NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 29 • Issue 49 • October 2025

ISSN 2046-4924

Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node-positive early breast cancer: a systematic review and economic evaluation

Paul Tappenden, Katy Cooper, Jean Hamilton, Gamze Nalbant, Munira Essat, Annabel Rayner, Ruth Wong, Nicolò Matteo Luca Battisti, Lynda Wyld and Uzma Asghar

DOI 10.3310/KGFD4040

NIHR | National Institute for Health and Care Research

Health Technology Assessment

Check for updates

Extended Research Article

Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node-positive early breast cancer: a systematic review and economic evaluation

Paul Tappenden[1*] Katy Cooper[1] Jean Hamilton[1] Gamze Nalbant[1] Munira Essat[1] Annabel Rayner[1] Ruth Wong[1] Nicolò Matteo Luca Battisti[2] Lynda Wyld[3] and Uzma Asghar[2]

[1]Sheffield Centre for Health and Related Research (SCHARR), School of Medicine and Population Health, University of Sheffield, Sheffield, UK
[2]Department of Medicine, Breast Unit, The Royal Marsden NHS Foundation Trust, London, UK
[3]Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, UK

[*]**Corresponding author** p.tappenden@sheffield.ac.uk

Published October 2025
DOI: 10.3310/KGFD4040

This report should be referenced as follows:

Tappenden P, Cooper K, Hamilton J, Nalbant G, Essat M, Rayner A, *et al.* Tumour profiling tests to guide adjuvant chemotherapy decisions in lymph node-positive early breast cancer: a systematic review and economic evaluation. *Health Technol Assess* 2025;**29**(49). https://doi.org/10.3310/KGFD4040

Health Technology Assessment

NICE TAR and DAR

ISSN 2046-4924 (Online)

Impact factor: 4

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 4 and is ranked 30th (out of 174 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal
Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as award number NIHR135822. The protocol was agreed in March 2023. The draft manuscript began editorial review in October 2023 and was accepted for publication in February 2025. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Background: Breast cancer is the most commonly diagnosed cancer in women in England. Breast cancer and chemotherapy treatment can impact upon patients' quality of life and survival. Tumour profiling tests can help to identify whether patients will benefit from chemotherapy.

Objectives: To evaluate the effectiveness and cost-effectiveness of four tumour profiling tests (Oncotype DX, Prosigna, EPclin and MammaPrint), compared with current decision-making (no testing), to guide use of adjuvant chemotherapy in people with hormone-receptor positive, human epidermal growth factor receptor 2 negative, early-stage breast cancer with one to three positive lymph nodes.

Methods and data sources: A systematic review identified studies via a literature search in April 2023 and from our previous review. The economic analysis included a review of existing models and development of an independent model.

Results: Fifty-five articles were included, 42 for prognostic and predictive ability and 13 for impact on chemotherapy decisions. All four tests showed prognostic ability for determining risk of relapse. The RxPONDER randomised controlled trial of Oncotype DX indicated no chemotherapy benefit in post-menopausal lymph node-positive patients with a recurrence score of 0–25, but a statistically significant benefit in pre-menopausal patients with a recurrence score of 0–25. An older randomised controlled trial reanalysis (Southwest Oncology Group-8814) indicated lower relative chemotherapy benefit with lower recurrence score, with statistically significant interactions between recurrence score and chemotherapy benefit in some but not all analyses. There was no clear evidence of prediction of relative chemotherapy benefit for Prosigna, EPclin or MammaPrint. Decision impact studies in lymph node-positive populations in the United Kingdom and Europe were only available for Oncotype DX, and they reported a reduction of 12–75% in chemotherapy recommendations following testing.

Based on the list prices of the tests and downstream treatments, the independent model suggests the following:

Oncotype DX: This test dominates current decision-making in post-menopausal lymph node-positive women, provided an assumption of predictive benefit holds, but the test is dominated if this assumption does not hold. The test is dominated by current decision-making in pre-menopausal lymph node-positive women.

Prosigna: The probabilistic incremental cost-effectiveness ratio for Prosigna versus current decision-making in post-menopausal lymph node-positive women is £39,357 per quality-adjusted life-year gained.

EPclin: The probabilistic incremental cost-effectiveness ratio for EPclin versus current decision-making in post-menopausal lymph node-positive women is £4113 per quality-adjusted life-year gained.

MammaPrint: Within clinical high-risk pre-/post-menopausal lymph node-positive women, MammaPrint is dominated by current decision-making.

Limitations: There are limited data on the prediction of chemotherapy benefit; evidence for Oncotype DX may support a predictive benefit, but this is uncertain. Decision impact studies in a lymph node-positive population were available only for Oncotype DX. The economic model relies on an assumption of predictive benefit for Oncotype DX, and broader assumptions around the way that Prosigna, MammaPrint and EPclin test results would affect chemotherapy decisions.

Conclusions: All four tests provide prognostic information on the risk of relapse. The evidence on prediction of relative chemotherapy benefit is weaker and mostly limited to Oncotype DX. The economic analyses indicate that Oncotype DX and EPclin may have favourable cost-effectiveness profiles in post-menopausal lymph node-positive subgroups, although this is uncertain.

Study registration: This study is registered as PROSPERO CRD42023425638.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme (NIHR award ref: NIHR135822) and is published in full in *Health Technology Assessment*; Vol. 29, No. 49. See the NIHR Funding and Awards website for further award information.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

AE	adverse event
AI	aromatase inhibitor
AJCC	American Joint Committee on Cancer
AML	acute myeloid leukaemia
AOL	Adjuvant! Online
BC	base case
BCSS	breast cancer-specific survival
BNF	*British National Formulary*
BRCA	breast cancer gene
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
CML	chronic myeloid leukaemia
CPCI	Conference Proceedings Citation Index
CS	company's submission
CTS	Clinical Treatment Score
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
ECG	electrocardiogram
EFS	event-free survival
EMBASE	Excerpta Medica Database
eMIT	electronic Market Information Tool
EQ-5D	EuroQol-5 Dimensions
EQ-5D-3L	EuroQol-5 Dimensions, three-level version
ER	oestrogen receptor
ERBB	erythroblastic oncogene B
ESBC	early-stage breast cancer
ET	endocrine therapy
FACT-B	Functional Assessment of Cancer Therapy – Breast cancer
FACT-G	Functional Assessment of Cancer Therapy – General
FEC75	fluorouracil, epirubicin and cyclophosphamide
FEC100-T	fluorouracil, epirubicin, cyclophosphamide and docetaxel
FFPE	formalin-fixed paraffin-embedded
FISH	fluorescence in situ hybridisation
FN	febrile neutropenia
G-CSF	granulocyte-colony stimulating factor
GP	general practitioner
HCHS	Hospital and Community Health Services
HER2	human epidermal growth factor receptor 2
HF	heart failure
HR	hormone receptor
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
IPD	individual patient data
IV	intravenous
LCIS	lobular carcinoma in situ
LN	lymph node
LN+	lymph node positive
LN0	lymph node negative
LR	local recurrence
LRR	locoregional recurrence
LYG	life year gained
mAOL	Modified Adjuvant! Online
MEDLINE	Medical Literature Analysis and Retrieval System Online
MeSH	medical subject heading
MIMS	*Monthly Index of Medical Specialities*
mRNA	messenger ribonucleic acid
NCCN	National Comprehensive Cancer Network
NCDB	National Cancer Database
NCRAS	National Cancer Registration and Analysis Service
NG	NICE guideline
NGS	next-generation sequencing
NHB	net health benefit
NHSCII	NHS Cost Inflation Indices
NHSE	National Health Service England
NICE	National Institute for Health and Care Excellence
NPI	Nottingham Prognostic Index
NR	not reported
OS	overall survival
PAM50	Prediction Analysis of Microarray 50
PARP	poly-ADP ribose polymerase
PAS	Patient Access Scheme
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
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
SEER	Surveillance Epidemiology and End Results
SLNB	sentinel lymph node biopsy
SLR	systematic literature review
STAI	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
UKBCG	UK Breast Cancer Group
VAT	value added tax
WTP	willingness to pay

Plain language summary

Breast cancer is the most common cancer in women in England. Breast cancer, and its treatment, can affect a person's quality of life and how long they live for. Most women with early-stage breast cancer which has spread to one to three lymph nodes receive chemotherapy to stop the cancer from coming back and spreading elsewhere in the body, but this treatment can cause side effects, including damage to the heart and secondary cancers. Currently, some doctors use computer tools which use information about the patient and the tumour to decide if chemotherapy is needed. Tumour profiling tests are used to help women with early breast cancer to decide whether they should have chemotherapy. They test small samples of a patient's tumour to find out whether the genes in it mean that a person has a high or low risk of the disease returning. If this risk is low, the patient might not need chemotherapy and, therefore, they can avoid its side effects. Some tests might also be able to identify which patients are more likely to respond to chemotherapy.

This study looked at the evidence for four tumour profiling tests. Fifty-four clinical studies were identified. The results suggest that all of the tests can give information on the risk of the cancer returning. There was some information about whether one of the tests (Oncotype DX) can predict which patients will respond to chemotherapy. There was information about how using one test (Oncotype DX) affects the decision to have chemotherapy. Our study also looked at whether or not these tests represent good value for money for the National Health Service through cost-effectiveness analyses. The analyses showed that two of the tests (Oncotype DX and EPclin) may represent a good use of National Health Service resources for some patient groups.

Scientific summary

Background

Breast cancer is the most commonly diagnosed cancer in women and the fourth most common cause of cancer-related death in the UK. Most people with lymph node-positive (LN+) breast cancer receive adjuvant chemotherapy due to their increased risk of recurrence. However, chemotherapy is associated with considerable adverse effects. Currently, adjuvant chemotherapy decisions may be informed by clinical and pathological information, sometimes via a risk prediction tool. 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 patients who will benefit most. Tumour profiling tests aim to improve decisions on chemotherapy use by improving the categorisation of patients according to risk and the identification of patients who will benefit most 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, human epidermal growth factor receptor 2 (HER2) negative, lymph node-negative (LN0) early breast cancer, including those with micrometastases. Two other tests assessed in DG34 (MammaPrint and immunohistochemical 4 (IHC4)) were not recommended in the LN0 population. While DG34 also assessed these tests in women with LN+ early breast cancer, the Appraisal Committee did not make any specific recommendations on the use of any test within LN+ patients. This assessment provides an updated systematic literature review and economic analysis of four tumour profiling tests (Oncotype DX, Prosigna, EPclin and MammaPrint) compared to current decision-making in women with ER-positive [and/or progesterone receptor (PR)-positive], HER2-negative, early breast cancer with one to three positive lymph nodes.

Objectives

The main research question 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 are:

- To conduct a systematic review of effectiveness and cost-effectiveness of four tumour profiling tests (Oncotype DX, Prosigna, EPclin and MammaPrint).
- To develop a health economic model to assess the cost-effectiveness of tumour profiling tests compared with current decision-making (no testing) on the use of chemotherapy from the perspective of the NHS and Personal Social Services (PSS).

Methods

Clinical evidence review methods

The External Assessment Group (EAG) undertook a systematic review of Oncotype DX, Prosigna, EPclin and MammaPrint for guiding adjuvant chemotherapy decisions in women with ER+/PR+, HER2− early breast cancer where the study population was at least 80% LN+. Studies were identified from the previous review which informed NICE DG34 (searches conducted in 2017) plus an updated search (April 2023) covering MEDLINE, EMBASE, Cochrane, and other sources. Eligible data types included prospective randomised controlled trials (RCTs) and studies of prognostic ability, prediction of relative chemotherapy benefit, impact of tests on chemotherapy decisions (restricted to UK and European studies), and health-related quality of life (HRQoL) and anxiety associated with testing.

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 and economic analyses in LN+ populations published since 2017. The EAG also 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 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. Model parameter values were informed by the RxPONDER, TransATAC, SWOG-8814 and MINDACT trials, a recent 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 including price discounts for the tests and downstream treatments were provided in a separate confidential appendix to NICE.

Results

Clinical evidence results

Overview of available evidence

In total, 55 articles were included, 42 relating to prognostic and predictive ability, and 13 relating to impact on chemotherapy decisions. Data were reported for two prospective RCTs (RxPONDER and MINDACT). In RxPONDER, LN+ patients with an Oncotype DX Breast Recurrence Score (RS) of 0–25 were randomised to chemotherapy versus no chemotherapy. In MINDACT, patients with discordant MammaPrint risk and clinical risk were randomised to chemotherapy versus no chemotherapy. In addition, the ongoing OPTIMA RCT compares Prosigna test-directed chemotherapy use versus standard chemotherapy use; however, results are not yet available.

Prognostic ability

All four tests demonstrated prognostic ability for determining risk of relapse in LN+ populations, both with and without adjustment for clinical factors.

Prediction of chemotherapy benefit

No predictive data in a LN+ population were identified for Prosigna or EPclin. For Oncotype DX, a reanalysis of the SWOG-8814 RCT using cut-offs of RS < 18 and > 30 indicated no effect of chemotherapy on 10-year disease-free survival 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, with statistically significant interaction tests in some but not all analyses. The RxPONDER prospective RCT reported no benefit of chemotherapy in post-menopausal patients with an RS of 0–25, but a statistically significant benefit in pre-menopausal patients with an RS of 0–25, while the test for interaction between RS (within the range 0–25) and effect of chemotherapy was not statistically significant in either group. The National Cancer Database reported 5-year overall survival within post-menopausal or older-age subgroups with RS ≤ 25; some analyses showed a statistically significant chemotherapy benefit while others did not. For MammaPrint, prediction of chemotherapy benefit could not be determined from the LN+ subgroup of the MINDACT prospective RCT, because all patients in the clinical high-risk, MammaPrint high-risk group, were offered chemotherapy (there was a non-significant effect of chemotherapy in the LN+ MammaPrint low-risk group). A cohort reanalysis from 2009 reported a non-significant interaction test between MammaPrint score and effect of chemotherapy on breast cancer-specific survival (p = 0.95).

Decision impact

Studies on chemotherapy decisions in LN+ populations in the UK and Europe indicated a net reduction in the percentage of patients recommended chemotherapy pre-test to post-test of between 12% and 75%, with greater reductions in groups with lower RS. All studies used Oncotype DX; no decision impact studies were identified for EPclin, Prosigna or MammaPrint.

Health-related quality of life and anxiety

No studies reported HRQoL or anxiety associated with using tumour profiling tests in a LN+ population. Therefore, studies in a LN0 or mixed nodal status population were briefly summarised, with mixed results regarding the impact of testing and anxiety.

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 current decision-making. 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 large 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, driven by a large reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment and an increase in the lifetime probability of developing distant metastases (DM). 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 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 £4113 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.

Discussion

Strengths and limitations in the clinical evidence base

Strengths of the clinical evidence base include the fairly substantial evidence for prognostic ability of all four tests. A major limitation is the difficulty in collecting new data on predictive ability, as it is not considered ethical to randomise

patients who are high risk on any test to chemotherapy versus no chemotherapy. Therefore, although there are prospective RCTs for the effect of chemotherapy in low-risk 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 only available for Oncotype DX. Anxiety and HRQoL data associated with testing 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; for each individual test, risk classification probabilities and distant recurrence-free interval estimates have been taken from same source where data permit, which avoids potential 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's economic analyses are subject to several weaknesses: the economic 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 provide a means for considering whether any missing health effects are likely to impact on the conclusions drawn from the economic analysis.

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 USA or through a decentralised testing service for laboratories with next-generation sequencing (NGS) capability. The per-sample pricing of MammaPrint remains the same regardless of testing location. 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 USA, whereas for Prosigna and EPclin, samples are processed in local laboratories.

Suggested research priorities

Research priorities include the following:

- Further studies assessing the ability of all four tests to predict long-term relative chemotherapy benefit in LN+ populations would help to address uncertainty. This may require observational or registry data to assess outcomes across the full range of test scores. In addition, the OPTIMA trial is ongoing, comparing Prosigna test-directed chemotherapy use versus standard chemotherapy use.

- Longer-term studies to further quantify the negative impact of adjuvant chemotherapy using a preference-based instrument would be valuable, to estimate both short-term toxicity and longer-term negative effects, including impacts on fertility in pre-menopausal women.
- Further UK and European studies assessing the impact of tumour profiling tests on recommendations for adjuvant chemotherapy in LN+ populations may reduce uncertainty around clinical impact and cost-effectiveness.
- The integration of tumour profiling tests with decision aid tools to support shared decision-making may constitute a useful research direction.
- The role of tumour profiling tests in older adults, who may be more prone to chemotherapy complications in the context of limited life expectancy, is also a research priority, as is research on test performance in males and in ethnically diverse populations.

Study registration

This study is registered as PROSPERO CRD42023425638.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme (NIHR award ref: NIHR135822) and is published in full in *Health Technology Assessment*; Vol. 29, No. 49. See the NIHR Funding and Awards website for further award information.

Chapter 1 Background and definition of the decision problem

This report includes some material which has been reproduced from the National Institute for Health and Care Excellence (NICE) scope and the protocol for this appraisal. © NICE 2024 Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer. Available from www.nice.org.uk/guidance/dg58. All rights reserved. Subject to Notice of rights (www.nice.org.uk/terms-and-conditions#notice-of-rights). NICE guidance is prepared for the National Health Service in England. All NICE guidance is subject to regular review and may be updated or withdrawn. NICE accepts no responsibility for the use of its content in this product/publication.

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 2016–8, 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 human epidermal growth factor receptor 2 (HER2)-positive breast cancers.

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 number 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 micrometastases depending on whether they are greater than or < 2 mm. 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 used to guide both 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.

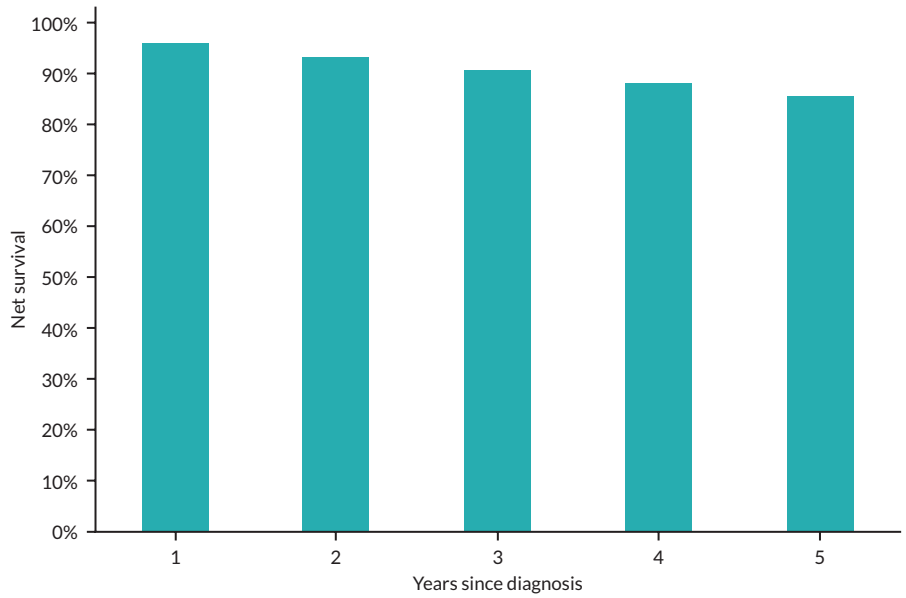


FIGURE 1 Five-year net survival by time since diagnosis, adults with breast cancer diagnosed in 2016–20, followed up to 2021. CDSR, Cochrane Database of Systematic Reviews; CENTRAL, Cochrane Central Register of Controlled Trials; INAHTA, International Network of Agencies for Health Technology Assessment. Source: NHS Digital, National Disease Registration Service.

