as longer term QALY decrements in the second and following years caused by chemotherapy are not considered in the model.</td><td></td></tr>
<tr><td>Agendia NV</td><td>52</td><td>141</td><td>4.4</td><td>“The EAG undertook a re-analysis of this model which removes the assumption of predictive benefit, down-weights the chemotherapy related QALY loss and corrects the programming errors; this analysis suggests that MammaPrint is dominated by current decision-making.”<br><br>As written in Comment #24, the re-analysis of the EAG did not present the adequate results with the intended changes of the EAG, partly due to the model being less user friendly than intended. When correcting for the issue addressed in Comment #24, results are as displayed in the table below, MammaPrint dominating.<br><br>
<table>
<tr><th>Option</th><th>LYGs</th><th>QALYs</th><th>Costs</th><th>Inc. LYGs</th><th>Inc. QALYs</th><th>Inc. Costs</th><th>ICER</th></tr>
<tr><td colspan="8">Company's Scenario Analysis 5 (LN+), other tests excluded</td></tr>
<tr><td>MammaPrint</td><td>23.45</td><td>11.57</td><td>23327</td><td>1.43</td><td>0.7</td><td>-£4,676</td><td>MammaPrint</td></tr>
<tr><td>Usual care</td><td>22.02</td><td>10.87</td><td>28003</td><td>-</td><td>-</td><td>-</td><td>dominating</td></tr>
<tr><td colspan="8">Company's Scenario Analysis 5 (LN+), other tests excluded, including EAG amendments (corrected)</td></tr>
<tr><td>Option</td><td>LYGs</td><td>QALYs</td><td>Costs</td><td>Inc. LYGs</td><td>Inc. QALYs</td><td>Inc. Costs</td><td>ICER</td></tr>
<tr><td>MammaPrint</td><td>15.56</td><td>11.78</td><td>23045</td><td>-0.04</td><td>0.03</td><td>-£1,307</td><td>MammaPrint</td></tr>
<tr><td>Usual care</td><td>15.61</td><td>11.75</td><td>24351</td><td>-</td><td>-</td><td>-</td><td>dominating</td></tr>
</table>
<br>The results of EAG’s additional analysis suggest that MammaPrint is less expensive and more effective than usual care; hence, it is dominating.</td><td>Please refer to response to comment #44. The EAG believes that the EAG’s model is a more suitable basis for informing decision-making.</td></tr>
<tr><td>Agendia NV</td><td>53</td><td>142</td><td>4.4</td><td>“MammaPrint: The model suggests that the use of MammaPrint will result in a large decrease in the use of chemotherapy in women who would have benefitted from it, an increase in the lifetime probability of developing DM and additional net costs due to the cost of the test. MammaPrint is dominated by current decision-making.”<br><br>As explained in Comment #5, the reason why the results indicate that MammaPrint is dominated and that women would have benefited from adjuvant chemotherapy, is because evidence from the MINDACT trial were disregarded, showing lack of chemotherapy benefit, and instead modelled a chemotherapy benefit from an indirect source. For patients with MammaPrint Low Risk, it is currently assumed a chemotherapy benefit of HR: 0.71, a 29% risk reduction, while this Hazard Ratio should have been set at 1.00 based on MINDACT results. This model is using MINDACT data as the main input for survival and predictive values, as such, it would seem natural to use a direct input from this source to build the model for MammaPrint.</td><td>Please refer to response to comment #25.</td></tr>
</table>

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Agendia has explored the model outcomes when correcting for this Hazard Ratio assumption in 32 model alterations. See Section B. | |
| Agendia NV | 54 | 142 | 4.4 | "For each individual test, risk classification probabilities and DRFI estimates have been taken from same source. This approach maintains correlation between these parameters and avoids the potential for spectrum bias."<br><br>The report writes that the DRFI estimates within each scenario have been taken from the same source, albeit not the case for BC1 and BC2. In these cases, the Oncotype DX RS>25 had to be informed by data from TransATAC, as written on page 118 of the report. For BC1 and BC2 there is spectrum bias. In addition, Oncotype DX high risk from TransATAC is modelled as RS>25, even though the prognostic results from TransATAC were based on the old RS cut-offs, where high risk was defined as RS≥31. With the change in cut-off from RS≥31 to RS>25, an improved prognosis for the new high-risk group should be expected. This means that with the prognostic input parameter from TransATAC, any chemotherapy benefit modelled using a Hazard Ratio in the G-high group of Oncotype DX, with 0.59 from Albain or 0.71 from EBCTCG will always overestimate chemotherapy benefit in this group, likely introducing an additional source of bias for BC1 and BC2. | This is repetition of comment #26. |
| Agendia NV | 55 | 142 | 4.4 | "For the analyses of Oncotype DX, the assumption of a predictive benefit of chemotherapy has been tested."<br><br>The inclusion of testing the assumption of predictive benefit of chemotherapy with Oncotype DX in the list of strengths seems to highlight a potential bias, as mentioned in Comment #1. There seem to be a more forgiving stance regarding the significant methodological flaws in the predictive claims of Oncotype DX, while simultaneously adopting a stricter approach when evaluating MammaPrint's data. | Please refer to response to comment #21. |
| Agendia NV | 56 | 143 | 4.4 | "Therefore, the assumption of predictive benefit applied in the Exact Sciences model and the EAG's model is hinged on a clinically plausible assumption about the benefit of chemotherapy benefit in women with an Oncotype DX RS of >25, rather than empirical studies which statistically demonstrate this interaction across the full range of RS scores."<br><br>Recognition that the chemotherapy benefit assumption is hinged on a clinically plausible assumption, which does not necessarily appear to be true in the TAILORx RS>26 exploratory analysis (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6777230/), leads to the question as to why such assumption could not be made for MammaPrint. In the spirit of balance and based on the additional evidence provided by Agendia to the EAG on May 30, 2023, and summarized in Comment #3, we hope that that a similar assumption to the one given to the Oncotype DX tests can be considered. | Please refer to response to comment #21. |