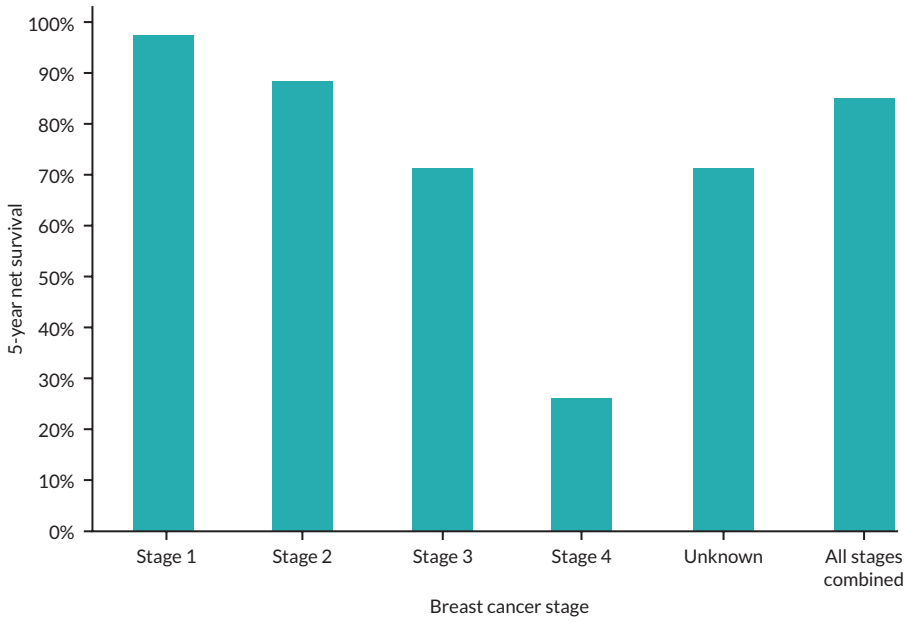


FIGURE 2 Five-year net survival by stage, adults with breast cancer diagnosed in 2016–20, followed up to 2021. Source: NHS Digital, National Disease Registration Service.

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 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 one to three positive lymph nodes (LN1–3).

Epidemiology and incidence

Figure 3 presents estimates of breast cancer incidence by age and sex for the UK, based on data from 2016 to 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.

Significance in terms of ill-health (burden of disease)

Breast cancer is the second most common 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 2017–9, an average of 9509 women and 69 men died from breast cancer in England each year.[7]

Current methods for staging of breast cancer

Breast cancer staging takes into account three main factors: (1) tumour size; (2) metastases to the regional lymph nodes; and (3) 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. 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. The overall TNM stage of the cancer is defined as shown in *Table 1*. 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.

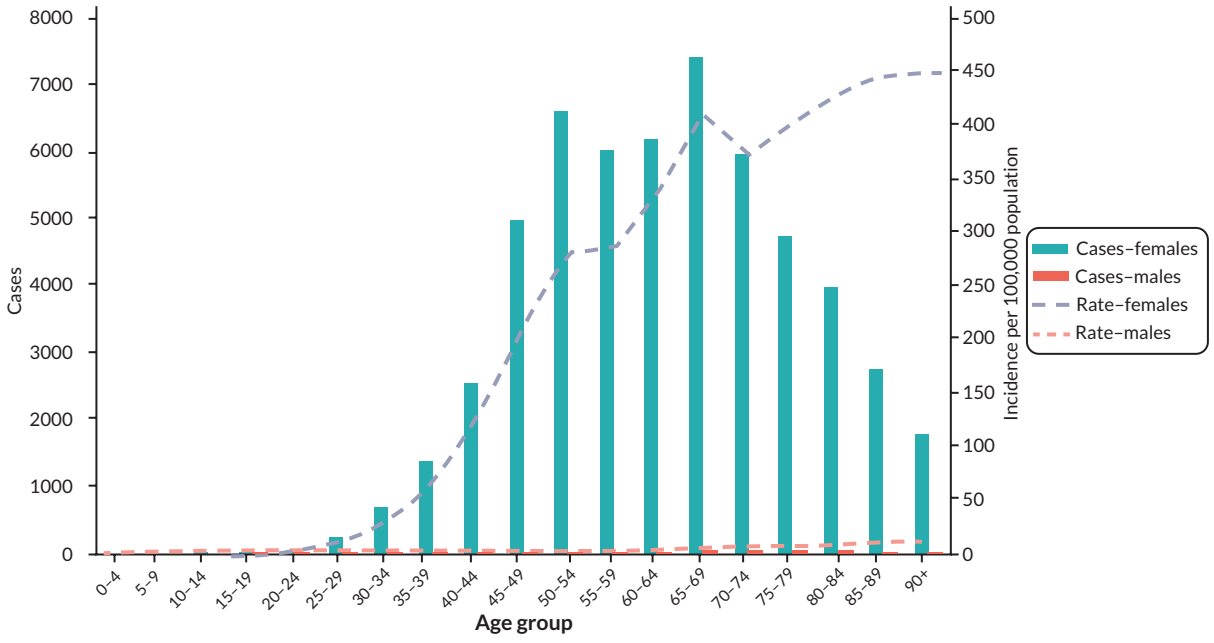


FIGURE 3 Breast cancer incidence by age and sex, 2016–8, UK. Source: Cancer Research UK.

TABLE 1 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

M, metastasis; mi, micrometastases; N, node; T, tumour.

Current service provision

Management of early breast cancer

National Institute for Health and Care Excellence 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).

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 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.

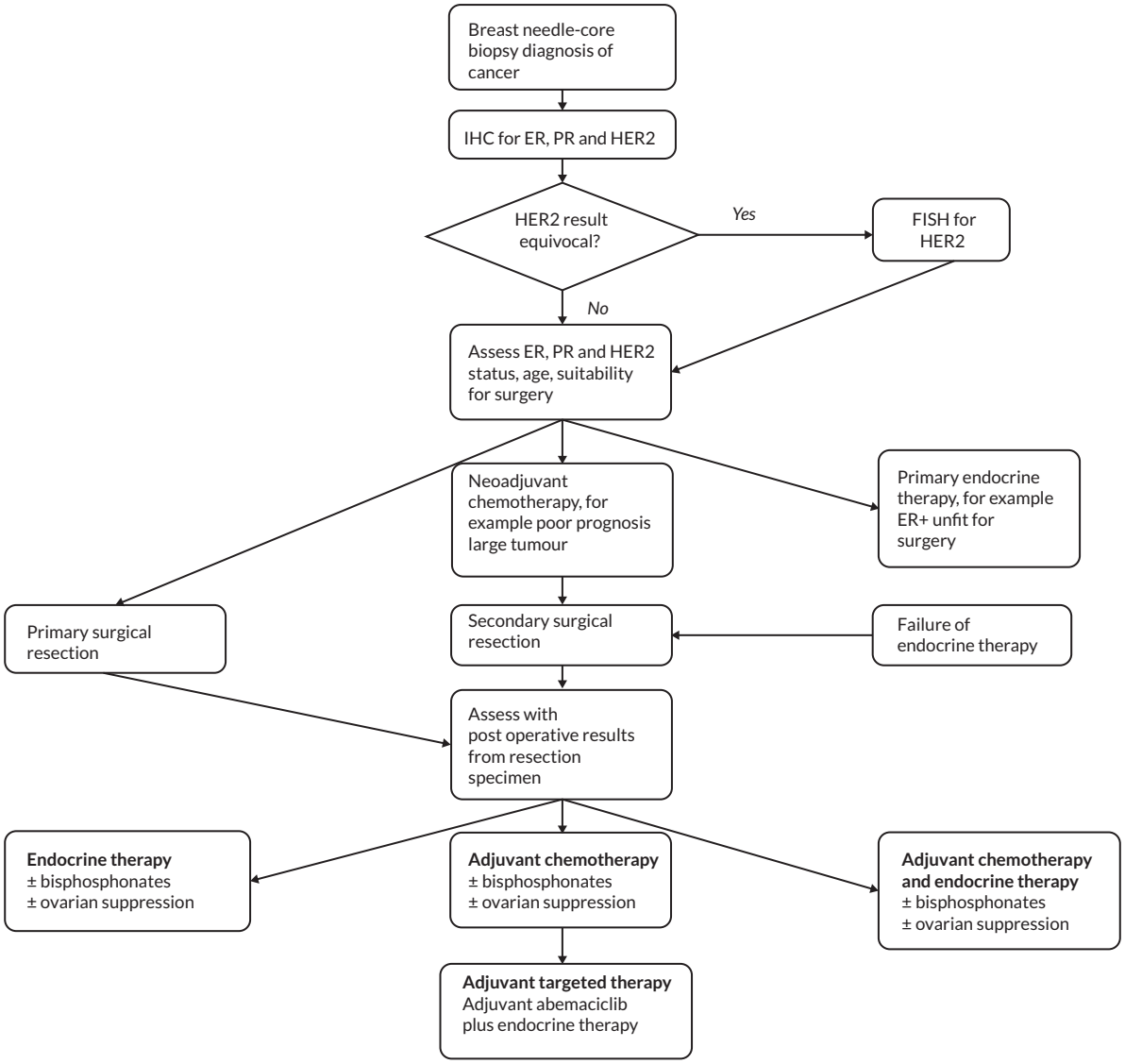


FIGURE 4 Diagnosis and management pathway for early breast cancer. FISH, fluorescence in situ hybridisation; IHC, immunohistochemistry. Notes: Ki67 is tested for following biopsy in some centres; however, the methodology for this is not standardised. Extended endocrine therapy may be given for 10 years. 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 postoperative 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. ER, oestrogen receptor; FISH, fluorescence in situ hybridisation; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; PR, progesterone receptor.

In women who have a sentinel lymph node biopsy, if there is heavy nodal disease then a subsequent clearance is performed. If only micrometastases, isolated tumour cells or just one or two 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.

While 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 *Prognostic risk prediction tools*.

National Institute for Health and Care Excellence Guideline 101[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 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 *Description of technologies under assessment*.

Endocrine therapy

Endocrine therapy 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 ER (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

2–5 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 2–5 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 (LR) 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 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 four positive axillary lymph nodes, or
- one to three 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 inhibitors and capecitabine are only relevant to women with triple negative or HER2+ breast cancer and, as such, are outside 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− 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.

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),[15] Adjuvant! Online (AOL) and PREDICT (University of Cambridge, Cambridge, Cambridgeshire). The factors included in the prediction algorithms and the outcomes predicted by these tools are summarised in *Table 2*.

Nottingham Prognostic Index

The NPI is a composite prognostic parameter involving both time-dependent factors and aspects of biological aggressiveness. The NPI score is calculated using a combination of tumour grade, lymph node involvement and tumour size. The score is calculated as follows: add numerical grade (1, 2 or 3), lymph node score (LN0 = 1, 1–3 nodes = 2, > 3 nodes = 3) and 0.2* tumour size in centimetres. Based on NPI, patients can be divided into three prognostic groups: (1) a good prognostic group (NPI < 3.4); (2) a moderate prognostic group (3.4 < NPI < 5.4); and (3) 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.

TABLE 2 Breast cancer risk prediction tools

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

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

Note

Adjuvant! Online

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 which provides 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)[16] 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 5000 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 (OS) 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.

Description of technologies under assessment

The potential value of tumour profiling tests to guide chemotherapy decisions for women with lymph node-positive early breast cancer

Meta-analyses of randomised clinical trials (RCTs) reported by the 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.[16] 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.[17,18] 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 healthcare 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.[19] Adjuvant chemotherapy may prevent distant recurrence for some women with early breast cancer, while others will not obtain benefit from treatment, with many women remaining recurrence-free at 10 years without chemotherapy.[20] 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.

Summary of tumour profiling tests included in the assessment

Oncotype DX (Exact Sciences)

Oncotype DX Breast Recurrence Score (RS) (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 three 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 RS.

Oncotype DX is offered as a test service to the NHS. Samples are processed centrally at the Exact Sciences centralised laboratory in the USA, which is accredited by the American Association for Laboratory Accreditation and the College of American Pathologists. 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 an RS 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 RS also predicts the benefit of chemotherapy in terms of reducing the risk of distant recurrence. For LN+ disease (one to three positive nodes), the Instructions For Use document states 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 27 February 2023) states that an RS 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 26–100 is inferred to predict substantial chemotherapy benefit.

The Oncotype DX Breast RS results are typically reported within 7–10 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 three positive nodes, or four 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 a 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 three positive nodes), a score of 0–15 indicates low risk, 16–40 indicates intermediate risk and 41–100 indicates high risk. For four or more positive nodes, any score

is assigned high risk. The EAG understands that most people with four 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 three 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 3–5 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 < 3.3 indicates low risk (< 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 < 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 5 cm
- LN0 or LN+ (up to three 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 USA. A decentralised version of the test is also available for local laboratories with next-generation sequencing (NGS) capability. The test requires a 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 < 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 while 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 2–5 years of tamoxifen treatment.

TABLE 3 Summary of tumour profiling tests

Test	Oncotype DX Breast RS	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–IIIa)	Early-stage (stage I–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	RS	Risk category (low, intermediate, high)	Risk category (low, high)	Risk category (low, ultra-low, high)
	Chemotherapy benefit	Intrinsic subtype	Chemotherapy benefit (%)	Chemotherapy benefit
	Probability of distant recurrence (%)	Probability of distant recurrence (%)	Probability of distant recurrence (%)	
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

LN, lymph node; mRNA, messenger ribonucleic acid.

Current usage of tumour profiling tests in the National Health Service

National Institute for Health and Care Excellence 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 immunohistochemical 4 (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.

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 one to three 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 (CEA) (Harnan *et al.*[10]) which informed considerations for the LN+ subgroup within NICE DG34.[13]

Interventions

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

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

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.

Population and important subgroups

The population of interest for this assessment relates to people with ER+ (and/or PR+), HER2−, early-stage breast cancer with one to three 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.[21]

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.

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:

- DRFS, distant recurrence-free interval (DRFI), distant metastasis-free survival (DMFS) and distant metastasis-free interval (DMFI).
- Disease-free survival (DFS) and BCSS.
- 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.

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 one to three 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.

Chapter 2 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.

Methods for clinical review

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 www.crd.york.ac.uk/prospero/display_record.php?RecordID=425638 (accessed 21 August 2023). The review was conducted following the general principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[22]

Inclusion criteria

Population and subgroups

The relevant population is people with ER+ (and/or PR+), HER2−, early-stage breast cancer (stage I, II or IIIA) with one to three 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), while if > 20% were out of scope, then the study was excluded. In particular, studies were required to include ≥ 80% LN+ patients since this was the focus of this updated review. Exceptions to this were that some studies did not report HER2 status, some studies had > 20% ER/PR− patients, and some studies included LN+ patients but > 20% had > 3 positive nodes or > 20% had micrometastases; these studies were included to ensure inclusion of sufficient relevant evidence, but these limitations were noted. Data for subgroups listed in *Description of decision problem* were included, where available.

Interventions

The following interventions were included: Oncotype DX Breast 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.

Comparators

The relevant comparator is current decision-making, which includes clinical and pathological features used to assess risk, and clinicopathological tools outlined in *Current service provision* (current tools include PREDICT and NPI while 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 *Independent External Assessment Group economic analysis*).

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:

- DRFS, DRFI, DMFS and DMFI.
- DFS.
- OS and BCSS.

Studies only reporting LR or LRR were excluded.

The different study types are linked via the EAG's health economic model in order to estimate the clinical effectiveness and cost-effectiveness of the tests (see *Independent External Assessment Group economic analysis*).

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), while studies published from 2017 onwards were identified via an update search. Studies not published in English were considered includable if sufficient data could be extracted.

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.[23–26] The databases, trial registers and websites searched in April 2023 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
- 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
- ClinicalTrials.gov (National Library of Medicine)
- American Society of Clinical Oncology
- European Society for Medical Oncology
- American Association for Cancer Research
- European CanCer Organisation.

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.

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 was 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 *Inclusion criteria*). Any studies causing uncertainty were checked by a second reviewer with involvement of a third reviewer when necessary. Data were extracted into Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) by one reviewer and checked

by a second reviewer. Data from studies published prior to 2017 were extracted from the existing review and checked by a second reviewer.[10]

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).[27] Prognostic and prediction studies were assessed using the Prediction model study Risk Of Bias Assessment Tool (PROBAST);[28] 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.

Methods of analysis/synthesis

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

Results of clinical review: overview

Quantity and type of included studies

The database search for the clinical review identified 4058 articles, of which 502 were checked as full texts and 42 were includable. In addition, 13 further articles were included from the previous review by Harnan *et al.*[10] Therefore, in total, 55 articles were included in the clinical review. Of these, 42 articles related to patient outcomes (prognostic, predictive and prospective use of test), while 13 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*).

Summary of evidence identified for each outcome type 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 CEA. The remainder of the clinical chapter presents the data on risk of bias in included studies (see *Risk of bias in included studies*), prognostic ability of the tests (see *Results: prognostic ability*), prediction of chemotherapy benefit (see *Results: prediction of chemotherapy benefit*), decision impact (see *Results: decision impact*), and HRQoL and anxiety (see *Results: health-related quality of life and anxiety*).

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[29] for Oncotype DX and MINDACT[30] for MammaPrint); these are described below
- reanalyses 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.[31] 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[29] for Oncotype DX and MINDACT[30] for MammaPrint) and are described below.
- Reanalyses 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)[32] and for MammaPrint from a reanalysis of two cohorts (Mook *et al.*, 2009).[33]
- 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 randomised controlled trial of Oncotype DX: RxPONDER

The RxPONDER[29] study of Oncotype DX randomised patients with Oncotype DX RS ≤ 25 to chemotherapy plus endocrine therapy (CET) versus 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 versus 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, that is for patients with an RS of 26–100. In addition, the majority of patients in RxPONDER had only one positive node (65% had one positive node, 25% had two positive nodes, 9% had three positive nodes). Furthermore, patients had knowledge of their RS result before agreeing to be randomised, which may have resulted in selection bias (of 9383 women screened, 4300 were excluded before randomisation, of which 1035 had RS > 25 but the remaining 3265 did not participate for other reasons). Results of RxPONDER are described in this report in *Results: prognostic ability* and *Results: prediction of chemotherapy benefit*. RxPONDER also informs the EAG's economic analyses of Oncotype DX (see *Independent External Assessment Group economic analysis*).

Prospective randomised controlled trial of MammaPrint: MINDACT

The MINDACT[30] 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 versus 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 versus no chemotherapy on patient outcomes (DMFS, DMFI, DFS, OS). Results for chemotherapy versus 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 *Independent External Assessment Group economic analysis*).

Ongoing prospective randomised controlled trial of Prosigna: OPTIMA

The ongoing OPTIMA study[34] is a RCT of test-directed chemotherapy use versus standard chemotherapy use. Included patients have high clinical risk of recurrence and are largely node-positive (one to nine 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, while the standard care (SC) 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 (within six references)[35–40] and seven other European studies.[41–47] 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.

Health-related quality of life 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.

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[29] and MINDACT[30]), assessed using the Cochrane RoB2 tool,[27] scored low risk of bias on all domains, and low risk of bias overall. As noted in *Summary of evidence identified for each outcome type*, 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.[28] 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 (see *Results: prognostic ability*). 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 inclusion 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 (see *Results: prognostic ability*).

For predictive studies, the following factors may have affected results to some extent. Only the SWOG-8814 study[32] was a reanalysis of a 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 (see *Results: prediction of chemotherapy benefit*). 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 (see *Results: prediction of chemotherapy benefit*).

Results: prognostic ability

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, that is, 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[29] for Oncotype DX and MINDACT[30] for MammaPrint).
- Reanalyses 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[20] 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,[20,29,32,48–51] 5 studies of MammaPrint,[30,33,52–55] 6 studies of Prosigna[20,48,56–62] and 5 studies of EPclin.[20,48,59,60,63–65]
- 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[66] in Israel, the Surveillance Epidemiology and End Results (SEER) registry[67–69] in the USA, the National Cancer Database (NCDB)[70–74] in the USA and a few smaller prospective studies.[75–77] 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 4*. 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[29,51] and MINDACT[30,52]) are presented in *Tables 5* and *6*. Data on observational studies of prospective use of Oncotype DX are provided in *Table 7*.

Summary of distribution of genomic risk groups and distant recurrence risk

Table 4 summarises prognostic data from studies reporting 10-year distant recurrence outcomes. For patients receiving endocrine monotherapy, the review identified one study of Oncotype DX,[20] no studies of MammaPrint, two studies of Prosigna[20,56] and three studies of EPclin.[20,63,65] A third study of Prosigna[58] which used different cut-offs is shown in *Table 4* for completeness, but is not included in this textual summary. In terms of distribution, the study of Oncotype DX[20] (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)[20,56] or EPclin (19–35% low risk).[20,63,65] Freedom from distant recurrence at 10 years in the low-risk group was 81% (one study of Oncotype DX),[20] 100% (two studies of Prosigna)[20,56] and 94–100% (three studies of EPclin).[20,63,65] Freedom from distant recurrence at 10 years in the high-risk group was 62% (one study of Oncotype DX),[20] 69–76% (two studies of Prosigna)[20,56] and 70–81% (three studies of EPclin).[20,63,65] Further details of the prognostic data can be found in *Appendix 4* (*Table 49* for Oncotype DX, *Table 50* for MammaPrint, *Table 51* for Prosigna and *Table 52* for EPclin).

Table 4 also presents 10-year distance recurrence data from further studies in which some or all patients received chemotherapy, including one study of Oncotype DX,[49] three studies of MammaPrint,[33,53,54] one study of Prosigna[59,60] and one study of EPclin.[59,60] 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, while freedom from distant recurrence at 10 years ranged from 79% to 95% in the low-risk group and 54–81% in the high-risk group.[33,53,54] Further details of the prognostic data can be found in *Appendix 4*.

Summary of prognostic ability across tests

Table 4 (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 4 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 0–10 years %			[a]Sig prog 10 years?	[a]Sig prog 10 years adj?
										Low	Int	High	Low	Int	High		
Oncotype DX	ET alone	Sestak 2018[20] (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[49] (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[53]	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[33]	Cohort-R	*N* = 241	DMFS	LN1–3 + Lnmicro	79% ER+ 84% HER2–	Pre/post	NR	41	–	59	91	–	76	Y	N
		Vliek 2017[54] (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[20] (TransATAC)	RCT-R	*N* = 183	DRFI	LN1–3	HR + HER2–	Post	16, 40	8	32	60	100	79	69	N	Y
		Gnant 2014[56]/ Filipits 2014[57] (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[58] (DBCG)	Cohort-R	*N* = 1395	DRFS	LN1–3	HR + HER2–	Post	Varies by *N* nodes	26	28	46	97	89	78	Y	Y
	All CT + ET	Martin 2016/14[59,60] (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[20] (TransATAC)	RCT-R	*N* = 183	DRFI	LN1–3	HR + HER2–	Post	3.3	23	–	77	94	–	70	Y	Y
		Filipits 2019[63] (ABCSG-6/8)	RCT-R	*N* = 453	DRFR	LN1–3	ER + HER2–	Post	3.3	35	–	65	96	–	81	Y	Y
		Constantinidou 2022[65]	Cohort-R	*N* = 62	DRFS	LN1–3	ER + HER2–	Pre	3.3	19	–	81	100	–	75	N	Y
	All CT + ET	Martin 2016/14[59,60] (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

Adj, adjusted; cohort-R, cohort reanalysis; CT, chemotherapy; DRFR, distant recurrence-free rate; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; NR, not reported; prog, prognostic; RCT-R, RCT reanalysis; sig, significant.

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).

TABLE 5 Prognostic data from prospective RCT of Oncotype DX (RxPONDER)

Reference Study	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Risk 0–5 years %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low RS ≤ 25		High		
						CT	No			
Oncotype DX: prospective RCT: distant recurrence										
Kalinsky 2021[29] RxPONDER	DRFI (0–5 years)	CT vs. none Prosp RCT *n* = 3353	LN1–3 (65% 1 node, 25% 2 nodes, 9% 3 nodes)	Post-meno	All ≤ 25	94.9	93.9	–	–	–
		n = 1665	100% HR+ 100% HER2–	Pre-meno	All ≤ 25	96.3	93.9	–	–	–
Oncotype DX: prospective RCT: IDFS										
Kalinsky 2021[29] RxPONDER	IDFS (0–5 years)	*n* = 5018	LN1–3	All meno (67% post)	All ≤ 25	92.2	91.0	–	*0–5 years: HR per unit-RS (within RS 0–25, adj for meno and CT): 1.05 (1.04 to 1.07), $p < 0.001$	Y*
			100% HR+	Post-meno	All ≤ 25	91.2	91.9	–	*0–5 years: 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*
			100% HER2–	Pre-meno	All ≤ 25	93.9	89.0	–	*0–5 years: 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 DX: prospective RCT: IDFS by ethnicity										
Abdou 2023[51] RxPONDER	IDFS	*n* = 4015 CT+ET vs. ET	LN1–3	White (*n* = 2833)	All ≤ 25	91.5		–	–	–
		Prosp RCT	100% HR+	Black (*n* = 248)	All ≤ 25	87.0		–	–	–
			100% HER2–							
				Asian (*n* = 324)	All ≤ 25	93.9		–	–	–
				Hispanic (*n* = 610)	All ≤ 25	91.4		–	–	–

Adj, adjusted; CT, chemotherapy; LN, lymph nodes (number positive); meno, menopausal; prosp, prospective; sig, significant; Y, yes.
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.

Prognostic data from prospective randomised controlled trial of Oncotype DX (RxPONDER)

The prospective RCT of Oncotype DX (RxPONDER)[29] randomised patients with an Oncotype DX RS of ≤ 25 to chemotherapy plus ET versus ET monotherapy. The publication mainly focuses on prediction of chemotherapy benefit; this is discussed in *Results: prediction of chemotherapy benefit*. RxPONDER also reports some prognostic data which are presented in *Table 5*. These data are not included in summary *Table 4* because prognostic data were not reported (NR) 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 5*).[29] A further RxPONDER publication[51] 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 5*).

Distant recurrence data in RxPONDER are also shown in *Table 5*. 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[20] and Penault-Llorca 2018,[50] *Appendix 4*).

Prognostic data from prospective randomised controlled trial of MammaPrint (MINDACT)

The prospective RCT of MammaPrint (MINDACT)[30] 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 versus no chemotherapy. All the MINDACT data presented in this chapter 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, while within the clinical low-risk group, 92% were MammaPrint low-risk and 8% were MammaPrint high-risk (*Table 6*).

Outcome data from the MINDACT LN+ subgroup are presented for patients in the different risk groups (see *Table 6*).[30] 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 4*. 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 versus 80.9% for MammaPrint high-risk, with other outcomes showing a similar pattern (see *Table 6*). 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[52] 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 versus no chemotherapy within each group is discussed in *Results: prediction of chemotherapy benefit*.

Ongoing prospective randomised controlled trial of Prosigna: OPTIMA

As described in *Results of clinical review: overview*, the ongoing OPTIMA study[34] is a RCT of Prosigna test-directed chemotherapy use versus standard chemotherapy use. The review did not identify any published results of OPTIMA so far.

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[66] in Israel ($n = 709$),

TABLE 6 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 0–8 years %		HR between groups	Sig?[a] *Adj
						Low	High	Low	High		
MammaPrint: prospective RCT: distant recurrence											
Piccart 2021[30,b] MINDACT	DMFI	*n* = 1176 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)	–[c]	–	–
	DMFS	*n* = 1176 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)	–[c]	–	–
Lopes Cardozo 2022[52] 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[30,b] MINDACT	DFS	*n* = 1176 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[30,b] MINDACT	OS	*n* = 1176 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)	–	–	–

Adj, adjusted; CT, chemotherapy; LN, lymph nodes (number positive); meno, menopausal; prosp, prospective; sig, significant.
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 Piccart 2021 data are from the Piccart *et al.* (2021)[30] supplement, Table S10.
c The mAOL low-risk, MammaPrint high-risk group was not analysed due to small numbers of LN+ patients (*n* = 15).

the SEER registry[67–69] in the USA (*n* = 6483), the NCDB[70–74] in the USA (*n* = 25,029) and a few smaller prospective studies.[75–77] An overview of results is described here, with full results in *Table 53* (see *Appendix 5*), while data on distant recurrence are shown in *Table 7*.

In terms of distribution using cut-offs of RS < 18 and > 30 (see *Appendix 5*), across the Clalit[66] and SEER[67,68] 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[20] (57% low, 32% intermediate, 11% high). A study in younger patients[75] (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[66] and NCDB[70] registries plus a German study,[77] the distribution was in the range of 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[66] and the Young Women's Breast Cancer Study[75]) are shown in *Table 7*. Within Clalit,[66] 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 7*). 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[66] and the Young Women's Breast Cancer Study[75] show a statistically significant prognostic effect (see *Table 7*). However, in older patients (≥ 70 years), there was no statistically significant prognostic effect on 5-year DRFI in Clalit[66] (see *Table 7*).

Data on other outcomes are shown in *Appendix 5*. For BCSS and OS, most analyses of the Clalit,[66] SEER[67,68] and NCDB[70,71,74] 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,[67] while statistically significant prognostic ability was reported in both men and women,[69] though these subgroups were based on small numbers. Analyses of NCBD reported statistically significant prognostic ability in patients aged 40–50 years[72] and in patients with lobular cancer.[73]

Conclusions for prognostic data

For all four tests, within reanalyses 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[66] 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,[29] 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,[30] within LN+ patients at high clinical risk, 8-year DMFI was 92.3% for MammaPrint low-risk versus 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.

Results: prediction of chemotherapy benefit

Overview of predictive data in this report

This section summarises two types of data: (1) the effect of chemotherapy versus no chemotherapy on patient outcomes within a test risk group or range; and (2) whether this effect of chemotherapy versus no chemotherapy differs significantly between test risk groups or ranges, that is whether the test is predictive of chemotherapy benefit, generally assessed via a test for interaction between chemotherapy effect and risk score.[31] 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[29,32,66,71–74,78–84] relating to 5 studies of Oncotype, and 2 publications[30,33] relating to 2 studies of MammaPrint, provided data on prediction and/or effect of chemotherapy.

TABLE 7 Observational data for Oncotype DX (distant recurrence)

Cohort	Reference	Nodal status HR, HER2	Outcome	N ET/CT	Meno Age Clin	Test cut-offs	Distribution %			% risk of outcome			HR between test risk groups (95% CI)	Sig?[a] *Adj
							Low	Int	High	Low	Int	High		
Oncotype: distant recurrence														
Clalit, Israel	Stemmer 2017[66]	LN1mic: 42%	DRFI (0–5 years)	*n* = 709	All meno	18, 30	53	36	10	96.8 (7% CT)	93.7 (40% CT)	83.1 (86% CT)	0–5 years: low vs. high: HR 0.19 (0.09 to 0.40)	Y
		LN1–3 : 58%		Var ET/CT									0–5 years: int vs. high: HR 0.39 (0.20 to 0.79), $p < 0.001$	Y
		100% ER+											*0–5 years: adj HR: low vs. high: HR 0.23 (0.11 to 0.50)	Y*
		100% HER2–											*0–5 years: adj HR: int vs. high: HR 0.42 (0.20 to 0.86), $p = 0.001$	Y*
						11, 25	≤ 25:81	19	95.7 (5% CT)	96.0 (18% CT)	86.9 (77% CT)	0–5 years: $p < 0.001$	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)	0–5 years: $p < 0.001$	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)	0–5 years: $p = 0.017$	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)	0–5 years: $p = 0.458$	N

Cohort	Reference	Nodal status HR, HER2	Outcome	N ET/CT	Meno Age Clin	Test cut-offs	Distribution %			% risk of outcome			HR between test risk groups (95% CI)	Sig?[a] *Adj
							Low	Int	High	Low	Int	High		
Young Women's Breast Cancer Study	Poorvu 2020[75]	LNmic, LN1–3	DRFS (0–6 years)	*n* = 163	Age ≤ 40	18, 30	33	42	25	0–6 years: 85.9 (83% CT)	0–6 years: 87.3 (97% CT)	0–6 years: 62.8 (98% CT)	0–6 years: $p = 0.004$	Y
		100% ER+		Var ET/CT										
		100% HER2–												
						11, 25	9	54	37	0–6 years: 92.3 (79% CT)	0–6 years: 85.2 (92% CT)	0–6 years: 71.3 (97% CT)	0–6 years: $p = 0.10$	N

Adj, adjusted; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); LNmic, lymph node micrometastases; meno, menopausal; N, no; sig, significant; var, variable; Y, yes.

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.

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

- A reanalysis of the SWOG-8814 RCT (Albain *et al.*, 2010),[32] in which Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy versus no chemotherapy. This study did not report distant recurrence but did report data for DFS, BCSS and OS. HRs for chemotherapy versus 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,[29,78] which reported the effect of chemotherapy versus 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,[66] SEER[67–69] and NCDB[70–74] 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)[33] which only reported a *p*-value for an interaction test for BCSS.
- The MINDACT prospective RCT,[30] which reported the effect of chemotherapy versus 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.

Prediction of chemotherapy benefit: randomised controlled trial reanalysis (Oncotype DX)

Albain *et al.* (2010)[32] reported a reanalysis of the SWOG-8814 RCT, in which Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy versus no chemotherapy. This study did not report on outcomes relating to distant recurrence, but did report DFS, BCSS and OS (*Table 8*). HRs for chemotherapy versus 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 versus 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 8*).[32] Similar data are presented in *Table 8* for DFS at different timepoints and for BCSS and OS.

Interaction tests were conducted for chemotherapy effect and risk group: some were statistically significant while others were not (see *Table 8*). 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 NR), but it 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; $p = 0.87$).[32]

Prediction of chemotherapy benefit: prospective randomised controlled trial of Oncotype DX (RxPONDER)

The RxPONDER[29] prospective RCT of Oncotype DX randomised patients with an RS of 0–25 to chemotherapy plus ET versus ET monotherapy (*Table 9*). 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 versus 96.6%

TABLE 8 Prediction of chemotherapy benefit: RCT reanalysis (Oncotype DX)

Study Reference Design	Nodal status HR, HER2	Outcome	*N*	Menopausal status	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)				Interaction	Pred[a] *Adj
						Low		Int		High										
						CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High	Adj		
Oncotype DX: RCT-reanalysis: distant recurrence																				
No data																				
Oncotype DX: RCT-reanalysis: DFS																				
SWOG-8814	LN1–3 : 62%	DFS	*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*
Albain 2010[32]	LN4 +: 38%	0–5 years																		
RCT-R	100% HR+																			
	88% HER2–																			
		DFS 0–10 years	*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 5–10 years	*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	LN1–3 : 62%	BCSS																		
Albain 2010[32]	LN4+: 38%																			
RCT-R	100% HR+																			
	88% HER2–	0–10vyears	*n* = 367	Post-meno	18, 30	–	–	–	–	73	54	–	–	19	*p* = 0.56	*p* = 0.89	*p* = 0.033	Y	–	–

continued

TABLE 8 Prediction of chemotherapy benefit: RCT reanalysis (Oncotype DX) (*continued*)

Study Reference Design	Nodal status HR, HER2	Outcome	*N*	Menopausal status	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)					Pred[a] *Adj
						Low		Int		High										
						CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High	Adj	Interaction	
		OS 0–10 years	*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): 0–10 years: *p* = 0.026 0–5 years: *p* = 0.016 5–10 years: *p* = 0.87	Y* Y* N*

Abs diff, absolute difference; adj, adjusted; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; N, no; pred, predictive of CT benefit; RCT-R, RCT reanalysis; sig, significant; Y, yes.

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.

TABLE 9 Prediction of chemotherapy benefit: prospective RCT of Oncotype DX (RxPONDER)

Study Reference Design	Nodal status HR, HER2	Outcome	*N*	Menopausal status	Test cut-offs	% risk of outcome				Abs diff CT vs. no CT		HR for CT vs. no CT (95% CI)	Interaction	Pred[a] *Adj
						Low RS ≤ 25		High		Low RS ≤ 25	High			
						CT	No	CT	No					
Distant recurrence: full population														
RxPONDER Kalinsky 2021[29] (Kalinsky SABCS 2021 slides[78])[b] Prosp RCT	LN1–3 (65% 1 node, 25% 2 nodes, 9% 3 nodes) 100% HR+ 100% HER2–	DRFS (0–5 years)	*n* = 5018	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 (0–5 years)	*n* = 3353	Post-meno	All ≤ 25	94.4 94.3[b]	94.4 94.8[b]	–	–	0.1 -0.5[b]	–	RS ≤ 25: HR 1.05 (0.81 to 1.37), *p* = 0.70 RS ≤ 25[b]: adj HR 1.12 (0.88 to 1.44), *p* = 0.35	–	–
		DRFI (0–5 years)	*n* = 3353	Post-meno	All ≤ 25	95.8[b]	96.6[b]	–	–	−0.8[b]	–	RS ≤ 25[b]: adj HR 1.12 (0.82 to 1.52), *p* = 0.49	–	–
Distant recurrence: pre-menopausal														
RxPONDER	LN1–3 100% HR+ 100% HER2–	DRFS (0–5 years)	*n* = 1665	Pre-meno	All ≤ 25	96.1 95.9[b]	92.8 93.4[b]	–	–	3.3 2.5[b]	–	RS ≤ 25: HR 0.58 (0.39 to 0.87), *p* = 0.009 RS ≤ 25[b]: adj HR 0.66 (0.45 to 0.97), *p* = 0.033	–	–
		DRFI (0–5 years)	*n* = 1665	Pre-meno	All ≤ 25	96.3[b]	93.9[b]	–	–	2.4[b]	–	RS ≤ 25[b]: 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 (0–5 years)	*n* = 5018	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

continued

TABLE 9 Prediction of chemotherapy benefit: prospective RCT of Oncotype DX (RxPONDER) (*continued*)

Study Reference Design	Nodal status HR, HER2	Outcome	*N*	Menopausal status	Test cut-offs	% risk of outcome: Low RS ≤ 25 CT	% risk of outcome: Low RS ≤ 25 No	% risk of outcome: High CT	% risk of outcome: High No	Abs diff CT vs. no CT: Low RS ≤ 25	Abs diff CT vs. no CT: High	HR for CT vs. no CT (95% CI)	Interaction	Pred[a] *Adj
IDFS: post-menopausal														
RxPONDER	LN1–3 100% HR+ 100% HER2–	IDFS (0–5 years)	*n* = 3353	Post-meno	All ≤ 25	91.3 91.2[b]	91.9 91.9[b]	–	–	−0.6 −0.7[b]	–	RS ≤ 25 : HR 1.02 (0.82 to 1.26), *p* = 0.89 RS ≤ 25[b] : 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 (0–5 years)	*n* = 1665	Pre-meno	All ≤ 25	93.9 93.9[b]	89.0 89.0[b]	–	–	4.9 4.9[b]	–	RS ≤ 25 : HR 0.60 (0.43 to 0.83), *p* = 0.002 RS ≤ 25[b] : 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)	–	–

Abs diff, absolute difference; adj, adjusted; CT, chemotherapy; LN, lymph nodes (number positive); meno, menopausal; N, no; prosp, prospective; pred, predictive of CT benefit; Y, yes.