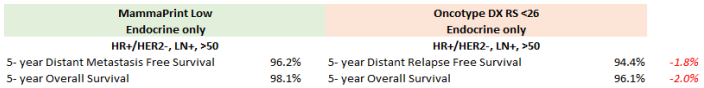
| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| Agendia NV | 57 | 143 | 4.4 | "As AOL is no longer available, this means that these women would not be identifiable current clinical practice in the NHS. This limits the applicability of the economic analyses of MammaPrint presented by both Agendia and the EAG."<br><br>As displayed in Comment #7, there is a 100% concordance between AOL and NPI.<br>Women are identifiable in current clinical practice in the NHS. Besides, patients are stratified in MINDACT based on a modified version of Adjuvant online (mAOL), that does not require web access to the any tool and can be used at any time by any physician based on already available pathology data. Therefore, the applicability of the economic analysis is not limited. | Please refer to response to comment #40. |
| Agendia NV | 58 | 144 | 4.4 | "Owing to the use of different studies across the EAG's base case analyses, and the inclusion of overlapping but non-identical samples used between the tests included in TransATAC,19 the EAG did not consider it appropriate to undertake indirect comparisons to compare tests incrementally."<br><br>Although no direct comparison is made, the conclusions drawn here indirectly suggest other tests are more cost-effective than MammaPrint. The report does not recognize that the cost-effectiveness of MammaPrint is tested in a clinically relevant population (i.e., patients with clinical high risk by AOL and NPI >3.4) and that MammaPrint is the only test that is supported by long-term follow-up of a prospective randomised trial.<br><br>Oncotype DX is tested using the RxPONDER data. The supplementary appendix of Kalinsky et al. (2021) show in Table S6 that 17.8% of post-menopausal patients enrolled were Clinical Low Risk (T-size <2cm and Grade 1). This means that 1 out of 5 patients modelled in the RxPONDER post-menopausal scenario (BC2) is not even considered for chemotherapy by most physicians, since these patients already have a good enough prognosis to forego chemotherapy based on clinical factors alone. Inclusion of this group in the CT vs no CT randomisation masks a potential benefit in the group that is truly considered as clinical high risk, and therefore potential candidate for chemotherapy. Furthermore, contamination of 17.8% clinical low risk in the RS 0-25 group, results in an improved prognosis as would be expected from a true clinical high risk cohort with RS 0-25, resulting in additional source of bias for BC2.<br><br>Prosigna and EndoPredict are modelled using TransATAC data, in which analysis women that received chemotherapy were excluded from the analysis (Sestak et al. 2018, page E2). This means that the model is exploring the effect of chemotherapy guidance with a tumour profiling test in patients who were all chemotherapy naïve as per physician choice based on clinical factors alone.<br><br>MINDACT has been the only study to address the relevant clinical question "If chemotherapy is considered, is it safe to forego chemotherapy?" and has shown that MammaPrint has Negative Predictive Value and is able to predict that patients with MammaPrint Low risk have no clinically relevant benefit of chemotherapy (also see Comment #5). | The EAG has not made any statement regarding the relative cost-effectiveness of competing tests.<br><br>The EAG report clearly states that MINDACT is a prospective RCT and describes the characteristics of the enrolled population.<br><br>RxPONDER meets the NICE scope as all patients were LN1-3. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | Unfortunately, the EAG disregarded the evidence of the Negative Predictive Value of MammaPrint from MINDACT and modelled a 29% risk reduction with the administration of chemotherapy in G-low patients, which was clearly not observed in the MINDACT trial. Disregarding this evidence has resulted in conclusions about MammaPrint in the EAG that do not do justice to the available data for MammaPrint, especially not in relation to the obvious weaknesses in clinical evidence of the other tumour profiling tests considered.<br><br>Agendia has explored the model outcomes when the model corrects the Hazard Ratio assumption in a manner that respect available data from MINDACT in 32 model alteration, which shows consistently that MammaPrint is a cost-effective method. See Section B. | |
| Agendia NV | 59 | 145 | 5.1.1 | The conclusion about the prognostic ability of MammaPrint is non-inclusive of the data provided in Table S15 of the MINDACT trial, as also commented in Comments #2, #11, #14. Acknowledgement of this data should also lead to a different conclusion in this segment.<br><br>**MammaPrint High Risk vs Low Risk – Hazard Ratio: 2.13 (P-value < 0.0001) for DMFS** | This is repetition of Comment #22. |
| Agendia NV | 60 | 146 | 5.1.1 | "Interaction tests for chemotherapy effect and risk group were statistically significant in some analyses but not others."<br><br>Several methodological flaws concerning the consideration of chemotherapy effect, risk groups and interaction tests, have been described in Comment #1. The most significant one from the Albain 2010 analysis (SWOG-8814) and not considered in the current assessment, is that in the P interaction analyses Oncotype DX **was not calculated as a chemotherapy effect and risk group.** The hazard ratio was calculated using patients with an increment of 50 units difference in recurrence score, i.e., only including the lowest and the highest recurrence scores (Albain 2010, Table 2). This is different from how the hazard ratio is applied in the model, where the Hazard ratio is used as if the hazard ratio has a significant p interaction in ODx "Low Risk group" versus ODx "High Risk group". From TAILORx we know that only 116 patients were classified as RS 51-100, that is 1.2% of the full TAILORx population (Sparano, 2020, Table 1 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6777230/). This means that the **interaction test for chemotherapy effect for risk score with an increment of 50**, already rightly described with uncertainty by EAG due to the borderline significance, is barely applicable to clinical practice, considering that only 1% of patients are classified as RS>50. | This is repetition of Comment #21. |
| Agendia NV | 61 | 146 | 5.1.1 | "Since no data for the LN+ MammaPrint high-risk group and no interaction tests were presented, it was not possible to determine from MINDACT whether MammaPrint was predictive for chemotherapy benefit."<br><br>The positive predictive value for mAOL high-risk / MP high risk was indeed not assessed in a randomised setting, as this has always been considered to be unethical, just like RxPONDER did not assess it. | As noted earlier, regarding the analysis of chemotherapy vs. no chemotherapy in C-high/G-high patients |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | However, what is available for MINDACT, is data from an analysis based on the MINDACT 2016 data cut-off that looked at the event rate of C-high/G-high HR+/HER2- patients based on actual treatment received. A multivariate analysis of distant recurrence free interval controlling for clinical risk features was performed and showed that chemotherapy was significantly associated with a 79% risk reduction (HR 0.21 [0.10-0.46], p < 0.001).The formal test of interaction between chemotherapy and genomic risk score was statistically significant (P interaction = 0.013). Data made available to EAG in Company's response on May 30th. Also, see Comment #3. | (company response 30 May), this analysis was not restricted to LN+ patients (over half were LN0). In addition, very few patients had no chemotherapy and this likely represented a biased selection, since most patients in this group were indicated for chemotherapy. Therefore the EAG does not believe this is a robust analysis for this appraisal. |
| Agendia NV | 62 | 148 | 5.1.2 | "MammaPrint is dominated by current decision-making."<br><br>As explained in Comment #5, the reason why the results indicate that MammaPrint is dominated and that women would have benefited from adjuvant chemotherapy, is because evidence from the MINDACT trial were disregarded, showing lack of chemotherapy benefit, and instead modelled a chemotherapy benefit from an indirect source. For patients with MammaPrint Low Risk, it is currently assumed a chemotherapy benefit of HR: 0.71, a 29% risk reduction, while this Hazard Ratio should have been set at 1.00 based on MINDACT results.<br><br>Agendia has explored the model outcomes when correcting for this Hazard Ratio assumption in 32 model alterations. See Section B. | The EAG disagrees with Agendia's view about the interpretation of non-significant HRs. Please refer to response to comment #25. |
| Agendia NV | 63 | 149 | 5.2.2 | "the EAG's analyses of Oncotype DX based on RxPONDER indirectly assume a predictive benefit which reflects a plausible clinical assumption about the effect of chemotherapy in women who were excluded from the trial (those with an RS of >25),"<br>Recognition that the chemotherapy benefit assumption is hinged on a clinically plausible assumption, which does not necessarily appear to be true in the TAILORx RS>26 exploratory analysis (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6777230/), leads to the question as to why such assumption could not be made for MammaPrint. In the spirit of balance and based on the additional evidence provided by Agendia to the EAG on May 30, 2023, and summarized in Comment #3, we hope that that a similar assumption to the one given to the Oncotype DX tests can be considered. | This is repetition of Comment #56. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| Agendia NV | 64 | 149 | 5.2.2 | Unavailability of AOL is listed as a limitation to interpretation.<br><br>As displayed in Comment #7, there is a 100% concordance between AOL and NPI.<br>Interpretation of the economic analysis of MammaPrint is not limited but should just be interpreted as if NPI is used. | This is repetition of Comment #27. |
| Agendia NV | 65 | 150 | 5.3 | "Limited data were available by clinical risk subgroups as defined by risk assessment tools such as NPI or PREDICT."<br><br>MINDACT has been the only study to address the relevant clinical question "If chemotherapy is considered, is it safe to forego chemotherapy?" and has shown that MammaPrint has Negative Predictive Value and is able to predict that patients with MammaPrint Low risk have no clinically relevant benefit of chemotherapy. This has been achieved by randomising patients with Clinical High risk and Genomic Low risk disease, where clinical high risk was defined by modified adjuvant online, shown to be 100% concordant with NPI (NPI >3.4). See Comment #7 | No response required. |
| Agendia NV | 66 | 150 | 5.5 | "Not all laboratories will have NGS capability which will impact how testing services are delivered."<br><br>In the event that NGS testing with MammaPrint would be set up in the UK, it will be done a centralized reference laboratory in the UK. This means that centres that do not have NGS testing capability will send samples to the NGS reference laboratory in the UK. This would be similar if all hospitals would send samples to the US for microarray testing, but then with the advantage of the NGS version being performed in the home country that does not necessitate a sample leaving the country. | No response required. |
| Agendia NV | 67 | 151 | 5.6 | "There remains some uncertainty around whether Oncotype DX is predictive of chemotherapy benefit."<br><br>As explained in Comment #1, the degree of uncertainty is considerably higher than currently reported for this assumption. This directly results in an unjustifiable advantage for Oncotype DX in this assessment as opposed to other GEP-tests. In particular unjustifiable versus MammaPrint, as MammaPrint also has evidence that would support to explore a scenario where MammaPrint is predictive for chemotherapy benefit in Genomic High risk (Comment #3). | Please refer to comment 21. |
| Agendia NV | 68 | General comment | General comment | **Final remark on the availability of clinical evidence of the four GEP-tests in HR+/HER2+/LN+ disease**<br><br>Agendia believes that it is important to consider that the first and foremost interest in utilisation of tumour profiling tests for chemotherapy decision making should be the available clinical evidence. Among the four tumour profiling tests available for HR+/HER2-/LN+ disease, only two generated prospective randomised evidence (MammaPrint and Oncotype DX), and two with only retrospective evidence available (Prosigna and EndoPredict). Of the two tests with prospective randomised evidence available, there is only one test | No further response required. |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | | that answered the relevant clinical question "If chemotherapy is considered, is it safe to forego chemotherapy?" with long-term follow-up available, which is MammaPrint.<br><br>That being said, of the two randomised trials, MINDACT is the only tests that randomised a 100% clinical high risk / 100% NPI >3.4 node positive cohort with long-term 8-year follow-up available. RxPONDER randomised a group with 1 out of 5 patients at clinical low risk (1 out of 5 at NPI ≤3.4), and with only 5-years of follow-up available.<br><br>Finally, it is important to also highlight that when comparing MINDACT and RxPONDER with comparable endpoints, patients with a MammaPrint Low Risk have a better survival at 5-year than their Oncotype DX RS < 26 counterparts. This applies both to the non-age stratified group and the older than 50 years old group.<br><br>**Clinical High / Genomic Low - HR+/HER2-/LN+**<br>End Point / Treatment / 5-year survival / Absolute difference / Source<br>DMFS: CT 96.0% / no CT 95.9% / 0.1% / Piccart 2021 Table S12<br>OS: CT 98.4% / no CT 98.8% / -0.4% / Piccart 2021 Table S12<br><br>**Clinical High / Genomic Low - HR+/HER2-/LN+ >50**<br>End Point / Treatment / 5-year survival / Absolute difference / Source<br>DMFS: CT 95.5% / no CT 96.2% / -0.7% / MINDACT Data on file<br>OS: CT 97.5% / no CT 98.1% / -0.6% / MINDACT Data on file<br><br>**Clinical Low risk and Clinical High Risk with RS 0-25 - HR+/HER2-/LN+**<br>End Point / Treatment / 5-year survival / Absolute difference / Source<br>DRFS: CT 94.9% / no CT 93.9% / 1.0% / Kalinsky 2021 Figure 2<br>OS: CT 97.8%* / no CT 96.9%* / 0.9% / Kalinsky 2020 SABCS symposium<br>* All ages was not presented, but calculated via 'number at risk' presented in Kaplan Meier curves from pre- and post-menopausal groups<br><br>**Clinical Low risk and Clinical High Risk with RS 0-25 - HR+/HER2-/LN+ >50**<br>End Point / Treatment / 5-year survival / Absolute difference / Source<br>DRFS: CT 94.4% / no CT 94.4% / 0.0% / Kalinsky 2021 Figure 2<br>OS: CT 96.2% / no CT 96.1% / 0.1% / Kalinsky 2020 SABCS symposium<br><br>A side-to-side comparison shows that for patients >50 with HR+/HER2-/LN+ breast cancer, the survival of the only two comparable endpoints in MINDACT and RxPONDER, are ~2% worse for Oncotype DX low risk, despite presenting results of a cohort with more favourable clinical factors (17.8% C-low / NPI ≤3.4). | |