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.[78]

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 versus 93.9% with no chemotherapy, an absolute difference of 2.4% favouring chemotherapy (adjusted HR 0.64; 95% CI 0.43 to 0.95; $p = 0.026$).[29] 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 9*).

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 9*).[29] 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 9*). 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.

Prediction of chemotherapy benefit: reanalysis of cohort (MammaPrint)

In terms of prediction of CT benefit, the only data identified for MammaPrint were a reanalysis of two cohorts (Mook *et al.*, 2009)[33] 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 (*Table 10*).

Chemotherapy effect within groups: prospective randomised controlled trial of MammaPrint (MINDACT)

As noted in *Results: prognostic ability*, the prospective RCT of MammaPrint (MINDACT)[30] randomised patients to chemotherapy versus no chemotherapy if they had a discordant genomic risk (via MammaPrint) and clinical risk (via mAOL). Data were presented for the effect of chemotherapy versus 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.[30] This is consistent with the company's focus on the clinical high-risk group in the Agendia submission to NICE.[85]

Within the clinical high-risk, MammaPrint low-risk, LN+, HR+, HER2− subgroup, 8-year DMFS was 91.2% with chemotherapy versus 89.9% with no chemotherapy, an absolute difference of 1.3% favouring chemotherapy,[30] with a non-significant HR (HR 0.84; 95% CI 0.51 to 1.37; p = NR; *Table 11*). 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 11*).[30]

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.

Effect of chemotherapy within recurrence score 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[66,79] in Israel ($n = 709$), the SEER registry[80] in the USA ($n = 2588$) and the NCDB[71–74,81–84] in the USA ($n = 28{,}591$). However, use or non-use of chemotherapy was not randomised, and it 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 *Table 54* (see *Appendix 6*). Data on distant recurrence are shown in *Table 12*, while *Table 13* 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 versus no chemotherapy are only reported for the Clalit registry[66,79] (SEER only reports BCSS while NCBD only reports OS). Within Clalit (see *Table 12*), using the cut-offs of RS < 18 and > 30, the

TABLE 10 Prediction of chemotherapy benefit: reanalysis of cohort (MammaPrint)

Cohort	Reference	Nodal status HR, HER2	Outcome	*N*	Menopausal status	Test cut-offs	% risk of outcome				Abs diff CT vs. no CT		HR for CT vs. no CT		Interaction	Pred[a] *Adj
							Low		High							
							CT	No	CT	No	Low	High	Low	High		
MammaPrint: cohort reanalysis: BCSS																
NKI, Italy, VdV Cohort-R	Mook 2009[33]	LN1micro to LN3 79% ER+ 84% HER2−	BCSS 0–10 years	*n* = 347	All meno	NR	–	–	–	–	–	–	–	–	Int *p* = 0.95 (adj)	N*

Abs diff, absolute difference; adj, adjusted; cohort-R, cohort reanalysis; CT, chemotherapy; LN, lymph nodes (number positive); meno, menopausal; N, no; pred, predictive of CT benefit.
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.

TABLE 11 Chemotherapy effect within risk groups: prospective RCT of MammaPrint (MINDACT)

Study Reference Design	Nodal status HR, HER2	Outcome	*N*	Clinical risk	Test cut-offs	% risk of outcome				Abs diff CT vs. no CT		HR for CT vs. no CT (95% CI)		Interaction	Pred[a] *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[30,b] Prosp RCT	LN1–3 100% HR+ 100% HER2−	DMFS (0–8 years)	*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 (0–8 years)	*N* = 658	High mAOL[c]	> 0 low, ≤ 0 high	92.3	90.9	–	–	1.4	–	–	–	–	–
MammaPrint: prospective RCT: other outcomes															
MINDACT Piccart 2021[30,b] Prosp RCT	LN1–3 100% HR+ 100% HER2−	DFS (0–8 years)	*N* = 658	High mAOL[c]	> 0 low, ≤ 0 high	85.3	82.8	–	–	2.5	–	–	–	–	–
		OS (0–8 years)	*N* = 658	High mAOL[c]	> 0 low, ≤ 0 high	95.5	94.9	–	–	0.6	–	–	–	–	–

Abs diff, absolute difference; adj, adjusted; CT, chemotherapy; LN, lymph nodes (number positive); meno, menopausal; prosp, prospective; pred, predictive of CT benefit; sig, significant.
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.* (2021)[30] supplement, Table S10.
c The mAOL low-risk, MammaPrint high-risk group was not analysed due to small numbers of LN+ patients.

TABLE 12 Effect of chemotherapy within risk groups: registry data for Oncotype DX (distant recurrence)

Cohort	Reference	Nodal status HR, HER2	Outcome	*N*	Menopausal status	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)			Inter-action	Pred[a] *Adj
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High		
Oncotype DX: observational registry: distant recurrence																				
Clalit, Israel	Stemmer 2017[66]	LN1mic : 42% LN1–3 : 58% 100% ER+ 100% HER2–	DRFI 0–5 years	*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)[79]	LN+ 100% ER+ 100% HER2–	DRFS 0–7 years	*n* = 140	All meno	All 26–30	–	–	–	–	89.4	78.0	–	–	11.4	–	–	Not sig	–	–

Abs diff, absolute difference; adj, adjusted; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; pred, predictive of CT benefit; sig, significant.

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.

TABLE 13 Chemotherapy effect within risk groups: registry data for Oncotype DX (post-menopausal or older age groups)

Cohort	Reference	Nodal status HR, HER2	Outcome	*N*	Age, Clinical factors	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)			Inter-action	Pred[a] *Adj
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High		
Oncotype: observational registry: OS																				
NCDB	Cao 2022 (abst)[82]	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[71]	LN1–3 : 97% LN4–9 : 3% 100% HR+ 100% HER2–	OS 0–5 years	*n* = 8886	Age > 50	All 11–25	–	–	–	–	–	–	–	–	–	–	Adj: 0.64 (0.47 to 0.86), *p* = 0.004	–	–	–
NCDB (cont)	Weiser 2022[73]	LN1–3 100% HR+ 100% HER2–	OS 0–5 years	NR	Age 50–75 Ductal	All ≤ 25	–	–	–	–	–	–	–	–	–	Adj: 1.12 (0.86 to 1.46)	–	–	–	–
NCDB (cont)	Weiser 2021[74]	LN1–3 100% HR+ 100% HER2–	OS 0–5 years	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	–	–	–

Abs diff, absolute difference; adj, adjusted; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; pred, predictive of CT benefit; unadj, unadjusted.
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.

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*.

Effect of chemotherapy for older patients with recurrence score ≤ 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 13*. 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 ($p = 0.006$),[74] patients aged > 50 years with an RS of 11–25 ($p = 0.004$)[71] and patients aged > 50 years with an RS of 20–25 ($p = 0.019$).[82] 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)[73] and patients aged > 70 years with an RS of ≤ 25 ($p = 0.69$).[74] When also considering the limitations of these studies, the results do not clearly either support or refute the RxPONDER findings.

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,[32] 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[29] 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 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[71,73,74,82] 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[33] 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,[30] within the mAOL high-risk, MammaPrint low-risk, LN+, HR+ HER2− subgroup, 8-year DMFS was 91.2% with chemotherapy versus 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.

Results: decision impact

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, 13 publications[35–47] relating to 12 studies reported decision impact data for Oncotype DX in a LN+ population. These included five UK studies[35–40] and seven other European (non-UK) studies[41–47] (*Table 14*). All studies included a combination of patients in both pre-menopausal and post-menopausal

TABLE 14 Decision impact: Oncotype DX (not split by test risk group)

Reference Country	Study, setting Years	HR, HER2	Nodal status Clinical risk	Recom/ decision	Menopausal 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)[35] UK	PONDx; 30 centres 2017–8	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)[36] Holt 2024[40] UK	14 centres 2017–22	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[37] UK	Greater Manchester (NR centres) 2012–5	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[38] UK	Norfolk and Norwich (1 centre) 2014–20	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)[39] UK	Oxford + Swansea (2 centres) 2013–9	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[47] Germany	15 centres 2010–1	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[41] Italy	PONDx; 27 centres 2016–7	ER+ HER2–	LN1–3	R-R	All (55% post)	All RS	–	414	–	–	–	–	258 (62%)	110 (28%)	−148 (−55%)
Dieci 2019[43] Italy	ROXANE; 9 centres 2017–8	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%)

Reference Country	Study, setting Years	HR, HER2	Nodal status Clinical risk	Recom/ decision	Menopausal 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
Dieci 2018[42] Italy	Breast DX, 9 centres 2014–6	ER+ HER2–	LN1–3 Int clin risk	R-R	All (55% post)	All RS	–	126	49 (39%)	5 (4%)	52 (41%)	20 (16%)	72 (57%)	57 (45%)	−15 (−12%)
				R-D	All (55% post)	All RS	–	126	–	–	–	–	72 (57%)	54 (43%)	−18 (−14%)
Zambelli 2020[46] Italy	BONDX (4 centres, Lombardy) 2017–8	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)[44] Spain	9 centres (Galicia) 2013–8	HR+ HER2–	LN1–3 (inc micromets)	R-R	All (50% post)	All RS	–	229	–	–	–	–	159 (69%)	59 (26%)	−100 (−44%)
Llombart-Cussac 2023[45] Spain	KARMA Dx (8 centres) 2016–7	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; LN, lymph nodes (number positive); meno, menopausal; Pre/post-RxP, pre/post publication of RxPONDER; R, recommendation.

stages, except for one study[37] which exclusively focused on post-menopausal patients. No UK and European studies assessed the decision impact of MammaPrint, Prosigna or EPclin in a LN+ population.

Prior to testing, patients were allocated 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). Post-testing, patients fell into four groups: those whose decision or 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 before and after testing, and as the net change in chemotherapy use (see *Table 14*).

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 five UK studies[35–40] and 38% to 100% among seven European (non-UK) studies.[41–47] The total proportion allocated to post-test chemotherapy ranged from 13% to 31% among five UK studies[35–40] and 20% to 57% among seven European (non-UK) studies[41–47] (see *Table 14*).

Among the five UK studies,[35–40] the net reduction in chemotherapy recommendations (pre-test to post-test) was 28%,[38] 37%[35] and 75%[39] across three studies, and the net reduction in chemotherapy decisions was 41%,[35] 52%[40] and 69%[37] across three studies (see *Table 14*). Two of these studies[37,39] assessed only patients with an initial recommendation for chemotherapy and so it may be misleading to calculate the absolute change. Also, in two studies,[37,38] the post-test decisions were based entirely on the test result and so their findings are less reliable. Across seven European studies,[41–47] the net reduction in chemotherapy recommendations (pre-test to post-test) ranged from 12%[42] to 73%.[45] Two of these studies also reported changes from pre-test chemotherapy recommendation to post-test decision with a net reduction of 14%[42] and 29%[47] in chemotherapy use.

Decision impact results for Oncotype DX: by risk group

Of the 12 Oncotype DX studies, 2 UK studies[36,37,40] and 4 European studies[43,45–47] presented data by Oncotype DX risk groups (*Table 15*). Five studies[36,37,40,45–47] used RS 18 and 30 cut-offs, while two studies used the newer cut-offs of RS 11 and 25 (one study[43]) or RS 13 and 25 (one study[36,40]).

Among the studies that used the cut-offs of RS 18 and 30,[36,37,40,45–47] the net change in chemotherapy recommendations or decisions (pre-test to post-test) was as follows: a decrease of 20%,[46] 68%,[40] 91%[45] and 93%[37] in the RS 0–17 risk group; a decrease of 19%,[46] 35%,[40] 37%[37] and 54%[45] in the RS 18–30 risk group; and either a 17%[37] decrease, no change,[45,46] or a 1.7%[40] increase in the RS > 30 risk group.

In the study that used cut-offs of RS 11 and 25,[43] 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),[36,40] 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.

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 28–75% across five Oncotype DX studies.[35–40] The net changes across European (non-UK) studies[41–47] were reductions of 12%[42] to 73% for Oncotype DX. Within studies reporting data by Oncotype DX risk group, there were greater reductions in chemotherapy recommendations in the low-risk and intermediate-risk groups than in the high-risk groups.

TABLE 15 Decision impact: Oncotype DX (results by test risk group)

Reference 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)[36] Holt 2024[40] UK	14 centres 2017–22	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[37] UK	Greater Manchester (NR centres) 2012–5	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[47] Germany	15 centres 2010–1	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[45] Spain	KARMA Dx (8 centres) 2016–7	ER+ HER2–	LN1–3 High clinical 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[46] Italy	BONDX (4 centres, Lombardy) 2017–8	ER+ HER2–	LN1–3 Int clinical 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

continued

TABLE 15 Decision impact: Oncotype DX (results by test risk group) (*continued*)

Reference 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 11 and 25															
Holt 2023 (abst)[36] Holt 2023 (unpub) UK	14 centres 2017–22	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%)
							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[43] Italy	ROXANE; 9 centres 2017–8	HR+ HER2–	LN1–3 Most high clinical 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%)

Reference 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
						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; LN, lymph nodes (number positive); meno, menopausal; Pre/post-RxP, pre/post publication of RxPONDER; R, recommendation.

Results: health-related quality of life and anxiety

Overview of data on health-related quality of life and anxiety

No studies (or subgroups) were identified which assessed HRQoL or anxiety associated with the 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*.

Overview of data on health-related quality of life and anxiety in a lymph node-negative or mixed population

Oncotype

Of two studies of Oncotype DX in LN0/LN+ patients in the USA, one (Evans *et al.*, 2016)[86] 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)[87] 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)[88] 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[89] in LN0/LN+ patients in England reported a statistically significant decrease in anxiety (measured via the STAI) for those whose treatment decision changed from chemotherapy to no chemotherapy on the basis of EndoPredict ($p = 0.045$), and an increase in anxiety (via the 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)[90] and Germany (Wuerstlein *et al.*, 2016).[91] 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[90] reported no change in overall scores, whereas the other[91] reported a statistically significant change in emotional and physical well-being ($p = 0.030$, $p = 0.005$, respectively).

Conclusions for health-related quality of life 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, while 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. It is unclear how far the results of these studies can be generalised to a LN+ population.

Chapter 3 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 (see *Review of existing economic analyses*), a summary and critique of the economic models submitted to NICE by the test manufacturers (see *Review and critique of economic analyses of tumour profiling tests submitted by the test manufacturers*) and the methods and results of an independent economic analysis undertaken by the EAG (see *Independent External Assessment Group economic analysis*). A discussion of the key issues around the cost-effectiveness of the tumour profiling tests is presented in *Discussion*. Details of the review and planned economic analyses can be found in the final EAG protocol (www.nice.org.uk/).

Review of existing economic analyses

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 and Other Non-Indexed Citations: Ovid, 1946–present.
- EMBASE: Ovid, 1974–present.
- Science Citation Index Expanded (SCI-E): Web of Science, 1900–present.
- Conference Proceedings Citation Index – Science (CPCI): Web of Science, 1990–present.

The search strategies comprised medical subject heading (MeSH) or Emtree Thesauri terms and free-text synonyms for: (1) 'tumour profiling tests' and 'breast cancer' and (2) '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 CEA report[85] [provided as part of the Agendia company submission (CS)], the Exact Sciences CS[23] and RFI documents provided to NICE by Veracyte[26] and Myriad[25] 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.

Cost-effectiveness review results: summary of included studies

A PRISMA flow diagram summarising study selection is presented in *Appendix 7*. 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.[10,92–95] Seven further studies[96–102] 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 hand-searching the reference lists for the systematic literature reviews (SLRs) reported in the submission from Exact Sciences,[23] the Cytel CEA report[85] or from the information provided to NICE by Veracyte and Myriad.[25,26] The scope addressed within the 12 included studies and the key aspects of the modelling approaches are summarised *Tables 16* and *17*, respectively.

The 12 included economic studies were undertaken to reflect a range of settings, including the UK, the USA, Canada and Germany. Most of the included studies adopted a direct healthcare 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[93,95] 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[93] 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,[101] 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: (1) relapse-free; (2) DM; and (3) 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 LR.[97] 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[95] 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 among 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, while 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[29] and longer-term follow-up data from MINDACT[30] – 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.

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).

TABLE 16 Existing economic evaluations – analytic scope

Author	Population	Per cent LN+	Age	Intervention	Comparator	Country	Perspective	Time horizon	Discount rate (%)
Berdunov *et al.* (2021)[92]	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 and PSS	Lifetime	3.5
Hinde *et al.* (2019)[93]	Women with ER+, HER2− EBC	95%[a]	Mean age 56.5 years	EndoPredict (EPclin)	Standard risk tools only	UK	NHS	Lifetime	3.5
Masucci *et al.* (2019)[94]	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 + sub-group	Mean age 58 years	Oncotype DX, EPclin, Prosigna, IHC4 + C, MammaPrint	Current practice	UK	NHS ans PSS	Lifetime	3.5
Hall *et al.* (2017)[95]	Women aged 40 years or older with ER+, HER2−, clinically high risk (1–9 axillary lymph nodes, or LN0 with a tumour size ≥ 30 mm) 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)[96]	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)[97]	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)[98]	Patients with ER+, HER2−, LN0 or LN+ (up to 3 nodes) ESBC.	100% in LN + sub-group	Mean age 56.3 years	Oncotype DX	Conventional diagnostic procedures	Germany	Health care payer	30 years	3.0
Lamond *et al.* (2012)[99]	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)[100]	LN+, ER+ ESBC	100%	Baseline age 60 years	Oncotype DX	SC (chemother-apy for all)	UK	NHS	Lifetime (up to maximum age 100 years)	3.5
Wong *et al.* (2012)[101]	Women with LN+ HR+ breast cancer (1–3 nodes)	100%	Reflective of RxPONDER	Oncotype DX	Current care (US NCCN guidelines)	USA	Payer	Lifetime (40 years)	3.0

continued

TABLE 16 Existing economic evaluations – analytic scope (*continued*)

Author	Population	Per cent LN+	Age	Intervention	Comparator	Country	Perspective	Time horizon	Discount rate (%)
Vanderlaan *et al.* (2011)[102]	Minimally LN+, ESBC	100%	Mean age 62 years	Oncotype DX	Current care (US NCCN guidelines)	USA	US payer (managed care) perspective	30-years	3.0

BC, base case; EBC, early breast cancer; ESBC, early-stage breast cancer; LN, lymph node; NCCN, National Comprehensive Cancer Network; NHS, National Health Service.

a 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.