| Stakeholder | Comment no. | Page no. | Section no. | Comment | EAG Response |
|---|---|---|---|---|---|
| | | | |  <br><br>**HR+/HER2-, LN+, >50**<br>**MammaPrint MINDACT Cohort** / **Oncotype DX RxPonder Cohort**<br>**Grade 1** 12.8% / 25.6%<br>**Grade 2** 71.6% / 62.6%<br>**Grade 3** 14.4% / 10.4%<br>**undefined** 1.2% / 1.4%<br><br>The comparative data presented here is a clear demonstration of the prognostic and predictive ability of MammaPrint in identifying patients as Genomic Low Risk who do not benefit from chemotherapy.<br><br>Unfortunately, the EAG disregarded the evidence of the Negative Predictive Value of MammaPrint from MINDACT and modelled a 29% risk reduction with the administration of chemotherapy in G-low patients, which was clearly not observed in the MINDACT trial. Disregarding this evidence has resulted in conclusions about MammaPrint in the EAG report that do no justice the available data, especially not in relation to the obvious weaknesses in clinical evidence of the other tumour profiling tests considered.<br><br>Agendia has explored the model outcomes when the model corrects the Hazard Ratio assumption in a manner that respects the available data from MINDACT in 32 model alteration, which shows consistently that MammaPrint is a cost-effective method. See Section B. | |

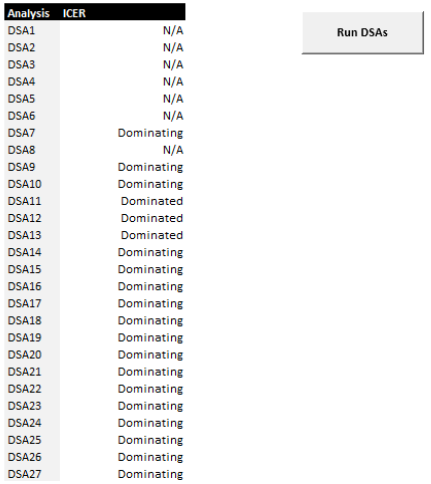
**Section B Economic model - Comments**

| Stakeholder | Comment | Description of problem | Description of proposed amendment | Result of amended model or expected impact on the result (if applicable) | EAG response |
|---|---|---|---|---|---|
| Veracyte Inc. | 1 | No issues identified | Veracyte wants to congratulate the EAR on a robust and effective modelling approach to a | NA | Thanks for the positive comment. |

| | | | | | |
|---|---|---|---|---|---|
| | | | complex topic. | | |
| Veracyte Inc | 2 | Prosigna list pricing | Veracyte would like to highlight that the Prosigna list price used in the model is the most expensive and least used option (namely the 1 kit version) and would like to request that a lower priced version be used given that it more realistically reflects what would be used in [this setting]. Prosigna is delivered and used in 1, 2, 3, 4 and 10 kit versions with an average list price of 1.296 GBP per test. 4 and 10 are mostly used in the UK (83% of total sales). Including other expenses to run the Prosigna assay (nCounter DX Flex, nCounter servicing, High pure RNA isolation kit and total laboratory staff costs the list price is 1.488 GBP per test. We would therefore ask the EAR to use this Prosigna price in the health economic modelling as the base case. Also to update Table 37 and Table 38 on the Prosigna list price so the total for | Using the 1.488 GBP price for Prosigna the Deterministic model result ICER is 25.403 GBP and the Probabilistic model result is 24.544 GBP. At the further discounted price the ICER will be further improved. | **The EAG has re-run the analysis using the updated list price. The updated results are presented in a separate addendum to the NICE report. This analysis will also be included in the final HTA monograph.** |

| | | | | | |
|---|---|---|---|---|---|
| | | | Prosigna is 1.488 GBP. The 1.488 GBP list price is not confidential in confidence so can be shared in the report (see break down of the list price for Prosigna below under Veracyte: supplementary information to the request for information process). Further Prosigna is already offered in the UK for the node negative patients at a further reduced discounted price and this confidential price will also be offered for the node positive patients as confirmed to NICE. | | |
| Agendia NV | 3 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial. (Section A, Comment #5)<br><br>The model didn't have the ability to explore a more conservative group (post-menopausal patients / >50 years). (Section A, Comment #5)<br><br>The model does not explore the positive | The subsequent issues display 32 alterations of the model with one or more model changed parameters.<br><br>These alterations present scenario's when the EAG model considers the results of MINDACT that proved lack of chemotherapy benefit in HR+/HER2-/N1 C-high/G-low breast cancer "all ages" and when restricted to patients >50. **G-low HR changed to 1.00, or** | Results of the 32 different alterations are **highly consistent** and show that in the base case with these alterations **MammaPrint is dominating in all of the deterministic and probabilistic analyses (100%).**<br><br>In the Deterministic Sensitivity Analyses, results for DSA11, DSA12, and DSA13 are mostly unchanged. These DSA's are cost-saving 33.3% of the time, and 16.7% of the time cost-saving and gains QALYs. It is logical for the remainder to be unchanged, as these DSA's undo the changes in HR proposed by Agendia based on MINDACT, by setting the chemotherapy HRs for all groups to 0.6, 0.71, and 0.8, respectively. | The EAG disagrees with the company's amendments for the reasons discussed in EAG report, Section 4.2.2.2. For brevity, the EAG's critique of the Agendia model assumptions are not repeated here. |