TABLE 17 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)[92]	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)[93]	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)[94]	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 FN; (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)[95]	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) LR; (4) disease-free after LR; (5) HF; (6) dead.
Stein *et al.* (2016)[96]	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) LR; (4) disease-free after LR; (5) CHF; (6) CML; (7) dead.
Hannouf *et al.* (2014)[97]	Markov	1 month	Classification to low-, intermediate- and high risk with separate Markov nodes for CT + ET vs. ET alone (accounting for chemotherapy-related AEs)	Unclear – appears to assume predictive benefit	Model assumes 50% IR patients receive chemotherapy	ET only model – 5 states: (1) remission; (2) LR; (3) distant recurrence; (4) dead. CT + ET model – 5 states: (1) remission with chemotherapy SAEs; (2) remission without chemotherapy SAEs; (3) LR; (4) distant recurrence; (5) dead.
Blohmer *et al.* (2013)[98]	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

continued

TABLE 17 Existing economic evaluations – modelling approach and assumptions regarding predictive benefit and chemotherapy use (*continued*)

Author	Model approach	Cycle length	Model type	Does model claim predictive benefit for test?	Assumptions on chemotherapy use	Long-term health states
Lamond *et al.* (2012)[99]	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)[100]	Decision tree and modified Markov model	NR	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) LR; (4) disease-free after LR; (5) CHF; (6) dead.
Wong *et al.* (2012)[101]	Decision tree with partitioned survival approach to determine sojourn time	NR	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)[102]	Appears to be Markov	NR	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.

CHF, congestive heart failure; CINV, chemotherapy-induced nausea and vomiting; CT, chemotherapy; MDS, myelodysplastic syndromes; NCCN, National Comprehensive Cancer Network; NCRAS, National Cancer Registration and Analysis Service; RFS, relapse-free survival; SAE, serious adverse event; SWOG, Southwest Oncology Group.

Exact Sciences model summary and critique (Oncotype DX)

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.[23]). The company also provided responses to clarification questions from the EAG in June 2023,[103] 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.*[92] (see Cost-effectiveness review results) 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,[78] 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[23] 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: (1) the overall ER+, HER2−, LN+ (one to three nodes) early breast cancer population; (2) pre-menopausal women with ER+, HER2−, LN+ early breast cancer; and (3) 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: (1) recurrence-free; (2) DM; (3) AML; and (4) 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.[32,78] 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.[32,78]

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[104] 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 while 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 18*.

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,[78] 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[20] and SWOG-8814[32]). 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 vs. 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 1 year.
- Health outcomes and costs differ between pre-menopausal and post-menopausal women, as reflected in the subgroup analyses of RxPONDER.[78]

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 19*, together with brief comments from the EAG.

Model evaluation methods

The Exact Sciences CS[23] 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 18*.

Results of the Exact Sciences model

The results presented in the Exact Sciences CS[23] are summarised in *Table 20*. 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 £9355 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.

External Assessment Group critique of the Exact Sciences model

The EAG's main concerns regarding the Exact Sciences model are summarised in *Box 1*. These concerns are discussed in detail in the subsequent sections.

TABLE 18 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 • Clinical-pathological tools alone	RxPONDER,[78] TransATAC,[20] SWOG-8814[32]	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),[78] TransATAC,[20] SWOG-8814[32]	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),[78] TransATAC,[20] SWOG-8814[32]	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) Data for comparators reflect exclusively or mostly post-menopausal patients	• Oncotype DX • MammaPrint • EndoPredict EPclin • Prosigna • Clinical-pathological tools alone All analyses presented as pairwise comparisons of test vs. clinical–pathological tools alone	RxPONDER,[78] TransATAC,[20] SWOG-8814,[32] MINDACT,[30] EBCTCG[16]	Predictive benefit assumed for Oncotype DX	Relevant to decision problem set out in final NICE scope.[11]

CS, company's submission; LN, lymph node; SWOG, Southwest Oncology Group.

TABLE 19 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.[23]
Test risk classification probabilities	RS 0–13 and RS 14–25: RxPONDER[29] RS > 25: Number of women who were excluded from RxPONDER[29] 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[29] had an RS of 0–25.
DRFI probabilities for ET alone	RS 0–25: RxPONDER, ET arm[29] RS > 25: TransATAC[20]	Use of external data from TransATAC[20] is necessary because patients with an RS of > 25 were excluded from RxPONDER.[29]
Risk tapering	Ward *et al.*[105] and expert opinion	Same as DG34 model up to 15 years.
Chemotherapy probability under current decision-making	Holt *et al.*[18]	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.*[18]	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[29]	The inclusion of the HR from SWOG-8814[32] for the RS > 25 group indirectly introduces an assumption of predictive benefit for Oncotype DX.
	RS > 25: HR from SWOG-8814[32] applied to ET risk from TransATAC[20]	
Death risk with DM	Abemaciclib plus fulvestrant arm of MONARCH2 trial[106]	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.*[107]	The Exact Sciences CS[23] justifies the use of Petrelli *et al.*[107] on the basis that it is more recent than Wolff *et al.*[108] However, Wolff *et al.* is more recent than Petrelli *et al.*
Death risk with AML	NICE TA552[109] (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.*[110]	Same as DG34 model.
HRQoL parameters		
Utility, recurrence-free	Lidgren *et al.*[111]	Same as DG34 model.
Utility, DM	Lidgren *et al.*[111]	Same as DG34 model.
Utility, AML	NICE TA552[109]	Consistent with source of modelled AML mortality risk.
QALY loss, chemotherapy	Campbell *et al.*[112]	Same as DG34 model.
QALY loss, LR	Campbell *et al.*[112]	Same as DG34 model.
Utility age adjustment	Ara and Brazier[104] (banded estimates)	Values reported by Hernández-Alava *et al.*[113] 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 agreed 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.

TABLE 19 Key evidence sources used to inform the Exact Sciences base-case analyses (overall LN+ population) (*continued*)

Parameter group	Source	EAG comments
ET usage	Ward *et al.*[105]	Based on DG34 model.
Treatments for DM	Clinical opinion, Kurosky *et al.*[114] and MONARCH2 trial[106]	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.*[115]	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[116]	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,[117,118] eMIT,[119] BNF,[120] PSSRU[121]	Appropriate sources applied
Cost of death	Hinde *et al.*[93]	–

BSC, best supportive care; BNF, *British National Formulary*; CDK4/6i, cyclin-dependent kinase 4/6 inhibitor; DM, distant metastases; eMIT, electronic Market Information Tool; FEC, fluorouracil, epirubicin and cyclophosphamide; FEC-D, fluorouracil, epirubicin and cyclophosphamide followed by docetaxel; LN, lymph node; PSSRU, Personal Social Services Research Unit; SWOG, Southwest Oncology Group.

TABLE 20 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 vs. clinical–pathological tools alone	Oncotype DX dominates clinical–pathological tools alone. MammaPrint: ICER vs. clinical–pathological tools alone = £50,626 per QALY gained EPclin: ICER vs. clinical–pathological tools alone = £9355 per QALY gained Prosigna: ICER vs. clinical–pathological tools alone = £41,773 per QALY gained	Not presented	Not presented

LN, lymph node; NMB, net monetary benefit; WTP, willingness to pay.

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.

Uncertainty surrounding the predictive benefit of Oncotype DX

All of the economic analyses of Oncotype DX presented in the Exact Sciences CS[23] are informed by RxPONDER[78] for patients with an RS of 0–25 and by TransATAC[20] and SWOG-8814[32] 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 *Conclusions for prediction of chemotherapy benefit data*, 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.[20] This type of analysis is presented by Berdunov *et al.*[92] but is not included in the Exact Sciences CS;[23] 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.

Analyses presented for the overall lymph node-positive 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 20*). 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[23] (p. 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,[29] 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,[29] 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.

Uncertainty around the probability of being in the Oncotype DX recurrence score > 25 group

The Exact Sciences model assumes that 11% of women in the target population will have an Oncotype RS of > 25. This estimated proportion 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* = 1035) and the overall number of women who were registered for screening in RxPONDER as the denominator (*N* = 9383).[29] Kalinsky *et al.*[29] report that a total of 4300 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* = 1035); declined to participate (*N* = 2372); had recurrence (*N* = 23) or had other or unknown reason (*N* = 622). It is likely that some of the 3265 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 of > 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,[103] 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.[29] The company's response also highlights that the proportion of women with an RS of > 25 in RxPONDER is broadly similar to that in Stemmer *et al.*[122] and SEER.[80]

Uncertainty around relevant cut-offs for National Institute for Health and Care Excellence decision-making

The Exact Sciences CS[23] 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 risk or high risk) based on an RS cut-off of 25. However, Table 2B of the Instructions For Use document for Oncotype DX[123] 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[29] (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,[103] 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.

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[23] 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.*[124]) 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 OS estimate of 9.6 months (based on the liposomal cytarabine–daunorubicin arm of Study 301[125]). 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 500 mg/m^2. The EAG's clinical advisors commented that doxorubicin should have been assumed to be given at a dose of 50 mg/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 utility values by age band using Ara and Brazier.[104] 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.*[113]
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.[16] 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.

Additional analyses undertaken using the Exact Sciences model

As part of their response to clarification questions from the EAG,[103] the company provided an updated version of the model which addresses errors (a) and (b) listed in External Assessment Group critique of the Exact Sciences model. The company's written response[103] 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.

Agendia model summary and critique (MammaPrint)

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[85]). The company also provided responses to clarification questions from the EAG in May and June 2023,[126] which included an updated version of the economic model. The Cytel CEA report (p. 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[85] 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 2021–2 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[85] presents a base-case analysis and five additional scenario analyses. The modelled population and the comparators under consideration differ between these analyses (*Table 21*). 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+ (one to three 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.[128] 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 21*). 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 risk 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: (1) recurrence free; (2) DM; (3) AML; and (4) 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 21*). 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 vs. 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.[112] The model includes age adjustment of utility values based on Ara and Brazier.[104]

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 while patients are receiving chemotherapy (applied in the first 6 months only).
- Additional resource use while patients remain recurrence-free (up to 3 years).
- Treatments for LRR (costed per local/contralateral recurrence event).
- Treatments for DM.
- Treatments for AML.
- End-of-life care.

TABLE 21 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. LN0 NPI > 3.4 and LN+ patients weighted in blended analysis. Clinical high-risk patients (ER+, HER2−).	• MammaPrint • Oncotype DX • Prosigna • EPclin • Usual care	TransATAC,[20] MINDACT,[30] EBCTCG[16]	Predictive benefit included for MammaPrint. All other tests assumed to be prognostic only.	Assumes mAOL high risk is equivalent to NPI > 3.4. 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. 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. 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. 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 • Oncotype DX • Usual care.	TAILORx[127] MINDACT[30]	Predictive benefit included for MammaPrint and Oncotype DX.	Clinical low-risk patients included, but only clinical high-risk patients receive the genomic test. 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[30] and TransATAC[20]			This is the only analysis which directly addresses the decision problem set out in the NICE scope.[11]

LN, lymph node.

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[20] is equivalent to mAOL high risk in MINDACT.[30]
- 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 22*, together with brief comments from the EAG.

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[85] 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 DSA 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[85] presents the results of a large number of analyses. For brevity, these are summarised in *Table 23*. Across the base-case and scenario analyses, the Agendia model suggests that MammaPrint dominates all other tumour profiling tests and usual care.

External Assessment Group 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.

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*

TABLE 22 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[129]	Reflects patients with LN0 disease.
Test risk classification probabilities	MINDACT,[30] TransATAC[20]	Analyses between MammaPrint and other tests assume equivalence of populations enrolled in TransATAC[20] and MINDACT.[30] MammaPrint data are based on the HR+/HER2− population of MINDACT (LN0 and/or LN+)[30]
DRFI probabilities	MINDACT,[30] TransATAC[20]	DRFI for MammaPrint low-risk ET and MammaPrint high-risk chemotherapy plus ET groups are based on reanalyses of MINDACT IPD.[85]
Risk tapering	Ward *et al.*[105]	Same as DG34 model.
Chemotherapy probability under usual care	NCRAS[17] and expert opinion[85]	The Cytel CEA report[85] 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.*[130] Oncotype DX – Crolley *et al.*[131] Prosigna – UKBCG survey (3-level)[10] 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. Crolley *et al.*,[131] was undertaken in a purely LN0 population and Kuijer *et al.*[130] 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[30]	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 are applied in Scenario 5 (LN+ subgroup).
Chemotherapy benefit for other tests	EBCTCG meta-analysis[16]	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.*[132]	This data set 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.*[108]	Same as DG34 model.
Death risk with AML	Edlin *et al.*[133]	Same as DG34 model.
LR transition probabilities	Geurts *et al.*[134]	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.*[111]	Disutility value applied for patients receiving CET for 3 years. The source of the assumed the duration of the disutility is unclear.
Utility value, DM	Lidgren *et al.*[111]	Same as DG34 model.
Utility value, AML	Younis *et al.*[135]	Same as DG34 model.
Disutility, LR	Campbell *et al.*[112]	Same as DG34 model.
Utility age adjustment	Ara and Brazier[104]	Values reported by Hernández-Alava *et al.*[113] 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.*[95]	Updated from DG34 model.
ET usage	Ward *et al.*[105]	Same as DG34 model.

TABLE 22 Key evidence sources used to inform Agendia model – base-case analysis and Scenario 5 (pure LN+ subgroup) (*continued*)

Parameter group	Source	EAG comments
Bisphosphonates	Ward *et al.*[105]	Similar to DG34 model.
Cost DM	Thomas *et al.*[136]	Same as DG34 model.
Cost AML	Russell-Smith *et al.*[137]	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[138–142]	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,[143] eMIT,[144] MIMS,[145] PSSRU[121]	Appropriate sources applied.
Cost of death	Georgiou and Bardsley[146]	–

CDK4/6i, cyclin-dependent kinase 4/6 inhibitor; DM, distant metastases; eMIT, electronic Market Information Tool; IPD, individual patient data; LN, lymph node; MIMS, *Monthly Index of Medical Specialities*; NCRAS, National Cancer Registration and Analysis Service; NHSE, National Health Service England; pN0, lymph node negative (pathological); pN1(mi), lymph node negative with micrometastases (pathological); PSSRU, Personal Social Services Research Unit; UKBCG, UK Breast Cancer Group.

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.[30] 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.

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[20] and MINDACT[30] 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.

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[85] 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 Scenarios 1–4, albeit with higher proportions of women (54% to 68%) not receiving the tumour profiling test.

Questionable assumptions around post-test chemotherapy probabilities

The post-test probabilities of receiving chemotherapy in the Agendia model are summarised in *Table 24*. 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

TABLE 23 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 = £3647 per QALY gained	Not presented	Not presented	Not presented	Not presented	Not presented

LN, lymph node; .

TABLE 24 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.*[130]
EPclin	0.15	0.92	UKBCG survey[10]
3-level tests			
Oncotype DX	0.03	0.91	Crolley *et al.*[131]
Prosigna	0.04	0.92	UKBCG survey[10]

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.*[130]), 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.*[131] was undertaken in a LN0 population and Kuijer *et al.*[130] 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]

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 21*). 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).[126] 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,[16] 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,[16] 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 × 1.00 + 0.39 × 0.38 = 0.76).

The EAG notes that the updated MINDACT publication by Piccart *et al.*[30] 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 × 0.74 + 0.39 × 0.79 = 0.76). This is similar to the HR for the MammaPrint low-risk group.

Piccart *et al.*[30] also report 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 11*). 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.

Concerns regarding health-related quality-of-life assumptions

The Cytel CEA report[85] (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.*,[111] 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 25*. The model suggests that a patient who is treated with CET who survives for 3 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 *Review of existing economic analyses*, *Table 26* and Summary of economic analysis submitted by Exact Sciences). The only studies in which a similar or higher chemotherapy-related QALY loss is applied are those reported by Vanderlaan *et al.*[102] and Wong *et al.*[101] In both 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;[104] the EAG believes that it would be more appropriate to use more recent estimates by Hernández-Alava *et al.*[113]

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

TABLE 25 Utility values associated with CET applied in the Agendia model

Model health state	Utility value	Source
Recurrence-free, CET, year 1	0.620	Lidgren *et al.*[111] State *p 'First year after primary breast cancer'*, receiving adjuvant chemotherapy
Recurrence-free, CET, year 2–3	0.743	Value used in model NR in Lidgren *et al.*[111]
Recurrence-free, CET, year 4+	0.824	Lidgren *et al.*[111] State S *'Second and following years after primary breast cancer/ recurrence'*, receiving ET
Recurrence-free, ET alone, year 1	0.744	Lidgren *et al.*[111] State *p 'First year after primary breast cancer'*, receiving ET
Recurrence-free, ET alone, year 2+	0.824	Lidgren *et al.*[111] State S *'Second and following years after primary breast cancer/ recurrence'*, receiving ET

TABLE 26 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[85]	QALY loss of 0.29	Lidgren *et al.*[111]
Harnan *et al.* (2019)[10]	QALY loss of 0.04	Campbell *et al.*[112]
Berdunov *et al.* (2021)[92]	Disutility of −0.04, 1-year duration	Campbell *et al.*[112]
Blohmer *et al.* (2013)[98]	QALY loss of 0.07	Peasgood *et al.*[147]
Hall *et al.* (2012)[100]	QALY loss of 0.08	Lidgren *et al.*[111]
Hall *et al.* (2017)[95]	Disutility of −0.096, 1-year duration	Campbell *et al.*[112]
Hannouf *et al.* (2014)[97]	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)[93]	QALY loss of 0.12	Lidgren *et al.*[111]
Lamond *et al.* (2012)[99]	QALY loss appears to be around 0.06 (excluding the impact of CINV and FN)	Tufts,[148] Tengs,[149] Ward[105]
Masucci *et al.* (2019)[94]	QALY loss of 0.06	Lidgren *et al.*[111]
Stein *et al.* (2016)[96]	Disutility of −0.096, 1-year duration	Lidgren *et al.*[111]
Vanderlaan *et al.* (2011)[102]	QALY loss of 0.50 over lifetime	Assumption
Wong *et al.* (2012)[101]	Disutility of −0.30, duration unclear	Assumption

CINV, chemotherapy-induced nausea and vomiting; FN, febrile neutropoenia.

AML compared with the DG34 model (Harnan *et al.*[10] cost = £10,600 vs. Russell-Smith *et al.*[137] cost = £132,039). This cost estimate has been taken from an economic modelling study of gemtuzumab ozogamicin plus 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 *Independent External Assessment Group economic analysis*, *Table 37*).

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.[126]
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.

Additional exploratory analysis undertaken by the External Assessment Group

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.[16]
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 27* 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 *Independent External Assessment Group economic analysis*). As such, the EAG believes that the results of this reanalysis should be interpreted with caution.

Independent External Assessment Group economic analysis

Scope of the External Assessment Group 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 model was programmed using Microsoft Excel. The scope of the EAG's model is summarised in *Table 28*. The model assesses the health outcomes and costs associated with each

TABLE 27 Results of additional analysis undertaken by the EAG

Option	LYGs[a]	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	−£4676	MammaPrint dominating
Usual care	22.02	10.87	£28,003	–	–	–	
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 dominating
Usual care	23.93	11.59	£21,596	–	–	–	

inc., incremental; LN+, lymph node positive.
a Undiscounted.

TABLE 28 Scope of the EAG economic analysis

Population	Women with ER+/PR+, HER2−, LN+ early breast cancer (1–3 nodes). 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. For the evaluation of MammaPrint, the modelled population reflects the mAOL clinical high-risk ER+, HER2−, LN+ subgroup within the MINDACT trial.[30]
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].0 2. Prosigna (cut-offs LN+: low 0–15, intermediate 16–40, high 41–100). 3. EPclin (cut-off: 3.3: low < 3.3; high ≥ 3.3). 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. For MammaPrint, the clinical high-risk subgroup is based on mAOL, as per the design of the MINDACT trial.[30,150] Due to evidence limitations, the tumour profiling tests are not compared incrementally against each other.[a]
Main economic outcome	Incremental cost per QALY gained
Additional model outcomes	• Incremental LYGs • Incremental QALYs gained • Incremental costs • Impact on chemotherapy use
Perspective	NHS and PSS
Time horizon	Lifetime
Discount rate	3.5% per annum
Price year	2022–3

LN, lymph node; .
a 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 data sets which feature different sample sizes between the tests.