| | | | | | |
|---|---|---|---|---|---|
| | | predictive value of MammaPrint in G-high patients because of uncertainty/lack-of evidence, while it does for Oncotype DX based on a study that has as many if not more uncertainties than MammaPrint data. (Section A, Comment #3)<br><br>The model removed the risk tapering assumption used in DG34 and might therefore overestimate chemotherapy benefit. (Section A, Comment #31)<br><br>The model is likely underestimating the utility decrement of chemotherapy. (Section A, Comment #32) | **more conservative, HR changed to HR from MINDACT 0.85 or 0.88, for the "all ages" and ">50" group, respectively**. Rational provided in Comment #5, Section A.<br><br>Furthermore, the results of scenario's are explored where positive predictive value of MammaPrint High Risk is assumed. **G-High HR changed to 0.35.** Rational provided in Comment #3 in Section A.<br><br>All above is repeated once more combined with making the risk tapering assumption part of the base case again **(risk of DM decreasing by 50% after 10 years, and dropping to 0% after 15 years)**, and then again once more with Lidgren as the utility decrement for chemotherapy **(disutility to -0.124 instead of -0.038)**. Rational provided in Comment #31 and #32, respectively, Section A. | Of the remaining DSA's applicable to the MammaPrint scenario, **MammaPrint is dominating in 92.7% of DSA's and dominating or cost-effective in 96.0% of DSA's.** | |

| | | | | | |
|---|---|---|---|---|---|
| | | | **Key changes for the 32 alterations are summarized in an overview table on the last page of this form.** | | |
| Agendia NV | 4 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial. | **"Pre-model" cell C65 should be changed to 1** | **Deterministic results:**<br>Inc LYGs : 0.16<br>Inc QALYs: 0.08<br>Inc costs: -£1476<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br>  <br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs : 0.26<br>Inc QALYs: 0.11<br>Inc costs: -£2087<br>**ICER: Dominating** | The EAG disagrees with this analysis as it assumes a predictive benefit for MammaPrint for which there is insufficient evidence. Please refer to EAG report Section 4.2.2.2 for further details. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | Dominating |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | Dominated |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| Agendia NV | 5 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation. | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Pre-model" cell C67 should be changed to 0.35** | **Deterministic results:**<br>Inc LYGs : 0.31<br>Inc QALYs: 0.13<br>Inc costs: -£2334<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs : 0.41<br>Inc QALYs: 0.17<br>Inc costs: -£2938<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |
| Agendia NV | 6 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial. | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Settings" cell C5 should be changed to 62** | **Deterministic results:**<br>Inc LYGs: 0.12<br>Inc QALYs: 0.07<br>Inc costs: -£2004<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | Dominating |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | Dominated |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | **"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>Analysis / ICER<br>DSA1 N/A<br>DSA2 N/A<br>DSA3 N/A<br>DSA4 N/A<br>DSA5 N/A<br>DSA6 N/A<br>DSA7 Dominating<br>DSA8 N/A<br>DSA9 Dominating<br>DSA10 Dominating<br>DSA11 Dominated<br>DSA12 Dominated<br>DSA13 8673 (SWQ) manually corrected changing parameter in model<br>DSA14 Dominating<br>DSA15 Dominating<br>DSA16 Dominating<br>DSA17 Dominating<br>DSA18 Dominating<br>DSA19 Dominating<br>DSA20 Dominating<br>DSA21 Dominating<br>DSA22 Dominating<br>DSA23 Dominating<br>DSA24 Dominating<br>DSA25 Dominating<br>DSA26 Dominating<br>DSA27 Dominating<br><br>**Probabilistic results:**<br>Inc LYGs: 0.19<br>Inc QALYs: 0.10<br>Inc costs: -£2587<br>**ICER: Dominating** | |
| Agendia NV | 7 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards | **"Pre-model" cell C65 should be changed to 1** | **Deterministic results:**<br>Inc LYGs: 0.22<br>Inc QALYs: 0.11<br>Inc costs: -£2806<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, | **"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>manually corrected changing parameter in model<br><br>**Probabilistic results:**<br>Inc LYGs: 0.29<br>Inc QALYs: 0.14<br>Inc costs: -£3365<br>**ICER: Dominating** | |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | Dominating |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | 8673 (SWQ) |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | | | |
| Agendia NV | 8 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial. | **"Pre-model" cell C65 should be changed to 0.85** | **Deterministic results:**<br>Inc LYGs: -0.03<br>Inc QALYs: 0.003<br>Inc costs: -£316<br>**ICER: Dominating**<br>DSA results summary for selected scenario<br>Run DSAs<br>**Probabilistic results:**<br>Inc LYGs: 0.04 | See responses to Comments #3 and #4 above. This still assumes a predictive effect (as the HRs differ between the genomic risk groups). |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | £43,815 |
| DSA10 | Dominating |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | Dominated |
| DSA14 | 17343 (SWQ) |
| DSA15 | 8031 (SWQ) |
| DSA16 | 5914 (SWQ) |
| DSA17 | 32720 (SWQ) |
| DSA18 | Dominating |
| DSA19 | 190989 (SWQ) |
| DSA20 | Dominating |
| DSA21 | Dominated |
| DSA22 | £364,031 |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | £117,604 |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | | | Inc QALYs: 0.03<br>Inc costs: -£765<br>**ICER: Dominating** | |
| Agendia NV | 9 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation. | **"Pre-model" cell C65 should be changed to 0.85**<br><br>**"Pre-model" cell C67 should be changed to 0.35** | **Deterministic results:**<br>Inc LYGs: 0.12<br>Inc QALYs: 0.06<br>Inc costs: -£1174<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs: 0.19<br>Inc QALYs: 0.09<br>Inc costs: -£1616<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | Dominating |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | Dominated |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | 485417 (SWQ) |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | £3,264 |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| Agendia NV | 10 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"** | **Deterministic results:**<br>Inc LYGs: 0.001<br>Inc QALYs: 0.02<br>Inc costs: -£1038<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | £14,239 | manually corrected changing parameter in model |<br>| DSA10 | Dominating | |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominated | |<br>| DSA13 | 8673 (SWQ) | manually corrected changing parameter in model |<br>| DSA14 | Dominating | manually corrected changing parameter in model |<br>| DSA15 | Dominating | manually corrected changing parameter in model |<br>| DSA16 | Dominating | manually corrected changing parameter in model |<br>| DSA17 | Dominating | manually corrected changing parameter in model |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | 230887 (SWQ) | manually corrected changing parameter in model |<br>| DSA22 | £37,871 | manually corrected changing parameter in model |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.06<br>Inc QALYs: 0.04<br>Inc costs: -£1497<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | input back to a >50 scenario. | | | |
| Agendia NV | 11 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"** | **Deterministic results:**<br>Inc LYGs: 0.10<br>Inc QALYs: 0.06<br>Inc costs: -£1840<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>manually corrected changing parameter in model<br><br>**Probabilistic results:**<br>Inc LYGs: 0.16<br>Inc QALYs: 0.08<br>Inc costs: -£2275<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | Dominating |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | 8673 (SWQ) |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | | | |
| Agendia NV | 12 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit. | "**Pre-model" cell C65 should be changed to 1**<br>"**Pre-model" cell C75 should be changed to 0.5**<br>"**Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.15<br>Inc QALYs: 0.07<br>Inc costs: -£1415<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | | | DSA results summary for selected scenario<br>Analysis ICER<br>DSA1 N/A<br>DSA2 N/A<br>DSA3 N/A<br>DSA4 N/A<br>DSA5 N/A<br>DSA6 N/A<br>DSA7 Dominating<br>DSA8 N/A<br>DSA9 Dominating<br>DSA10 N/A Risk tapering already standard part of model alteration<br>DSA11 Dominated<br>DSA12 Dominated<br>DSA13 18478 (SWQ)<br>DSA14 Dominating<br>DSA15 Dominating<br>DSA16 Dominating<br>DSA17 Dominating<br>DSA18 Dominating<br>DSA19 Dominating<br>DSA20 Dominating<br>DSA21 Dominating<br>DSA22 Dominating<br>DSA23 Dominating<br>DSA24 Dominating<br>DSA25 Dominating<br>DSA26 Dominating<br>DSA27 Dominating<br>Run DSAs<br>**Probabilistic results:**<br>Inc LYGs: 0.22<br>Inc QALYs: 0.10<br>Inc costs: -£1751<br>**ICER: Dominating** | |
| Agendia NV | 13 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring | **"Pre-model" cell C65 should be changed to 1**<br>**"Pre-model" cell C67 should be changed to 0.35**<br>"**Pre-model" cell C75 should be changed to 0.5**<br>"**Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.27<br>Inc QALYs: 0.12<br>Inc costs: -£2028<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit. | | DSA results summary for selected scenario<br><br>Analysis ICER<br>DSA1 N/A<br>DSA2 N/A<br>DSA3 N/A<br>DSA4 N/A<br>DSA5 N/A<br>DSA6 N/A<br>DSA7 Dominating<br>DSA8 N/A<br>DSA9 Dominating<br>DSA10 N/A<br>DSA11 Dominated<br>DSA12 Dominated<br>DSA13 18478 (SWQ)<br>DSA14 Dominating<br>DSA15 Dominating<br>DSA16 Dominating<br>DSA17 Dominating<br>DSA18 Dominating<br>DSA19 Dominating<br>DSA20 Dominating<br>DSA21 Dominating<br>DSA22 Dominating<br>DSA23 Dominating<br>DSA24 Dominating<br>DSA25 Dominating<br>DSA26 Dominating<br>DSA27 Dominating<br><br>Run DSAs<br><br>Risk tapering already standard part of model alteration<br><br>**Probabilistic results:**<br>Inc LYGs: 0.34<br>Inc QALYs: 0.14<br>Inc costs: -£2371<br>**ICER: Dominating** | |
| Agendia NV | 14 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38** | **Deterministic results:**<br>Inc LYGs: 0.11<br>Inc QALYs: 0.06<br>Inc costs: -£1946<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | modelled was limited to patient >50.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | **"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>Risk tapering already standard part of model alteration<br>manually corrected changing parameter in model<br>manually corrected changing parameter in model<br><br>**Probabilistic results:**<br>Inc LYGs: 0.17<br>Inc QALYs: 0.09<br>Inc costs: -£2283<br>**ICER: Dominating** | |
| Agendia NV | 15 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial. | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Pre-model" cell C67 should be changed to 0.35** | **Deterministic results:**<br>Inc LYGs: 0.20<br>Inc QALYs: 0.10<br>Inc costs: -£2547<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | N/A |
| DSA11 | Dominated |
| DSA12 | 9593 (SWQ) |
| DSA13 | 54931 (SWQ) |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed | **"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs: 0.25<br>Inc QALYs: 0.12<br>Inc costs: -£2875<br>**ICER: Dominating** | |