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 2022–3, including uplifting of older cost estimates using inflation indices[151] where necessary.

Population

Overall, the population reflected in the economic model relates to women with ER+/PR+, HER2−, LN+ (one to three 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,[85] the analysis of MammaPrint is focused 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,[20] RxPONDER[29,78] and MINDACT.[30,150] 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 3*).

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 28*. 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: (1) the newer cut-offs of RS 0–25 and RS > 25, as assessed in RxPONDER,[29] and (2) 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,[30,150] 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,[20] 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 External Assessment Group

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: (1) the analyses presented by the EAG to inform DG34;[10] (2) more recent evidence on the tests published since DG34; and (3) key scenarios presented in the Cytel CEA report[85] and the Exact Sciences CS.[23] 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,[78] supplemented using external data on women with an RS of > 25 (thereby assuming predictive benefit).[20,32] This scenario is similar to the pre-menopausal LN+ subgroup analysis presented in the Exact Sciences CS.[23]
- **BC2** – Oncotype DX versus current decision-making, RxPONDER post-menopausal LN+ subgroup,[78] supplemented using external data on women with an RS of > 25 (thereby assuming predictive benefit).[20,32] This scenario is similar to the post-menopausal LN+ subgroup analysis presented in the Exact Sciences CS.[23]
- **BC3** – Oncotype DX versus current decision-making, TransATAC post-menopausal LN+ population,[20] assuming predictive benefit based on SWOG-8814.[32] 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,[20] assuming prognostic benefit only.[16] 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,[20] assuming prognostic benefit only.[16] This scenario is similar to the EAG non-predictive base-case scenario for Prosigna in the LN+ population in Harnan *et al.*[10]
- **BC6** – EPclin versus current decision-making, TransATAC post-menopausal LN+ population,[20] assuming prognostic benefit only.[16] This scenario is similar to the EAG non-predictive base-case scenario for EPclin 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,[85] but excludes the company's assumption of a predictive benefit for MammaPrint [see *Agendia model summary and critique (MammaPrint)*]. 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 *Evidence sources used to inform the model parameters*.

The EAG notes that Exact Sciences have indicated that the validated RS result cut-offs used in RxPONDER[78] 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.[103] 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.

Model structure

The EAG's model is based on the economic analysis used to inform DG34,[13] together with updated evidence 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 *Review of existing economic analyses*) as well as the models submitted by Agendia[85] and Exact Sciences[23] (see *Review and critique of economic analyses of tumour profiling tests submitted by the test manufacturers*). 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 risk 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 (*Figure 5*). 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.

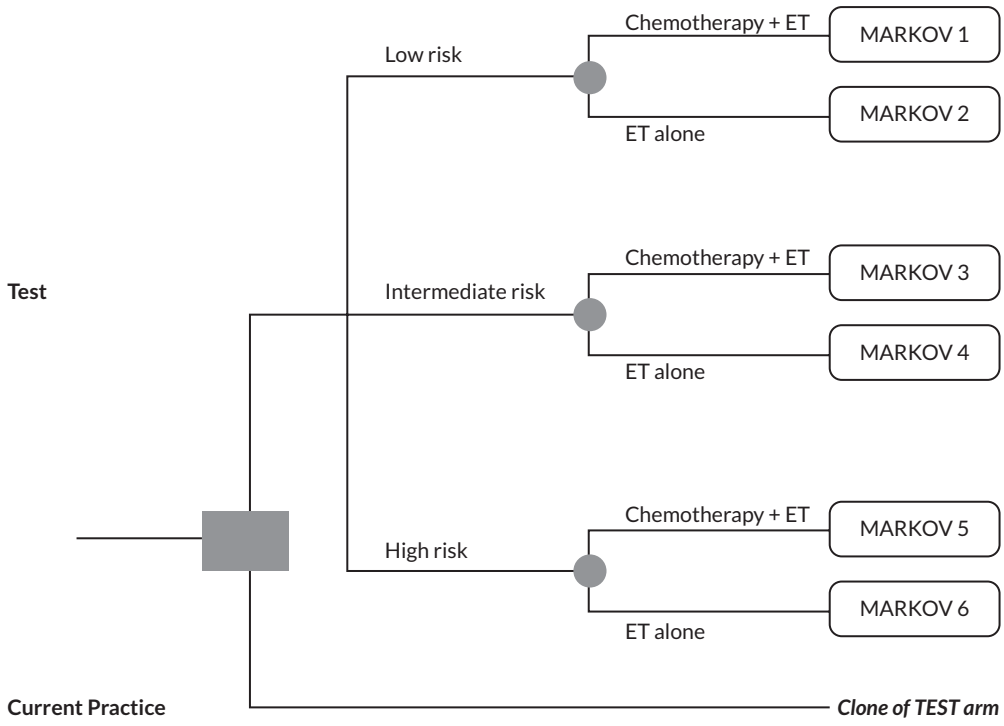


FIGURE 5 External Assessment Group's model – decision tree component. Notes: For Oncotype DX under the new cut-offs, EPclin, MammaPrint, four branches are used due to the absence of an intermediate-risk category for these tests. All patients are also assumed to receive ET.

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.

The structure of the long-term Markov submodels is illustrated in *Figure 6*. Each Markov submodel is evaluated using 6-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 submodel includes four mutually exclusive and jointly exhaustive health states: (1) recurrence-free; (2) DM; (3) AML; and (4) dead. Each submodel 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 submodels, 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. Congestive heart failure (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.[152] 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.

QALYs gained are estimated by assigning health utility values to each of the Markov submodel 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.

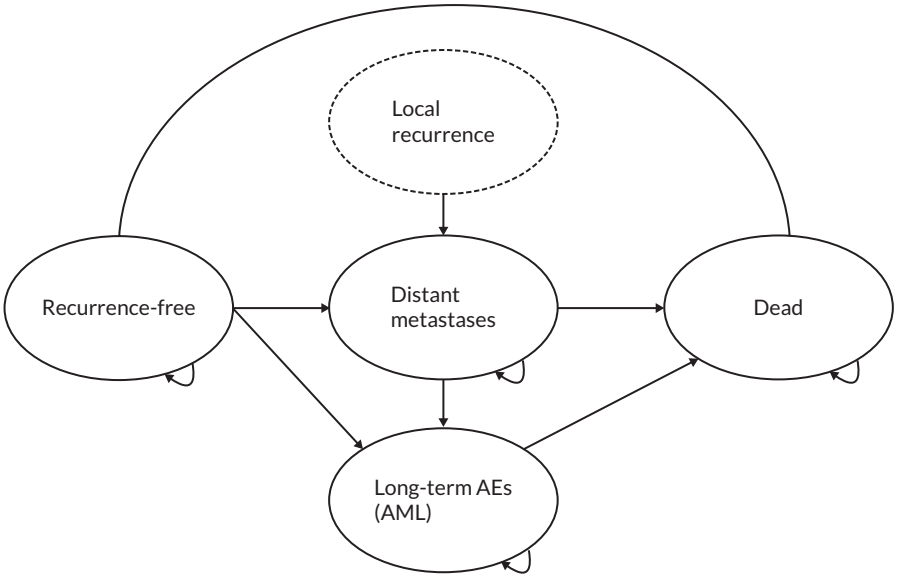


FIGURE 6 External Assessment Group's model – long-term Markov model component. Note: In line with the model developed to inform DG34, once-only costs and QALY losses associated with LR are modelled for women who develop DM (see *Evidence sources used to inform the model parameters*)

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 External Assessment Group 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 risk 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 for at least 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.

Evidence sources used to inform the model parameters

Table 29 summarises the key evidence sources used to inform the model. Further details on the individual model parameters are provided in the subsequent sections.

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.*[18] The analysis of MammaPrint (BC7) applies a mean age of 56 years, assuming that one-third of women are pre-menopausal.[150] Patients are assumed to have a mean body surface area 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 30*. 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 on RxPONDER.[29] 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* = 9112; see *Figure 1* in Kalinsky *et al.*[29]).
- Within BC3–6, risk classification probabilities for Oncotype DX, EPclin and Prosigna were taken from the published analysis of TransATAC.[20]
- 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.*,[30] Supplementary Appendix, page 29, Table S10).

Distant recurrence-free interval on endocrine therapy alone

Distant recurrence-free interval estimates for women receiving ET alone were taken from RxPONDER, TransATAC and MINDACT (*Table 31*).[20,78,85] 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.*[78] This source was used because it reports DRFI estimates by menopausal status.
- The time points for reporting DRFI differ between the three trials, with RxPONDER,[78] MINDACT[30] and TransATAC[20] 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.[85] 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.*[30] (Supplementary Appendix, p. 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[16] (1/0.71) to the DRFI estimate for the clinical high-risk genomic high-risk group. This approach assumes no predictive benefit for MammaPrint.

Hormone receptor 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 (*Table 32*).

TABLE 29 Evidence sources used in the EAG's base-case model

Parameter group	Source
Patient characteristics	
Patient age	Holt *et al.*[18]
Mean BSA	Assumption
Clinical parameters	
Risk classification probabilities	TransATAC,[20] RxPONDER,[29] MINDACT[30]
6-month DRFI on ET alone	TransATAC,[20] RxPONDER,[78] MINDACT IPD[85]
Pre-test probability of receiving adjuvant chemotherapy	Holt *et al.*[18,36]
Post-test probability of receiving chemotherapy (3-level tests)	Holt *et al.*[18,36]
Post-test probability of receiving chemotherapy (2-level tests)	Holt *et al.*[18,36]
HRs for DM, CET vs. ET	EBCTCG,[16] RxPONDER,[78] SWOG-8814[32]
6-month probability of death due to DM	Rebuilt model based on Suri *et al.*[153]
Probability of developing LR	De Bock *et al.*[110]
6-month probability of developing AML	Wolff *et al.*[108]
6-month probability of death due to AML	Rebuilt model based on Bewersdorf *et al.*[154]
Other-cause mortality	ONS life tables for England[155]
HRQoL parameters	
Utility, recurrence-free	Lidgren *et al.*[111]
Utility, distant recurrence	Lidgren *et al.*[111]
Utility AML	Rebuilt model based on Bewersdorf *et al.*[154]
QALY loss due to chemotherapy-related AEs	Campbell *et al.*[112]
QALY loss due to LR	Campbell *et al.*[112]
Cost parameters	
Tumour profiling test costs	Test manufacturers[23–26]
Adjuvant chemotherapy and supportive medications	Proportions based on expert opinion.[92] Unit costs taken from eMIT[156] and BNF.[157]
ET	Distribution and duration of treatments based on expert opinion. Drug costs taken from eMIT[156] and BNF
Bisphosphonates	Proportion based on expert opinion. Unit costs taken from eMIT[156] and BNF.[157]
AEs	Frequency based on Ellis *et al.*[116] Costs taken from NHS Reference Costs 2021–2.[143]
Routine follow-up	Frequency based on expert opinion. Unit costs taken from Ward *et al.*[105] and NHS Reference Costs 2021–2.[143]
LR (once-only cost)	Karnon *et al.*[124]
DM (lifetime cost)	Mean cost reported by Suri *et al.*[153]
AML (lifetime cost)	Costs of intensive therapy, HSCT and subsequent BSC from Zeidan *et al.*[115] applied to rebuilt model based on Bewersdorf *et al.*[154]

BNF, *British National Formulary*; BSA, body surface area; BSC, best supportive care; DM, distant metastases; eMIT, electronic Market Information Tool; HSCT, haematopoietic stem cell transplantation; ONS, Office for National Statistics.

TABLE 30 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[a]	Kalinsky *et al.*[29]
BC2	RxPONDER post-menopausal	Oncotype DX	0.89	N/A	0.11[a]	Kalinsky *et al.*[29]
BC3	TransATAC, predictive	Oncotype DX	0.57	0.32	0.11	Sestak *et al.*[20]
BC4	TransATAC, non-predictive	Oncotype DX	0.57	0.32	0.11	Sestak *et al.*[20]
BC5	TransATAC, non-predictive	Prosigna	0.08	0.32	0.60	Sestak *et al.*[20]
BC6	TransATAC, non-predictive	EPclin	0.23	N/A	0.77	Sestak *et al.*[20]
BC7	MINDACT clinical high-risk	MammaPrint	0.69	N/A	0.31	Piccart *et al.*[30]

BC, base case; Int, intermediate; N/A, not applicable.
a 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.

TABLE 31 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.*[78] RS > 25: Sestak *et al.*[20]
BC2	RxPONDER post-menopausal	Oncotype DX	RS0-25 : 5 years	0.03	N/A	0.38	RS0-25: Kalinsky *et al.*[78]
			RS > 25 : 10 years				RS > 25: Sestak *et al.*[20]
BC3	TransATAC, predictive	Oncotype DX	10 years	0.19	0.29	0.38	Sestak *et al.*[20]
BC4	TransATAC, non-predictive	Oncotype DX	10 years	0.19	0.29	0.38	Sestak *et al.*[20]
BC5	TransATAC, non-predictive	Prosigna	10 years	0.00	0.21	0.31	Sestak *et al.*[20]
BC6	TransATAC, non-predictive	EPclin	10 years	0.06	N/A	0.30	Sestak *et al.*[20]
BC7	MINDACT clinical high-risk	MammaPrint	8 years	0.09	N/A	0.26[a]	MINDACT IPD[85]

BC, base case; int, intermediate; N/A, not applicable.
a 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.

- 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.*[78] 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.*[32]). 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.[32]

TABLE 32 Hormone receptors for DM for chemotherapy vs. 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.*[78] RS > 25: Albain *et al.*[3]
BC2	RxPONDER post-menopausal	Oncotype DX	1.12	N/A	0.59	RS 0–25: Kalinsky *et al.*[78] RS > 25: Albain *et al.*[32]
BC3	TransATAC, predictive	Oncotype DX	1.02	0.72	0.59	Albain *et al.*[32]
BC4	TransATAC, non-predictive	Oncotype DX	0.71	0.71	0.71	EBCTCG[16]
BC5	TransATAC, non-predictive	Prosigna	0.71	0.71	0.71	EBCTCG[16]
BC6	TransATAC, non-predictive	EPclin	0.71	N/A	0.71	EBCTCG[16]
BC7	MINDACT clinical high-risk	MammaPrint	0.71	N/A	0.71	EBCTCG[16]

BC, base case; Int, intermediate.

Note
Values shown in the table are median estimates. These are converted to mean values within the model.

TABLE 33 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.*[18]

BC, base case.

- For the analyses of all tests without predictive benefit (BC4–7), the model applies an HR for DRFI based on the EBCTCG meta-analysis.[16] 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,[16] Web Extra Material, Analysis P11, p. 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.*

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.*[18] (*Table 33*). 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.

Post-test probability of receiving chemotherapy

The post-test chemotherapy probabilities applied in the model are shown in *Table 34*. Post-test chemotherapy probabilities were selected based on consideration of the studies included in the systematic review of decision impact studies (see *Results: decision impact*):

- 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.*[18] 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

TABLE 34 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.*,[18] pre-menopausal subgroup
BC2	RxPONDER post-menopausal	Oncotype DX	0.11	N/A	0.96	Holt *et al.*,[18] post-menopausal subgroup
BC3	TransATAC, predictive	Oncotype DX	0.08	0.49	0.98	Holt *et al.*[36]
BC4	TransATAC, non-predictive	Oncotype DX	0.08	0.49	0.98	Holt *et al.*[36]
BC5	TransATAC, non-predictive	Prosigna	0.08	0.49	0.98	Holt *et al.*[36]
BC6	TransATAC, non-predictive	EPclin	0.11	N/A	0.96	Holt *et al.*,[18] post-menopausal subgroup
BC7	MINDACT clinical high-risk	MammaPrint	0.19	N/A	0.96	Holt *et al.*,[18] post-menopausal subgroup

BC, base case; Int, intermediate risk; N/A, not applicable.

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.*[36] 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 reanalysis of the Holt *et al.* data[36] 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.*[18] 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 *Results: decision impact*). These alternative sources include Llombart Cussac *et al.*,[45] Loncaster *et al.*,[37] Zambelli *et al.*,[46] Dieci *et al.*[43] and the UKBCG survey reported in Harnan *et al.*[10]

Long-term risk of distant metastases on endocrine therapy 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 time points. However, a meta-analysis of 88 trials involving 62,923 women with ER+ breast cancer reported by Pan *et al.*[158] suggests that the risk of DM in women with breast cancer with one to three 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 distant metastases

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.*[136]). 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 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.[153] 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 of ribociclib plus letrozole versus placebo plus letrozole based on the MONALEESA-2 and PALOMA-1 studies.[159,160] 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 local recurrence conditional on distant metastases

The probability of LR was based on a multistate modelling study reported by de Bock *et al.*[110] Within this study, the authors analysed 3601 women enrolled in three European Organisation for Research and Treatment of Cancer RCTs. The study included both LN0 and LN+ women who had been treated for early-stage breast cancer. Of the 1224 women who developed DM, 129 women experienced a previous LRR. 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.[23,85]

Six-month probability of developing acute myeloid leukaemia

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.*[108]). 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.[85]

Six-month probability of death due to acute myeloid leukaemia

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, chronic myelomonocytic leukaemia and AML (TA218).[133] 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.*[154]). The published model presents plots of cumulative survival probabilities based on log-logistic models fitted to data on event-free survival (EFS) and OS from Study 301[125] 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 of 2.3 based on Martin *et al.*[161] 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 life tables for England (years 2018–20).[155] 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 35*. 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 EuroQol-5 Dimensions, three-level version (EQ-5D-3L) estimates for the general population of the UK reported by Hernández-Alava *et al.*[113]

TABLE 35 Utility values and QALY losses applied in the EAG's model

Parameter	Mean value	Source
Utility, recurrence-free	0.824	Lidgren *et al.*[111]
Utility, DM	0.685	Lidgren *et al.*[111]
Utility AML	0.59	Estimated by dividing the mean QALYs by mean the LYGs in the rebuilt model based on Bewersdorf *et al.*[154]
QALY loss chemotherapy	−0.038	Campbell *et al.*[112]
QALY loss LR	−0.108	Campbell *et al.*[112]

DM, distant metastases.

Utility values associated with recurrence-free and distant metastases states and quality-adjusted life-year loss associated with chemotherapy-related toxicity

The model developed to inform DG34[85] applied utility values to the recurrence-free and DM health states based a cross-sectional observational study of 361 patients with a 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.*[111]). Within this study, patients were asked to complete both the EQ-5D-3L and a direct time trade-off 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.*[112]

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–present.
- EMBASE: Ovid, 1974–7 July 2017.
- SCI-E: Web of Science, 1900–present.
- CPCI – Science (CPCI): Web of Science, 1990–present.

The searches focused 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.*[162]). 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.*[162] 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.*[111] 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.*[112] This disutility value is also used in the Exact Sciences model[23] and is the same as the value used in the EAG's model used to inform NICE DG34.[10]

Utility value associated with acute myeloid leukaemia

The utility value for the AML state was estimated based on the same model used to estimate survival with AML (Bewersdorf *et al.*[154]). 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).