| Analysis | ICER | |
|---|---|---|
| DSA1 | N/A | |
| DSA2 | N/A | |
| DSA3 | N/A | |
| DSA4 | N/A | |
| DSA5 | N/A | |
| DSA6 | N/A | |
| DSA7 | Dominating | |
| DSA8 | N/A | |
| DSA9 | Dominating | |
| DSA10 | N/A | Risk tapering already standard part of model alteration |
| DSA11 | Dominated | |
| DSA12 | 9593 (SWQ) | manually corrected changing parameter in model |
| DSA13 | 54931 (SWQ) | manually corrected changing parameter in model |
| DSA14 | Dominating | |
| DSA15 | Dominating | |
| DSA16 | Dominating | |
| DSA17 | Dominating | |
| DSA18 | Dominating | |
| DSA19 | Dominating | |
| DSA20 | Dominating | |
| DSA21 | Dominating | |
| DSA22 | Dominating | |
| DSA23 | Dominating | |
| DSA24 | Dominating | |
| DSA25 | Dominating | |
| DSA26 | Dominating | |
| DSA27 | Dominating | |

| | | | | | |
|---|---|---|---|---|---|
| | | group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | | | |
| Agendia NV | 16 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit. | **"Pre-model" cell C65 should be changed to 0.85**<br><br>"**Pre-model" cell C75 should be changed to 0.5**<br><br>"**Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.01<br>Inc QALYs: 0.02<br>Inc costs: -£637<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>Risk tapering already standard part of model alteration<br><br>**Probabilistic results:**<br>Inc LYGs: 0.05<br>Inc QALYs: 0.03<br>Inc costs: -£888 | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | £12,986 |
| DSA10 | N/A |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | 18478 (SWQ) |
| DSA14 | 113815 (SWQ) |
| DSA15 | 24921 (SWQ) |
| DSA16 | 16388 (SWQ) |
| DSA17 | 235691 (SWQ) |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | 30239 (SWQ) |
| DSA22 | £45,990 |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

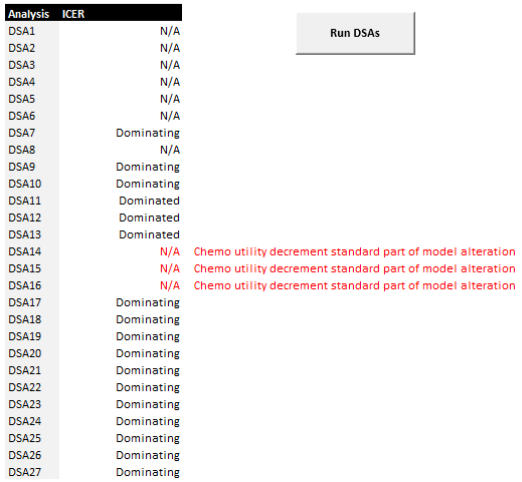
| | | | | | |
|---|---|---|---|---|---|
| | | | | **ICER: Dominating** | |
| Agendia NV | 17 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit. | **"Pre-model" cell C65 should be changed to 0.85**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.13<br>Inc QALYs: 0.06<br>Inc costs: -£1286<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>Risk tapering already standard part of model alteration<br><br>**Probabilistic results:**<br>Inc LYGs: 0.17<br>Inc QALYs: 0.08<br>Inc costs: -£1509<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |
| Agendia NV | 18 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards | **"Pre-model" cell C65 should be changed to 0.88** | **Deterministic results:**<br>Inc LYGs: 0.02<br>Inc QALYs: 0.02<br>Inc costs: -£1281<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | N/A |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | 18478 (SWQ) |
| DSA14 | Dominating |
| DSA15 | Dominating |
| DSA16 | Dominating |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | £1,351 |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | **"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>"**Pre-model" cell C75 should be changed to 0.5**<br><br>"**Pre-model" cell C76 should be changed to 0** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | £7,151 | manually corrected changing parameter in model |<br>| DSA10 | N/A | Risk tapering already standard part of model alteration |<br>| DSA11 | Dominated | |<br>| DSA12 | 9593 (SWQ) | manually corrected changing parameter in model |<br>| DSA13 | 54931 (SWQ) | manually corrected changing parameter in model |<br>| DSA14 | Dominating | |<br>| DSA15 | Dominating | manually corrected changing parameter in model |<br>| DSA16 | Dominating | manually corrected changing parameter in model |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | £13,562 | manually corrected changing parameter in model |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.06<br>Inc QALYs: 0.04<br>Inc costs: -£1520<br>**ICER: Dominating** | |

| | | | | | |
|---|---|---|---|---|---|
| Agendia NV | 19 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>"**Pre-model" cell C75 should be changed to 0.5**<br><br>"**Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.10<br>Inc QALYs: 0.06<br>Inc costs: -£1882<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | Dominating | |<br>| DSA10 | N/A | Risk tapering already standard part of model alteration |<br>| DSA11 | Dominated | |<br>| DSA12 | 9593 (SWQ) | manually corrected changing parameter in model |<br>| DSA13 | 54931 (SWQ) | manually corrected changing parameter in model |<br>| DSA14 | Dominating | |<br>| DSA15 | Dominating | |<br>| DSA16 | Dominating | |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | Dominating | |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.15<br>Inc QALYs: 0.08<br>Inc costs: -£2112<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario. | | | |
| Agendia NV | 20 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br>The model is likely underestimating the utility decrement of chemotherapy. | Use the data from the Randomised Clinical Trial MINDACT to inform the Chemotherapy, since it is a direct data source with the highest level of evidence.<br>Data was provided by Agendia in Company's response to the EAG on June 26<br>Intention to treat C-high/G-low HR+/HER2+ LN+<br>Hazard ratio CT vs no CT:<br>Non-significant 0.8497 (0.5012 – 1.441) | **Deterministic results:**<br>Inc LYGs: 0.16<br>Inc QALYs: 0.11<br>Inc costs: -£1476<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | | Because the Hazard Ratio for chemotherapy is non-significant the model input on sheet "**Pre-model" cell C65 should be changed to 1**, as no chemotherapy benefit was observed for the C-high/G-low HR+/HER2+/LN+ in MINDACT.<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br>  <br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs: 0.26<br>Inc QALYs: 0.15<br>Inc costs: -£2087<br>**ICER: Dominating** | |
| Agendia NV | 21 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | **Deterministic results:**<br>Inc LYGs: 0.31<br>Inc QALYs: 0.17<br>Inc costs: -£2334<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | | DSA results summary for selected scenario<br><br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs: 0.41<br>Inc QALYs: 0.20<br>Inc costs: -£2938<br>**ICER: Dominating** | |
| Agendia NV | 22 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38** | **Deterministic results:**<br>Inc LYGs: 0.12<br>Inc QALYs: 0.10<br>Inc costs: -£2004<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER | |
|---|---|---|
| DSA1 | N/A | |
| DSA2 | N/A | |
| DSA3 | N/A | |
| DSA4 | N/A | |
| DSA5 | N/A | |
| DSA6 | N/A | |
| DSA7 | Dominating | |
| DSA8 | N/A | |
| DSA9 | Dominating | |
| DSA10 | Dominating | |
| DSA11 | Dominated | |
| DSA12 | Dominated | |
| DSA13 | Dominated | |
| DSA14 | N/A | Chemo utility decrement standard part of model alteration |
| DSA15 | N/A | Chemo utility decrement standard part of model alteration |
| DSA16 | N/A | Chemo utility decrement standard part of model alteration |
| DSA17 | Dominating | |
| DSA18 | Dominating | |
| DSA19 | Dominating | |
| DSA20 | Dominating | |
| DSA21 | Dominating | |
| DSA22 | Dominating | |
| DSA23 | Dominating | |
| DSA24 | Dominating | |
| DSA25 | Dominating | |
| DSA26 | Dominating | |
| DSA27 | Dominating | |