Quality-adjusted life-year loss due to local recurrence

The model applies a QALY loss of −0.108 for patients experiencing LR. This estimate was also taken from Campbell *et al.*[112] This value is also used in the Exact Sciences model[23] and the Agendia model[85] 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 36* 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.

Tumour profiling tests

The list prices of the tests applied in the EAG's model are summarised in *Table 37*. 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.

Adjuvant chemotherapy and supportive medications

The costs associated with adjuvant chemotherapy and supportive medications are summarised in *Table 38*. The proportionate use of each chemotherapy regimen was taken from Berdunov *et al.*,[92] which in turn, was based on the costing approach used by Hall *et al.*[95] 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 comorbidities 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. Granulocyte-colony stimulating factor (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 those associated with pharmacy preparation, outpatient monitoring visits and tests (full blood counts, liver function tests and urea and electrolytes) and electrocardiograms (ECGs) in 25% of patients]. Drug acquisition costs were taken from eMIT.[156] The costs of delivering chemotherapy and tests were taken from NHS Reference Costs 2021–2.[143] The cost of pharmacy preparation was taken from Ward *et al.*[105]

TABLE 36 Summary of costs applied in the EAG's model

Resource use component	Mean cost
Tumour profiling tests (list prices)	
Oncotype DX	£2580
Prosigna	£1896
EPclin	£1500
MammaPrint	£2616
Adjuvant chemotherapy (once-only)	£7410.48
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)[a]	£320.84
Ovarian suppression (per cycle)[b]	£496.73
AEs (once-only)	£1249.58
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)	£4898.17

DM, distant metastases.
a Applied to post-menopausal women only.
b Applied to pre-menopausal women only.

TABLE 37 Costs of tumour profiling tests

Test	List price excluding VAT	Source
Oncotype DX	£2580.00	Exact Sciences CS,[23] 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	£1896.00	Veracyte RFI document,[26] 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	£1500.00	Myriad RFI document,[25] March 2023. Price includes all reaction agents and consumables. Price reflects locally run testing service.
MammaPrint	£2616.00	Agendia value dossier,[24] February 2023. Price includes transport, specimen processing and all other costs associated with reporting the result.

CS, company's submission; VAT, value added tax.

TABLE 38 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	£4110.70	£4931.83
FEC100-T (3 + 3 cycles)	23.75	£1581.22	£4110.70	£5691.92
TC (4 cycles)	10.00	£1640.50	£2735.53	£4376.03
EC90/T75 (4 + 4 cycles)	28.75	£2101.75	£5485.87	£7587.62
EC90 (4 cycles)	0.00	£547.90	£2735.53	£3283.43
C-D (6 cycles)	2.50	£2458.68	£4110.70	£6569.38
TAC (6 cycles)	1.25	£2509.74	£4110.70	£6620.44
Accelerated EC90/P (4 + 4 cycles)	23.75	£3381.97	£5485.87	£8867.84
Weekly P (12 weeks)	2.50	£155.13	£8236.21	£8391.34
EC/weekly P (4 cycles, 12 weeks)	7.50	£703.03	£10,986.55	£11,689.58
Weighted cost	–	£2096.50	£5313.98	£7410.48

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

Within the economic model, a weighted mean cost of £7410.48 is applied to all patients who receive chemotherapy. This cost is applied in the first model cycle as a once-only cost.

Adverse events associated with adjuvant chemotherapy

The frequency of grade 3/4 AEs was informed by the TACT trial[116] (*Table 39*). Unit costs were taken from NHS Reference Costs 2021–2[143] based on the same service codes as those used in the Exact Sciences model.[23] 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 38*. The model applies a weighted cost of £1249.58 to all patients receiving adjuvant chemotherapy. This cost is applied as a once-only cost in the first 6-month model cycle.

Endocrine therapy

The model assumes that, while 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,[156] whereas the price of tamoxifen was taken from the British National Formulary (BNF).[157] Monthly pharmacy preparation and dispensing costs were taken from the Personal Social Services Research Unit (PSSRU).[151] 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.

TABLE 39 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	£1439.66	NHS Reference Costs 2021–2:[143] SA04G-L, non-elective short and long stay
Febrile neutropoenia	0.071	0.029	£3676.55	NHS Reference Costs 2021–2:[143] SA35A-E, non-elective long stay
Leucopoenia	0.246	0.175	£501.80	NHS Reference Costs 2021–2:[143] weighted average of JA12D-L, non-elective short and long stay and NHS Reference Costs 2020–1: consultant-led outpatient visit, WF01A medical oncology
Neutropoenia	0.455	0.384	£501.80	NHS Reference Costs 2021–2:[143] weighted average of JA12D-L, non-elective short and long stay and NHS Reference Costs 2020–1: consultant-led outpatient visit, WF01A medical oncology
Thrombocytopenia	0.006	0.013	£2163.16	NHS Reference Costs 2021–2:[143] SA12G-K, non-elective short and long stay
Alopecia	0.102	0.103	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Diarrhoea	0.037	0.028	£1446.84	NHS Reference Costs 2021–2:[143] FD10J-M, non-elective short and long stay
Infection	0.142	0.088	£1628.07	NHS Reference Costs 2021–2:[143] DZ22K-Q, non-elective short and long stay
Lethargy	0.221	0.131	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Musculoskeletal (other)	0.070	0.015	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Myalgia/arthralgia	0.050	0.001	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Nausea/vomiting	0.097	0.099	£1579.14	NHS Reference Costs 2021–2:[143] FD11K, non-elective short and long stay
Neuropathy	0.048	0.005	£1886.35	NHS Reference Costs 2021–2:[143] AA26C-H, non-elective short and long stay
Oedema	0.008	0.003	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Pain	0.028	0.001	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Skin disorder (including nail changes)	0.033	0.012	£221.48	NHS Reference Costs 2021–2:[143] consultant-led outpatient visit, WF01A medical oncology
Stomatitis	0.076	0.036	£1978.14	NHS Reference Costs 2021–2:[143] CB01F, non-elective short and long stay

FEC-D, fluorouracil, epirubicin and cyclophosphamide followed by docetaxel.

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 2021–2.[143] The unit cost for mammograms is not listed in NHS Reference Costs 2021–2; instead, this was taken from Ward *et al.*[105] These costs are applied in each 6-monthly cycle for up to 5 years while patients remain recurrence-free.

Cost of bisphosphonates

The model assumes that 60% of post-menopausal women who are recurrence-free receive bisphosphonates (4 mg zoledronic acid) every 6 months for 3 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[156] and the cost of administration was taken from NHS Reference Costs 2021–2.[143] 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.[157] Administration costs were taken from NHS Reference Costs 2021–2[143] and the PSSRU.[151] 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.[23,85]

Cost of treating local recurrence

The cost of treating LR was taken from a breast cancer costing study reported by Karnon *et al.*[124] (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.*,[110]).

Lifetime cost of treating distant metastases

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.*[153] The EAG's model applies a once-only cost of £117,482 to patients entering the DM health state.

Lifetime cost of treating acute myeloid leukaemia

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.*[154]), together with treatment costs reported by Zeidan *et al.*[115] 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[125] (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 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.*,[93] which, in turn, was based on Karnon *et al.*[124] Within the model, a once-only cost of £4898 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 External Assessment Group critique of the Exact Sciences model, 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.*[56] rather than TransATAC.[20] 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.*[163] rather than TransATAC.[20] The test risk classification probabilities for low risk 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.*[45] rather than Holt *et al.*[18] (values presented in *Table 15*).
- *DSA5:* The post-test chemotherapy probabilities for 3-level tests were based on estimates reported by Loncaster *et al.*[37] rather than Holt *et al.*[18] (values presented in *Table 15*).
- *DSA6:* The post-test chemotherapy probabilities for 3-level tests were based on estimates reported by Zambelli *et al.*[90] rather than Holt *et al.*[18] (values presented in *Table 15*).
- *DSA7:* The post-test chemotherapy probabilities for 2-level tests were based on estimates reported by Dieci *et al.*,[43] rather than Holt *et al.*[18] (values presented in *Table 15*).
- *DSA8*: The post-test chemotherapy probabilities for 3-level tests were based on the UKBCG survey reported by Harnan *et al.*,[10] rather than Holt *et al.*[18]
- *DSA9*: The post-test chemotherapy probabilities for 2-level tests were based on the UKBCG survey reported by Harnan *et al.*,[10] rather than Holt *et al.*[18]
- *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.*[162] (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 £7410 to £3705).
- *DSA23:* The cost of adjuvant chemotherapy was doubled (from £7410 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*, *Table 55*).

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[164] and published checklists for economic evaluations and models.[165,166]
- 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.

Results of the External Assessment Group 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 *Tables 40* and *41*, respectively. The results of the DSAs are presented in *Table 42*. A summary of the model-predicted impact of tumour profiling testing on chemotherapy use, clinical outcomes, costs and net health benefits (NHBs) per 1000 women tested is presented in *Table 43*. CEACs for each comparison are shown in *Appendix 9*, *Figures 7–13*. 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 £1810 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 £4273 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 £9772 to £279,599 saved per QALY lost. The analyses of Oncotype DX within the post-menopausal LN+ subgroup based on TransATAC[20] 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 £1084 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

TABLE 40 Central estimates of cost-effectiveness, all EAG base-case comparisons, probabilistic

Option	LYGs[a]	QALYs	Costs	Inc. LYGs[a]	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	£1810	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	-£4273	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	−£1942	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	£1811	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	£1084	£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	£4113
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	–	–	–	–

DM, decision-making; Inc., incremental.
a Undiscounted.

TABLE 41 Central estimates of cost-effectiveness, all base-case comparisons, deterministic

Option	LYGs[a]	QALYs	Costs	Inc. LYGs[a]	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	£1787	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	−£4283	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	−£2300	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	£1862	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	£1108	£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	£5580
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	–	–	–	–

DM, decision-making; Inc., incremental.
a Undiscounted.

TABLE 42 Deterministic sensitivity analysis results for all base-case comparisons – test vs. 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**	**£5580**	**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.*,[56]	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.*[163]	N/A	N/A	N/A	N/A	N/A	Dominated	N/A
DSA4: 3-level post-test chemotherapy probabilities – Llombart Cussac *et al.*[45]	N/A	N/A	Dominating	Dominated	£37,959	N/A	N/A
DSA5: 3-level post-test chemotherapy probabilities – Loncaster *et al.*[37]	N/A	N/A	Dominating	Dominated	Dominated	N/A	N/A
DSA6: 3-level post-test chemotherapy probabilities – Zambelli *et al.*[46]	N/A	N/A	Dominating	Dominated	£62,801	N/A	N/A
DSA7: 2-level post-test chemotherapy probabilities – Dieci *et al.*[43]	Dominated	Dominating	N/A	N/A	N/A	£6448	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	£7097	Dominated
DSA11: CET vs. ET HR = 0.60 in all genomic risk groups (non-predictive)	Dominated	£9772 (SWQ)	Dominated	Dominated	£24,584	Dominating	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
DSA12: CET vs. ET HR = 0.71 in all genomic risk groups (non-predictive)	Dominated	£42,518 (SQW)	Dominated	Dominated	£40,220	£5580	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	£5820	Dominated
DSA15: Chemotherapy QALY loss doubled	Dominated	Dominating	£757,556 (SWQ)	Dominated	£49,618	£6080	Dominated
DSA16: Chemotherapy QALY loss tripled	Dominated	Dominating	£106,021 (SWQ)	Dominated	£53,808	£6267	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	£8137	Dominated
DSA19: Start age – 5 years	Dominated	Dominating	Dominating	Dominated	£33,567	£4379	Dominated
DSA20: Utility values from Verrill *et al.*[162]	Dominated	Dominating	Dominating	Dominated	£44,393	£6172	Dominated
DSA21: AML removed from model	Dominated	Dominating	Dominating	Dominated	£47,629	£7274	Dominated
DSA22: Chemotherapy cost halved	Dominated	Dominating	Dominating	Dominated	£46,376	£8253	Dominated
DSA23: Chemotherapy cost doubled	£5007 (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	£1239 (SWQ)
DSA25: DM lifetime cost doubled	Dominated	Dominating	Dominating	Dominated	£28,111	Dominating	Dominated
DSA26: AML costs halved	Dominated	Dominating	Dominating	Dominated	£41,175	£6066	Dominated
DSA27: AML costs doubled	Dominated	Dominating	Dominating	Dominated	£38,311	£4608	Dominated

BC, base case; DM, distant metastases; N/A, not applicable; SWQ, South-West quadrant.

TABLE 43 Model-predicted incremental clinical and economic outcomes per 1000 women tested – test vs. current decision-making

Incremental model outcome (test vs. 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	−1854	−3051	−2520	−2520	−235	−203	−1900
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.

£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 £23,853 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 £4113 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.

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: (1) pre-menopausal women with LN+ early breast cancer; (2) post-menopausal women with LN+ early breast cancer; and (3) a blended analysis which reflects a mixed pre- and post-menopausal LN+ population. The model is informed by RxPONDER[29,78] in women with an Oncotype DX RS of 0–25 and by external data (TransATAC[20] and SWOG-8814[32]) 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.[18] 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 reanalysis of IPD for women with HR+, HER2− disease (LN0 or LN+) from the MINDACT trial.[30] 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 reanalysis 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 reanalysis of this model which removes the assumption of predictive benefit, down-weights the chemotherapy-related QALY loss and corrects the programming errors. This reanalysis 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.[29,78]
- 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.[30]
- The incorporation of published analyses of TransATAC.[20]
- The incorporation of estimates of pre- and post-test chemotherapy use, based on Holt *et al.*[18] 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). This result is driven by a large reduction in the use of adjuvant chemotherapy in women who would have benefitted from treatment and an increase in the lifetime probability of developing DM.

- *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 £4113 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[164] 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.*[18]).
- 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, and 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 *Results: prediction of chemotherapy benefit*, tests for interaction between Oncotype DX RS and chemotherapy benefit on DFS in SWOG-8814[32] were statistically significant for some analyses, but not others. RxPONDER[29] 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 *Conclusions for prediction of chemotherapy benefit data*). 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 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 *Results: decision impact*) 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.*[18]). 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[20] 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,[20] 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.
- Among 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,[103] 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 (see *Table 43*) 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 43*). 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).

Chapter 4 Discussion and conclusions

Statement of principal findings

Clinical effectiveness – principal findings

Overview of evidence

The search identified 4058 articles. In total, 55 articles were included, 42 relating to prognostic and predictive ability and 13 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[29] for Oncotype DX and MINDACT[30] for MammaPrint. In RxPONDER,[29] patients with an Oncotype DX RS of ≤ 25 were randomised to chemotherapy versus no chemotherapy. In RxPONDER, 65% of patients had one positive node, 25% had two positive nodes, and 9% had three positive nodes. The MINDACT study[30] 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 versus no chemotherapy. The ongoing OPTIMA RCT compares Prosigna test-directed chemotherapy use versus standard chemotherapy use; however, results are not yet available.[34]

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 reanalyses 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[66] 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[29] 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,[30] within LN+ patients at high clinical risk, 8-year DMFI was 92.3% for MammaPrint low-risk versus 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 versus no chemotherapy on patient outcomes differs between test risk groups or ranges, and is 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,[32] Oncotype DX was conducted retrospectively on tumour samples from patients randomised to chemotherapy versus 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[29] reported no benefit of chemotherapy in post-menopausal patients with an RS of 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 an RS of 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[71,73,74,82] 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[33] 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[30] prospective RCT, within the clinical high-risk, MammaPrint low-risk, LN+, HR+ HER2− subgroup, 8-year DMFS was 91.2% with chemotherapy versus 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 (5 in the UK and 7 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,[35–40] and a reduction of 12% to 73% across seven European studies.[41–47] Within studies reporting data by Oncotype DX risk group, there were greater reductions in chemotherapy recommendations in the low-risk and intermediate-risk groups than in the high-risk groups.

Health-related quality of life and anxiety

No studies reported HRQoL or anxiety associated with use of tumour profiling tests in a LN+ population. Therefore, studies in a LN0 or mixed nodal status population were briefly summarised. 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. It is unclear how far the results of these studies can be generalised to a LN+ population.

Evidence on clinical subgroups

The NICE scope for this appraisal[11] 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, while 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[69] reported significant prognostic ability of Oncotype DX in both men and women. In terms of ethnicity, one RxPONDER publication[51] 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[67] 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.

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 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. 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 £4113 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.

Strengths and limitations of the assessment

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 versus 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.

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 the 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.

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.

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.

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 improvements.

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 USA or 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.

Suggested research priorities

Research priorities include the following:

- There remains uncertainty around the ability of all four tests to predict relative chemotherapy benefit in LN+ populations. Further studies demonstrating a statistical interaction between test score and long-term chemotherapy benefit across the full range of test scores would help to address this uncertainty. This may require observational or registry data, despite the limitation that test results may influence chemotherapy use, due to the ethical issues in withholding chemotherapy from test high-risk patients.
- 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 and European studies assessing the impact of tumour profiling tests on recommendations for adjuvant chemotherapy in LN+ populations may help to reduce uncertainty around the clinical impact and cost-effectiveness of these tests.
- The integration of tumour profiling tests with decision aid tools to support shared decision-making may constitute a useful research direction.
- The role of tumour profiling tests in older adults, who may be more prone to chemotherapy complications in the context of limited life expectancy, is also a research priority, as is research on test performance in ethnically diverse populations.

The use of patient and public involvement

There was no patient and public involvement in producing the draft version of this report. However, the report was circulated for consultation to stakeholders, which included patients and the public, and the final EAG report was discussed by the NICE Technology Appraisal Committee at a meeting which included representation by patients and the public. The report was amended and further analyses were conducted to address points raised during these stages of the appraisal.

Equality, diversity and inclusion

As this report is secondary research, no patient participation was involved and the EAG did not need to consider the equality, diversity and inclusion of participants. The primary research team was part of the SCHARR Technology Assessment Group contracted by the Department of Health, and this team is a diverse group representing a wide range of protected characteristics, consisting of seniority, ages, ethnicity and religious beliefs, and including both males and females. The clinical team represents experts within their field who have successfully worked with the SCHARR Technology Assessment Group.

Additional information

CRediT contribution statement

Paul Tappenden (https://orcid.org/0000-0001-6612-2332): Conceptualisation, Formal analysis, Methodology, Project administration, Validation, Supervision, Writing – original draft, Writing – reviewing and editing.

Katy Cooper (https://orcid.org/0000-0002-7702-8103): Conceptualisation, Formal analysis, Methodology, Project administration, Validation, Supervision, Writing – original draft, Writing – reviewing and editing.

Jean Hamilton (https://orcid.org/0000-0003-3326-9842): Formal analysis, Writing – reviewing and editing.

Gamze Nalbant (https://orcid.org/0000-0002-5414-9383): Formal analysis, Writing – original draft, Writing – reviewing and editing.

Munira Essat (https://orcid.org/0000-0003-2397-402X): Formal analysis, Writing – reviewing and editing.

Annabel Rayner (https://orcid.org/0000-0002-6458-1663): Formal analysis, Writing – reviewing and editing.

Ruth Wong (https://orcid.org/0000-0002-4536-4794): Formal analysis, Writing – original draft, Writing – reviewing and editing.

Nicolò Matteo Luca Battisti (https://orcid.org/0000-0002-1063-1717): Resources (expert clinical opinion), Writing – reviewing and editing.

Lynda Wyld (https://orcid.org/0000-0002-4046-5940): Resources (expert clinical opinion), Writing – reviewing and editing.