| | | | | | |
|---|---|---|---|---|---|
| | | modelled was limited to patient >50.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | **"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | Dominating | |<br>| DSA10 | Dominating | |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominated | |<br>| DSA13 | Dominating | manually corrected changing parameter in model |<br>| DSA14 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA15 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA16 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | Dominating | |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.19<br>Inc QALYs: 0.14<br>Inc costs: -£2587<br>**ICER: Dominating** | |
| Agendia NV | 23 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards | **"Pre-model" cell C65 should be changed to 1** | **Deterministic results:**<br>Inc LYGs: 0.22<br>Inc QALYs: 0.15<br>Inc costs: -£2806<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, | **"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>\|---|---|---|<br>\| DSA1 \| N/A \| \|<br>\| DSA2 \| N/A \| \|<br>\| DSA3 \| N/A \| \|<br>\| DSA4 \| N/A \| \|<br>\| DSA5 \| N/A \| \|<br>\| DSA6 \| N/A \| \|<br>\| DSA7 \| Dominating \| \|<br>\| DSA8 \| N/A \| \|<br>\| DSA9 \| Dominating \| \|<br>\| DSA10 \| Dominating \| \|<br>\| DSA11 \| Dominated \| \|<br>\| DSA12 \| Dominated \| \|<br>\| DSA13 \| Dominating \| manually corrected changing parameter in model \|<br>\| DSA14 \| N/A \| Chemo utility decrement standard part of model alteration \|<br>\| DSA15 \| N/A \| Chemo utility decrement standard part of model alteration \|<br>\| DSA16 \| N/A \| Chemo utility decrement standard part of model alteration \|<br>\| DSA17 \| Dominating \| \|<br>\| DSA18 \| Dominating \| \|<br>\| DSA19 \| Dominating \| \|<br>\| DSA20 \| Dominating \| \|<br>\| DSA21 \| Dominating \| \|<br>\| DSA22 \| Dominating \| \|<br>\| DSA23 \| Dominating \| \|<br>\| DSA24 \| Dominating \| \|<br>\| DSA25 \| Dominating \| \|<br>\| DSA26 \| Dominating \| \|<br>\| DSA27 \| Dominating \| \|<br><br>**Probabilistic results:**<br>Inc LYGs: 0.29<br>Inc QALYs: 0.18<br>Inc costs: -£3365<br>**ICER: Dominating** | |

| | | | | | |
|---|---|---|---|---|---|
| | | the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | | | |
| Agendia NV | 24 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | **"Pre-model" cell C65 should be changed to 0.85**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | **Deterministic results:**<br>Inc LYGs: -0.03<br>Inc QALYs: 0.03<br>Inc costs: -£316<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>Analysis ICER<br>DSA1 N/A<br>DSA2 N/A<br>DSA3 N/A<br>DSA4 N/A<br>DSA5 N/A<br>DSA6 N/A<br>DSA7 Dominating<br>DSA8 N/A<br>DSA9 £14,439<br>DSA10 Dominating<br>DSA11 Dominated<br>DSA12 Dominated<br>DSA13 Dominated<br>DSA14 N/A Chemo utility decrement standard part of model alteration<br>DSA15 N/A Chemo utility decrement standard part of model alteration<br>DSA16 N/A Chemo utility decrement standard part of model alteration<br>DSA17 Dominating<br>DSA18 Dominating<br>DSA19 Dominating<br>DSA20 Dominating<br>DSA21 £4,216<br>DSA22 £30,400<br>DSA23 Dominating<br>DSA24 Dominating<br>DSA25 £9,821<br>DSA26 Dominating<br>DSA27 Dominating | See responses to Comments #3 and #4 above. |

<table>
<tr><td></td><td></td><td></td><td></td><td>Probabilistic results:<br>Inc LYGs: 0.04<br>Inc QALYs: 0.07<br>Inc costs: -£765<br>ICER: Dominating</td><td></td></tr>
<tr><td>Agendia NV</td><td>25</td><td>Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model is likely underestimating the utility decrement of chemotherapy.</td><td>“Pre-model” cell C65 should be changed to 0.85<br><br>“Pre-model” cell C67 should be changed to 0.35<br><br>“Pre-model” cell C124 should be changed to 0.124</td><td>Deterministic results:<br>Inc LYGs: 0.12<br>Inc QALYs: 0.09<br>Inc costs: -£1174<br>ICER: Dominating<br><br>DSA results summary for selected scenario<br><br>
<table>
<tr><th>Analysis</th><th>ICER</th><th></th></tr>
<tr><td>DSA1</td><td>N/A</td><td></td></tr>
<tr><td>DSA2</td><td>N/A</td><td></td></tr>
<tr><td>DSA3</td><td>N/A</td><td></td></tr>
<tr><td>DSA4</td><td>N/A</td><td></td></tr>
<tr><td>DSA5</td><td>N/A</td><td></td></tr>
<tr><td>DSA6</td><td>N/A</td><td></td></tr>
<tr><td>DSA7</td><td>Dominating</td><td></td></tr>
<tr><td>DSA8</td><td>N/A</td><td></td></tr>
<tr><td>DSA9</td><td>Dominating</td><td></td></tr>
<tr><td>DSA10</td><td>Dominating</td><td></td></tr>
<tr><td>DSA11</td><td>Dominated</td><td></td></tr>
<tr><td>DSA12</td><td>Dominated</td><td></td></tr>
<tr><td>DSA13</td><td>Dominated</td><td></td></tr>
<tr><td>DSA14</td><td>N/A</td><td>Chemo utility decrement standard part of model alteration</td></tr>
<tr><td>DSA15</td><td>N/A</td><td>Chemo utility decrement standard part of model alteration</td></tr>
<tr><td>DSA16</td><td>N/A</td><td>Chemo utility decrement standard part of model alteration</td></tr>
<tr><td>DSA17</td><td>Dominating</td><td></td></tr>
<tr><td>DSA18</td><td>Dominating</td><td></td></tr>
<tr><td>DSA19</td><td>Dominating</td><td></td></tr>
<tr><td>DSA20</td><td>Dominating</td><td></td></tr>
<tr><td>DSA21</td><td>Dominating</td><td></td></tr>
<tr><td>DSA22</td><td>£2,139</td><td></td></tr>
<tr><td>DSA23</td><td>Dominating</td><td></td></tr>
<tr><td>DSA24</td><td>Dominating</td><td></td></tr>
<tr><td>DSA25</td><td>Dominating</td><td></td></tr>
<tr><td>DSA26</td><td>Dominating</td><td></td></tr>
<tr><td>DSA27</td><td>Dominating</td><td></td></tr>
</table>
Run DSAs<br><br>Probabilistic results:<br>Inc LYGs: 0.19<br>Inc QALYs: 0.12<br>Inc costs: -£1616<br>ICER: Dominating</td><td>See responses to Comments #3 and #4 above.</td></tr>
</table>

| | | | | | |
|---|---|---|---|---|---|
| Agendia NV | 26 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | **Deterministic results:**<br>Inc LYGs: 0.001<br>Inc QALYs: 0.05<br>Inc costs: -£1038<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs: 0.06<br>Inc QALYs: 0.08<br>Inc costs: -£1497<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER | |
|---|---|---|
| DSA1 | N/A | |
| DSA2 | N/A | |
| DSA3 | N/A | |
| DSA4 | N/A | |
| DSA5 | N/A | |
| DSA6 | N/A | |
| DSA7 | Dominating | |
| DSA8 | N/A | |
| DSA9 | £6,790 | manually corrected changing parameter in model |
| DSA10 | Dominating | |
| DSA11 | Dominated | |
| DSA12 | Dominated | |
| DSA13 | Dominating | manually corrected changing parameter in model |
| DSA14 | N/A | Chemo utility decrement standard part of model alteration |
| DSA15 | N/A | Chemo utility decrement standard part of model alteration |
| DSA16 | N/A | Chemo utility decrement standard part of model alteration |
| DSA17 | Dominating | |
| DSA18 | Dominating | |
| DSA19 | Dominating | |
| DSA20 | Dominating | |
| DSA21 | Dominating | |
| DSA22 | £10,235 | manually corrected changing parameter in model |
| DSA23 | Dominating | |
| DSA24 | Dominating | |
| DSA25 | Dominating | |
| DSA26 | Dominating | |
| DSA27 | Dominating | |

| | | | | | |
|---|---|---|---|---|---|
| | | input back to a >50 scenario.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | | | |
| Agendia NV | 27 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>Running the deterministic sensitivity analysis defaults settings on the | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | **Deterministic results:**<br>Inc LYGs: 0.10<br>Inc QALYs: 0.10<br>Inc costs: -£1840<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | Dominating | |<br>| DSA10 | Dominating | |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominated | |<br>| DSA13 | Dominating | manually corrected changing parameter in model |<br>| DSA14 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA15 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA16 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | Dominating | |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.16<br>Inc QALYs: 0.12<br>Inc costs: -£2275<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario.<br>The model is likely underestimating the utility decrement of chemotherapy. | | | |
| Agendia NV | 28 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br>The model removed the risk tapering assumption used in DG34 and is might therefore | "**Pre-model" cell C65 should be changed to 1**<br>"**Pre-model" cell C75 should be changed to 0.5**<br>"**Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.15<br>Inc QALYs: 0.10<br>Inc costs: -£1415<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | overestimate chemotherapy benefit.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | **"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>Risk tapering already standard part of model alteration<br>Chemo utility decrement standard part of model alteration<br>Chemo utility decrement standard part of model alteration<br>Chemo utility decrement standard part of model alteration<br><br>**Probabilistic results:**<br>Inc LYGs: 0.22<br>Inc QALYs: 0.13<br>Inc costs: -£1751<br>**ICER: Dominating** | |
| Agendia NV | 29 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>"**Pre-model" cell C75 should be changed to 0.5**<br><br>"**Pre-model" cell C76 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.27<br>Inc QALYs: 0.15<br>Inc costs: -£2028<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | N/A |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | Dominating |
| DSA14 | N/A |
| DSA15 | N/A |
| DSA16 | N/A |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | **"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>Risk tapering already standard part of model alteration<br><br>Chemo utility decrement standard part of model alteration<br><br>**Probabilistic results:**<br>Inc LYGs: 0.34<br>Inc QALYs: 0.18<br>Inc costs: -£2371<br>**ICER: Dominating** | |
| Agendia NV | 30 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38** | **Deterministic results:**<br>Inc LYGs: 0.11<br>Inc QALYs: 0.10<br>Inc costs: -£1946<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER |
|---|---|
| DSA1 | N/A |
| DSA2 | N/A |
| DSA3 | N/A |
| DSA4 | N/A |
| DSA5 | N/A |
| DSA6 | N/A |
| DSA7 | Dominating |
| DSA8 | N/A |
| DSA9 | Dominating |
| DSA10 | N/A |
| DSA11 | Dominated |
| DSA12 | Dominated |
| DSA13 | Dominating |
| DSA14 | N/A |
| DSA15 | N/A |
| DSA16 | N/A |
| DSA17 | Dominating |
| DSA18 | Dominating |
| DSA19 | Dominating |
| DSA20 | Dominating |
| DSA21 | Dominating |
| DSA22 | Dominating |
| DSA23 | Dominating |
| DSA24 | Dominating |
| DSA25 | Dominating |
| DSA26 | Dominating |
| DSA27 | Dominating |

| | | | | | |
|---|---|---|---|---|---|
| | | modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high patients because of uncertainty/lack-of evidence, while it does for Oncotype DX based on a study that has as many if not more uncertainties than MammaPrint data.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, | **"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>"**Pre-model" cell C75 should be changed to 0.5**<br><br>"**Pre-model" cell C76 should be changed to 0**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | Dominating | |<br>| DSA10 | N/A | Risk tapering already standard part of model alteration |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominating | manually corrected changing parameter in model |<br>| DSA13 | Dominating | |<br>| DSA14 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA15 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA16 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | Dominating | |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.17<br>Inc QALYs: 0.13<br>Inc costs: -£2283<br>**ICER: Dominating** | |

| | | | | | |
|---|---|---|---|---|---|
| | | the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | | | |
| Agendia NV | 31 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more | **"Pre-model" cell C65 should be changed to 1**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"** | **Deterministic results:**<br>Inc LYGs: 0.20<br>Inc QALYs: 0.14<br>Inc costs: -£2547<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario.<br><br>The model is likely underestimating the utility | **"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | Dominating | |<br>| DSA10 | N/A | Risk tapering already standard part of model alteration |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominating | manually corrected changing parameter in model |<br>| DSA13 | Dominating | |<br>| DSA14 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA15 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA16 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | Dominating | |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.25<br>Inc QALYs: 0.17<br>Inc costs: -£2875<br>**ICER: Dominating** | |

| | | | | | |
|---|---|---|---|---|---|
| | | decrement of chemotherapy. | | | |
| Agendia NV | 32 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>The model is likely underestimating the utility decrement of chemotherapy. | **"Pre-model" cell C65 should be changed to 0.85**<br><br>**"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | **Deterministic results:**<br>Inc LYGs: 0.01<br>Inc QALYs: 0.05<br>Inc costs: -£637<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>**Probabilistic results:**<br>Inc LYGs: 0.05<br>Inc QALYs: 0.07<br>Inc costs: -£888<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |
| Agendia NV | 33 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards | **"Pre-model" cell C65 should be changed to 0.85** | **Deterministic results:**<br>Inc LYGs: 0.13<br>Inc QALYs: 0.09<br>Inc costs: -£1286<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| Analysis | ICER | |
|---|---|---|
| DSA1 | N/A | |
| DSA2 | N/A | |
| DSA3 | N/A | |
| DSA4 | N/A | |
| DSA5 | N/A | |
| DSA6 | N/A | |
| DSA7 | Dominating | |
| DSA8 | N/A | |
| DSA9 | £5,936 | |
| DSA10 | N/A | Risk tapering already standard part of model alteration |
| DSA11 | Dominated | |
| DSA12 | Dominated | |
| DSA13 | Dominating | |
| DSA14 | N/A | Chemo utility decrement standard part of model alteration |
| DSA15 | N/A | Chemo utility decrement standard part of model alteration |
| DSA16 | N/A | Chemo utility decrement standard part of model alteration |
| DSA17 | Dominating | |
| DSA18 | Dominating | |
| DSA19 | Dominating | |
| DSA20 | Dominating | |
| DSA21 | Dominating | |
| DSA22 | £14,852 | |
| DSA23 | Dominating | |
| DSA24 | Dominating | |
| DSA25 | Dominating | |
| DSA26 | Dominating | |
| DSA27 | Dominating | |

| | | | | | |
|---|---|---|---|---|---|
| | | the findings of the MINDACT trial.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit..<br><br>The model is likely underestimating the utility decrement of chemotherapy. | **"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>Analysis ICER<br>DSA1 N/A<br>DSA2 N/A<br>DSA3 N/A<br>DSA4 N/A<br>DSA5 N/A<br>DSA6 N/A<br>DSA7 Dominating<br>DSA8 N/A<br>DSA9 Dominating<br>DSA10 N/A Risk tapering already standard part of model alteration<br>DSA11 Dominated<br>DSA12 Dominated<br>DSA13 Dominating<br>DSA14 N/A Chemo utility decrement standard part of model alteration<br>DSA15 N/A Chemo utility decrement standard part of model alteration<br>DSA16 N/A Chemo utility decrement standard part of model alteration<br>DSA17 Dominating<br>DSA18 Dominating<br>DSA19 Dominating<br>DSA20 Dominating<br>DSA21 Dominating<br>DSA22 £897<br>DSA23 Dominating<br>DSA24 Dominating<br>DSA25 Dominating<br>DSA26 Dominating<br>DSA27 Dominating<br><br>**Probabilistic results:**<br>Inc LYGs: 0.17<br>Inc QALYs: 0.12<br>Inc costs: -£1509<br>**ICER: Dominating** | |
| Agendia NV | 34 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0** | **Deterministic results:**<br>Inc LYGs: 0.02<br>Inc QALYs: 0.06<br>Inc costs: -£1281<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

| | | | | | |
|---|---|---|---|---|---|
| | | modelled was limited to patient >50.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity analysis defaults settings on the "Settings" sheet for "initial age", "Proportion of women pre-menopausal", "Time Horizon" and "Data source for post-test chemo probabilities" back to BC7's original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the "settings" input back to a >50 scenario.<br><br>The model is likely underestimating the utility | **"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>**"Pre-model" cell C75 should be changed to 0.5**<br><br>**"Pre-model" cell C76 should be changed to 0**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | £3,636 | manually corrected changing parameter in model |<br>| DSA10 | N/A | Risk tapering already standard part of model alteration |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominating | manually corrected changing parameter in model |<br>| DSA13 | Dominating | |<br>| DSA14 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA15 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA16 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | £5,236 | manually corrected changing parameter in model |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.06<br>Inc QALYs: 0.08<br>Inc costs: -£1520<br>**ICER: Dominating** | |

| | | | | | |
|---|---|---|---|---|---|
| | | decrement of chemotherapy. | | | |
| Agendia NV | 35 | Chemotherapy Hazard Ratios for BC7 MammaPrint for the G-low group, disregards the findings of the MINDACT trial.<br><br>The EAG model did not have access to model input from Agendia to explore a scenario where the population modelled was limited to patient >50.<br><br>The model does not explore the positive predictive value of MammaPrint in G-high, while it does for Oncotype DX. Exploring both would be a more balanced approach that could provide a more comprehensive evaluation.<br><br>The model removed the risk tapering assumption used in DG34 and is might therefore overestimate chemotherapy benefit.<br><br>Running the deterministic sensitivity | **"Pre-model" cell C65 should be changed to 0.88**<br><br>**"Pre-model" cell C67 should be changed to 0.35**<br><br>**"Settings" cell C5 should be changed to 62**<br><br>**"Settings" cell C6 should be changed to 0**<br><br>**"Settings" cell C7 should be changed to 38**<br><br>**"Settings" cell C15 should be changed to "Holt et al., 2023, LN+ post-menopausal, 2-level"**<br><br>"**Pre-model" cell C75 should be changed to 0.5**<br><br>"**Pre-model" cell C76 should be changed to 0**<br><br>**"Pre-model" cell C124 should be changed to 0.124** | **Deterministic results:**<br>Inc LYGs: 0.10<br>Inc QALYs: 0.10<br>Inc costs: -£1882<br>**ICER: Dominating**<br><br>DSA results summary for selected scenario<br><br>Run DSAs<br><br>| Analysis | ICER | |<br>|---|---|---|<br>| DSA1 | N/A | |<br>| DSA2 | N/A | |<br>| DSA3 | N/A | |<br>| DSA4 | N/A | |<br>| DSA5 | N/A | |<br>| DSA6 | N/A | |<br>| DSA7 | Dominating | |<br>| DSA8 | N/A | |<br>| DSA9 | Dominating | |<br>| DSA10 | N/A | Risk tapering already standard part of model alteration |<br>| DSA11 | Dominated | |<br>| DSA12 | Dominating | manually corrected changing parameter in model |<br>| DSA13 | Dominating | |<br>| DSA14 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA15 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA16 | N/A | Chemo utility decrement standard part of model alteration |<br>| DSA17 | Dominating | |<br>| DSA18 | Dominating | |<br>| DSA19 | Dominating | |<br>| DSA20 | Dominating | |<br>| DSA21 | Dominating | |<br>| DSA22 | Dominating | |<br>| DSA23 | Dominating | |<br>| DSA24 | Dominating | |<br>| DSA25 | Dominating | |<br>| DSA26 | Dominating | |<br>| DSA27 | Dominating | |<br><br>**Probabilistic results:**<br>Inc LYGs: 0.15<br>Inc QALYs: 0.12<br>Inc costs: -£2112<br>**ICER: Dominating** | See responses to Comments #3 and #4 above. |

<table>
<tr><td></td><td></td><td>analysis defaults settings on the “Settings” sheet for “initial age”, “Proportion of women pre-menopausal”, “Time Horizon” and “Data source for post-test chemo probabilities” back to BC7’s original inputs accounting for a pre- and post-menopausal mixed group. For this reason, the DSA for this model alteration had to be run manually and lead to the need to manually correct DSA results as the results differed when correcting the “settings” input back to a >50 scenario.<br><br>The model is likely underestimating the utility decrement of chemotherapy.</td><td></td><td></td><td></td></tr>
<tr><td>Agendia NV</td><td>36</td><td>Alterations key change overview:
<table>
<tr><th></th><th>Key change #1</th><th>Key change #2</th><th>Key change #3</th><th>Key change #4</th><th>Key change #5</th></tr>
<tr><td>Alteration 1</td><td>MP Low Risk – CT HR to 1.00</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Alteration 2</td><td>MP Low Risk – CT HR to 1.00</td><td></td><td>MP High Risk – CT HR Predictive explored</td><td></td><td></td></tr>
<tr><td>Alteration 3</td><td>MP Low Risk – CT HR to 1.00</td><td>Population restricted to >50</td><td></td><td></td><td></td></tr>
</table></td><td></td><td></td><td>See responses to Comments #3 and #4 above.</td></tr>
</table>

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | **Alteration 4** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | | |
| | | | **Alteration 5** | MP Low Risk – CT HR to 0.85 | | | | |
| | | | **Alteration 6** | MP Low Risk – CT HR to 0.85 | | MP High Risk – CT HR Predictive explored | | |
| | | | **Alteration 7** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | | | |
| | | | **Alteration 8** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | | |
| | | | **Alteration 9** | MP Low Risk – CT HR to 1.00 | | | Risk tapering assumption restored | |
| | | | **Alteration 10** | MP Low Risk – CT HR to 1.00 | | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | |
| | | | **Alteration 11** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | | Risk tapering assumption restored | |
| | | | **Alteration 12** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | |
| | | | **Alteration 13** | MP Low Risk – CT HR to 0.85 | | | Risk tapering assumption restored | |
| | | | **Alteration 14** | MP Low Risk – CT HR to 0.85 | | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | |
| | | | **Alteration 15** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | | Risk tapering assumption restored | |
| | | | **Alteration 16** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | |
| | | | **Alteration 17** | MP Low Risk – CT HR to 1.00 | | | | Lidgren – CT utility decrement |
| | | | **Alteration 18** | MP Low Risk – CT HR to 1.00 | | MP High Risk – CT HR Predictive explored | | Lidgren – CT utility decrement |
| | | | **Alteration 19** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | | | Lidgren – CT utility decrement |
| | | | **Alteration 20** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | | Lidgren – CT utility decrement |
| | | | **Alteration 21** | MP Low Risk – CT HR to 0.85 | | | | Lidgren – CT utility decrement |
| | | | **Alteration 22** | MP Low Risk – CT HR to 0.85 | | MP High Risk – CT HR Predictive explored | | Lidgren – CT utility decrement |
| | | | **Alteration 23** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | | | Lidgren – CT utility decrement |
| | | | **Alteration 24** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | | Lidgren – CT utility decrement |
| | | | **Alteration 25** | MP Low Risk – CT HR to 1.00 | | | Risk tapering assumption restored | Lidgren – CT utility decrement |
| | | | **Alteration 26** | MP Low Risk – CT HR to 1.00 | | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | Lidgren – CT utility decrement |
| | | | **Alteration 27** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | | Risk tapering assumption restored | Lidgren – CT utility decrement |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **Alteration 28** | MP Low Risk – CT HR to 1.00 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | Lidgren – CT utility decrement |
| | | **Alteration 29** | MP Low Risk – CT HR to 0.85 | | | Risk tapering assumption restored | Lidgren – CT utility decrement |
| | | **Alteration 30** | MP Low Risk – CT HR to 0.85 | | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | Lidgren – CT utility decrement |
| | | **Alteration 31** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | | Risk tapering assumption restored | Lidgren – CT utility decrement |
| | | **Alteration 32** | MP Low Risk – CT HR to 0.88 | Population restricted to >50 | MP High Risk – CT HR Predictive explored | Risk tapering assumption restored | Lidgren – CT utility decrement |