Uzma Asghar (https://orcid.org/0000-0001-5103-1673): Resources (expert clinical opinion), Writing – reviewing and editing.

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.

Data-sharing statement

All data from the systematic review synthesis are available in appendices to this report. All other data requests should be submitted to the corresponding author for consideration.

Ethics statement

This review did not involve the collection or analysis of any data that were not included in previously published research in the public domain. Therefore, it was exempt from formal ethical review.

Information governance statement

No personal information was collected or stored during this project.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/KGFD4040.

Primary conflicts of interest: Nicolò Matteo Luca Battisti has received consulting fees from Pfizer, Abbott, Sanofi and Astellas, honoraria from Pfizer, AbbVie, Roche, Sanofi, Novartis, Exact Sciences, Servier, Gilead and AstraZeneca and support for attending meetings from Exact Sciences, Pfizer, Lilly, and Novartis.

Lynda Wyld has received honoraria for lectures to ESTRO and ESO and was a trustee of the Association for Breast Surgery and a member of the HTA Surgery Board from 2013 to 2023.

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

References

1. Cancer Research UK. *Breast Cancer Incidence (Invasive) Statistics – Breast Cancer Incidence by Gender and UK Country*. London, UK: Cancer Research UK; 2023. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-Zero (accessed 7 August 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*. London: NICE; 2018.
4. Chua B, Ung O, Taylor R, Boyages J. Frequency and predictors of axillary lymph node metastases in invasive breast cancer. *ANZ J Surg* 2001;**71**:723–8.
5. NHS Digital. *Cancer Survival in England, Cancers Diagnosed 2016 to 2020, Followed up to 2021*. 2023. URL: https://digital.nhs.uk/data-and-information/publications/statistical/cancer-survival-in-england/cancers-diagnosed-2016-to-2020-followed-up-to-2021 (accessed 8 August 2023).
6. Cancer Research UK. *Breast Cancer Incidence (Invasive) Statistics – Breast Cancer Incidence by Age*. 2023. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-One (accessed 7 August 2023).
7. Cancer Research UK. *Breast Cancer Mortality Statistics – Breast Cancer Mortality by Age and UK Country*. 2023. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/mortality#heading-Zero (accessed 8 August 2023).
8. Rosen RD, Sapra A. TNM Classification. *StatPearls*. 2023. URL: www.ncbi.nlm.nih.gov/books/NBK553187/ (accessed 8 August 2023).
9. Amin MB, Gress DM, Meyer Vega LR, Edge SB, Greene FL, Byrd DR, *et al*. *AJCC Cancer Staging Manual*. 8th edn. New York, NY: 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 Technol Assess* 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*. London: NICE; 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 Res* 2017;**19**:58.
13. National Institute for Health and Care Excellence. *Diagnostics Guidance 34. Tumour Profiling Tests to Guide Adjuvant Chemotherapy Decisions in Early Breast Cancer*. London: NICE; 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*. London: NICE; 2022.
15. Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, Nicholson RI, Griffiths K. A prognostic index in primary breast cancer. Br J Cancer. 1982 Mar;45(3):361-6
16. 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.

17. National Cancer Registration and Analysis Service. *Bespoke Data Request – Chemotherapy Use Amongst Women with Early Breast Cancer (Data Held on File)*. London; 2017.
18. 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.
19. Ewertz M, Jensen AB. Late effects of breast cancer treatment and potentials for rehabilitation. *Acta Oncol* 2011;**50**:187–93.
20. 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* 2018;**4**:545–53.
21. Cancer Research UK. *Number Stages of Breast Cancer*. 2023. URL: www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/number-stages (accessed 9 March 2023).
22. 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. *BMJ* 2009;**339**:b2700.
23. Exact Sciences. *Tumour Profiling Tests to Guide Adjuvant Chemotherapy Decisions in Lymph Node Positive Early Breast Cancer. Company evidence submission*. London; 2023.
24. Agendia. *MammaPrint: Reimbursement Dossier United Kingdom*. Amsterdam, Netherlands: Agendia, Inc.; 2023.
25. Myriad. *Additional Information Provided to NICE*. Salt Lake City, Utah: Myriad Genetics; 2023.
26. Veracyte I. *Additional Information Provided to NICE*. San Francisco, CA: Veracyte Inc.; 2023.
27. 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. *BMJ* 2019;**366**:l4898.
28. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, *et al*.; PROBAST Group. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;**170**:51–8.
29. 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. *N Engl J Med* 2021;**385**:2336–47.
30. Piccart M, van 't Veer LJ, Poncet C, Lopes Cardozo JMN, Delaloge S, Pierga JY, *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 Oncol* 2021;**22**:476–88.
31. Ballman KV. Biomarker: predictive or prognostic? *J Clin Oncol* 2015;**33**:3968–71.
32. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, *et al*.; Breast Cancer Intergroup of North America. 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 Oncol* 2010;**11**:55–65.
33. Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A, *et al*.; TRANSBIG Consortium. 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 Res Treat* 2009;**116**:295–302.
34. Stein RC, Makris A, MacPherson IR, Hughes-Davies L, Marshall A, Dotchin G, *et al*.; OPTIMA Trial Management Group. Optima: optimal personalised treatment of early breast cancer using multi-parameter analysis, an international randomized trial of tumor gene expression test directed chemotherapy treatment in a largely node-positive population. *J Clin Oncol* 2021;**39**:TPS599.

35. 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.

36. 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 Res* 2023;**83**.

37. 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. *Eur J Surg Oncol* 2017;**43**:931–7.

38. 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 Rep (Hoboken)* 2022;**5**:e1546.

39. 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. *Eur J Surg Oncol* 2021;**47**:e327.

40. Holt S, Verrill M, Pettit L, Rigg A, Hickish T, Archer C, *et al*. 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. *Br J Cancer* 2024;**130**:1149–56.

41. 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.

42. 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. *Oncologist* 2018;**23**:297–305.

43. Dieci MV, Guarneri V, Zustovich F, Mion M, Morandi P, Bria E, *et al*.; Veneto Oncology Network. 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. *Oncologist* 2019;**24**:1424–31.

44. 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). *Ann Oncol* 2021;**32**:S46.

45. 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**:1529.

46. 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.

47. 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. *Ann Oncol* 2013;**24**:618–24.

48. Sestak I, Dowsett M, Cuzick J. *NICE Request – TransATAC Data Analysis*. London, UK: Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University; 2017.

49. Mamounas EP, Tang G, Paik S, Baehner FL, Liu Q, Jeong JH, *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 Res Treat* 2018;**168**:69–77.

50. 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.

51. 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 Res* 2023;**83**.
52. 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. *J Clin Oncol* 2022;**40**:1335–45.
53. Drukker CA, van Tinteren H, Schmidt MK, Rutgers EJ, Bernards R, van de Vijver MJ, Van't Veer LJ. Long-term impact of the 70-gene signature on breast cancer outcome. *Breast Cancer Res Treat* 2014;**143**:587–92.
54. 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. *Ann Oncol* 2017;**28**:v60.
55. 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. *Ann Oncol* 2022;**33**:S612.
56. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, *et al*.; Austrian Breast and Colorectal Cancer Study Group. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinico-pathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol* 2014;**25**:339–45.
57. Filipits M, Nielsen TO, Rudas M, Greil R, Stöger H, Jakesz R, *et al*.; Austrian Breast and Colorectal Cancer Study Group. The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. *Clin Cancer Res* 2014;**20**:1298–305.
58. Laenkholm AV, Jensen MB, 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. *J Clin Oncol* 2018;**36**:735–40.
59. 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* 2016;**156**:81–9.
60. 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* 2014;**16**:R38.
61. 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 Res Treat* 2020;**179**:197–206.
62. Lundgren C, Bendahl PO, 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.
63. 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* 2019;**25**:3865–72.
64. 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 Re* 2020;**26**:4682–7.
65. 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* 2022;**28**:4435–43.
66. 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.

67. 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.

68. 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 Res Treat* 2017;**163**:303–10.

69. Massarweh SA, Sledge GW, Miller DP, McCullough D, Petkov VI, Shak S. Molecular characterization and mortality from breast cancer in men. *J Clin Oncol* 2018;**36**:1396–404.

70. 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.

71. 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.

72. 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. *Ann Surg Oncol* 2023;**30**:2130–9.

73. 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.

74. Weiser R, Haque W, Polychronopoulou E, Hatch SS, Kuo YF, Gradishar WJ, Klimberg VS. The 21-gene recurrence score in node-positive, hormone receptor-positive, HER2-negative breast cancer: a cautionary tale from an NCDB analysis. *Breast Cancer Res Treat* 2021;**185**:667–76.

75. 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. *J Clin Oncol* 2020;**38**:725–33.

76. 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 Res Treat* 2017;**165**:573–83.

77. 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 (Basel)* 2022;**17**:288–95.

78. 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 Res* 2022;**82**.

79. 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. *Ann Oncol* 2022;**33**:S607–8.

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

81. Abel MK, Shui AM, Chien AJ, Rugo HS, Melisko M, Baehner F, Mukhtar RA. The 21-gene recurrence score in clinically high-risk lobular and ductal breast cancer: a National Cancer Database study. *Ann Surg Oncol* 2022;**29**:7739–47.

82. 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. *J Clin Oncol* 2022;**40**:540.

83. 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 Precis Oncol* 2019;**3**:PO.19.00029.

84. 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 Res* 2023;**83**.

85. 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.

86. 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 Res Treat* 2016;**156**:549–55.

87. 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. *J Clin Oncol* 2010;**28**:1671–6.

88. 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.

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

90. 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. *Curr Med Res Opin* 2015;**31**:1129–37.

91. 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. *Curr Med Res Opin* 2016;**32**:1217–24.

92. 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. *J Med Econ* 2022;**25**:591–604.

93. Hinde S, Theriou C, May S, Matthews L, Arbon A, Fallowfield L, Bloomfield D. The cost-effectiveness of EndoPredict to inform adjuvant chemotherapy decisions in early breast cancer. *Health Policy Technol* 2019;**8**:75–83.

94. 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. *Curr Oncol* 2019;**26**:307–18.

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

96. Stein RC, Dunn JA, Bartlett JM, Campbell AF, Marshall A, Hall P, *et al*.; on behalf of the OPTIMA Trial Management Group. OPTIMA Prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. *Health Technol Assess* 2016;**20**:1–202.

97. 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.

98. 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. *J Med Econ* 2013;**16**:30–40.

99. 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 Res Treat* 2012;**133**:1115–23.

100. Hall PS, McCabe C, Stein RC, Cameron D. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. *J Natl Cancer Inst* 2012;**104**:56–66.

101. 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). *Contemp Clin Trials* 2012;**33**:1117–23.

102. Vanderlaan BF, Broder MS, Chang EY, Oratz R, Bentley TG. Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. *Am J Manag Care* 2011;**17**:455–64.

103. Exact Sciences. *Company's Response to Clarification Questions from the EAG*. London; 2023.

104. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;**13**:509–18.

105. Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P, Wyld L. 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 Technol Assess* 2013;**17**:1–302.

106. 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 Oncol* 2020;**6**:116–24.

107. 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 Res Treat* 2012;**135**:335–46.

108. 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. *J Clin Oncol* 2015;**33**:340–8.

109. National Institute for Health and Care Excellence. *NICE Technology Appraisal No. 552. Liposomal Cytarabine–daunorubicin for Untreated Acute Myeloid Leukaemia. Committee papers*. London: NICE; 2018.

110. 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 Res Treat* 2009;**117**:401–8.

111. Lidgren M, Wilking N, Jönsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Qual Life Res* 2007;**16**:1073–81.

112. 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. *Eur J Cancer* 2011;**47**:2517–30.

113. Hernández Alava M, Pudney S, Wailoo A. *Estimating EQ-5D by Age and Sex for the UK*. Sheffield: University of Sheffield; 2022.

114. 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. *Clin Breast Cancer* 2018;**18**:e529–38.

115. Zeidan AM, Mahmoud D, Kucmin-Bemelmans IT, Alleman CJ, Hensen M, Skikne B, Smith BD. Economic burden associated with acute myeloid leukemia treatment. *Expert Rev Hematol* 2016;**9**:79–89.

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

117. NHS England. *National Schedule of NHS Costs 2018/19*. London: NHS; 2020.

118. NHS England. *National Schedule of NHS Costs 2019/20*. London: NHS; 2021.

119. Department of Health and Social Care. *Drugs and Pharmaceutical Electronic Market Information Tool (eMIT)*. 2021.

120. Joint Formulary Committee. *British National Formulary* (Online). London: BMJ Group and Pharmaceutical Press. URL: https://bnf.nice.org.uk/ (accessed 17 June 2025).

121. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2020*. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2020. https://doi.org/10.22024/UniKent/01.02.84818.

122. 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.

123. Genomic Health Inc. *Oncotype DX Breast Recurrence Score Test – Instructions for Use*. California, USA; 2022.

124. 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. *Br J Cancer* 2007;**97**:479–85.

125. 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. *J Clin Oncol* 2018;**36**:2684–92.

126. 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*. 2023.

127. 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. *N Engl J Med* 2018;**379**:111–21.

128. 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*. London: NICE; 2013.

129. NHS England. *NHS England Oncotype DX Access Scheme Dataset*. London: NHS; 2016.

130. 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. *J Clin Oncol* 2017;**35**:2814–9.

131. 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 Res Treat* 2020;**180**:809–17.

132. 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**.

133. Edlin R, Connock M, Tubeuf S, Round J, Fry-Smith A, Hyde C, Greenheld W. Azacitidine for the treatment of myelodysplastic syndrome, chronic myelomonocytic leukaemia and acute myeloid leukaemia. *Health Technol Assess* 2010;**14**:69–74.

134. Geurts YM, Witteveen A, Bretveld R, Poortmans PM, Sonke GS, Strobbe LJA, Siesling S. Patterns and predictors of first and subsequent recurrence in women with early breast cancer. *Breast Cancer Res Treat* 2017;**165**:709–20.

135. 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 Oncol* 2011;**18**:e288–96.

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

137. 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. *PharmacoEcon Open* 2021;**5**:677–91.

138. 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. *J Clin Oncol* 2006;**24**:5664–71.

139. Martín M, Ruiz Simón A, Ruiz Borrego M, Ribelles N, Rodríguez-Lescure A, 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. *J Clin Oncol* 2015;**33**:3788–95.

140. 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 Res* 2013;**33**:3407–13.

141. 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. *J Clin Oncol* 2009;**27**:1177–83.

142. 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.

143. NHS England. *National Schedule of NHS Costs 2021/22*. London: NHS; 2023.

144. Commercial Medicines Unit. *Commercial Medicines Unit. Drugs and Pharmaceutical Electronic Market Information Tool (eMIT)*. 2021. URL: www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (accessed 17 June 2025).

145. Haymarket Media Group Ltd. *Monthly Index of Medical Specialities*. Gloucester; 2022.

146. Georghiou T, Bardsley M. *Exploring the Cost of Care at the End of Life*. London: Nuffield Trust; 2014.

147. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. *Expert Rev Pharmacoecon Outcomes Res* 2010;**10**:553–66.

148. Tufts-New England Medical Center IfCRaHPS. *The CEA Registry*. 2010.

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

150. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, *et al*.; MINDACT Investigators. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med* 2016;**375**:717–29.

151. Jones K, Burns A. *Unit Costs of Health and Social Care 2021*. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2021. https://doi.org/10.22024/UniKent/01.02.92342

152. Lyon AR, López-Fernández T, Couch LS, Asteggiano R, Aznar MC, Bergler-Klein J, *et al*.; ESC Scientific Document Group. 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). *Eur Heart J* 2022;**43**:4229–361.

153. 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 Health* 2017;**20**:A436.

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

155. Office for National Statistics. *National Life Tables: England*. London; 2021.

156. Commercial Medicines Unit. *Drugs and Pharmaceutical Electronic Market Information Tool (eMIT)*. 2023. URL: www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (accessed 9 March 2023).

157. Joint National Formulary. *British National Formulary (online)*. London: BMJ Group and Pharmaceutical Press; 2023. URL: https://bnf.nice.org.uk/ (accessed 9 March 2023).

158. 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. *N Engl J Med* 2017;**377**:1836–46.

159. Hortobagyi GN, Stemmer SM, Burris HA, Yap YS, Sonke GS, Hart L, *et al*. Overall survival with ribociclib plus letrozole in advanced breast cancer. *N Engl J Med* 2022;**386**:942–50.

160. 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 Res Treat* 2020;**183**:419–28.

161. 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. *J Clin Oncol* 2010;**28**:1011–6.

162. 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 Qual Life Outcomes* 2020;**18**:353.

163. 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 Res* 2019;**79**.

164. National Institute for Health and Care Excellence. *NICE Health Technology Evaluations: The Manual*. London: NICE; 2022.

165. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. New York: Oxford University Press; 2005.

166. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce BR; ISPOR Task Force on Good Research Practices – Modeling Studies. 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 Health* 2003;**6**:9–17.

167. 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. https://doi.org/10.1136/bmj.n71

Appendix 1 Literature search strategies

Searches

- Clinical effectiveness searches
- Cost-effectiveness searches
- EQ-5D searches

Clinical effectiveness searches

Sources searched

Host	Database	Dates covered by database[a]	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

a Indicates dates covered by database as a whole. Searches were then restricted to years 2017 onwards, as indicated in the search strategies below.

Search strategies

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946–25 April 2023
21 April 2023
1191 records

#	Searches	Results
1	exp Breast Neoplasms/	339,439
2	exp mammary neoplasms/	23,367
3	exp breast/	52,644
4	exp neoplasms/	3,822,842
5	3 and 4	32,428
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.	471,637

#	Searches	Results
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.	44,846
8	1 or 2 or 5 or 6 or 7	495,731
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.*[10] © Queen's Printer and Controller of HMSO 2019.

EMBASE 1974–2023 Week 16
21 April 2023
3184 records

#	Searches	Results
1	breast tumor/	94,110
2	exp breast/	127,654
3	exp neoplasm/	5,482,710
4	2 and 3	82,927
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.	748,134
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.	43,926
7	1 or 4 or 5 or 6	766,300
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.*[10] © Queen's Printer and Controller of HMSO 2019.

Cochrane (CDSR and CENTRAL)
21 April 2023
639 records

#	Searches	Results
1	MeSH descriptor: [Breast Neoplasms] explode all trees	17,635
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	110,452
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))	44,803
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	45,124
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.	19,573
13	#9 or #10 or #11 or #12	19,949
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.*[10] © Queen's Printer and Controller of HMSO 2019.

Web of Science Science Citation Index Expanded (1900–), Conference Proceedings Citation Index – Science (1990–)
26 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)	1676
6	(oncotype OR "recurrence score" OR 21-gene OR "21 gene") (Topic)	4301
7	(prosigna OR pam50 OR 50-gene OR "50 gene") (Topic)	1777
8	#7 OR #6 OR #5 OR #4	7417
9	#8 AND #3	3322
10	#8 AND #3 and 2023 or 2022 or 2021 or 2020 or 2019 or 2018 or 2017 (Publication Years)	1846

Search strategy adapted from Harnan *et al.*[10] © Queen's Printer and Controller of HMSO 2019.

INAH
023
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.*[10] © Queen's Printer and Controller of HMSO 2019.

WHO International Clinical Trials Registry Platform
28 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
28 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) www.asco.org/
19 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) www.esmo.org/
23 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) www.aacr.org/
25 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 Organisation (ECO) www.europeancancer.org/
25 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/
1 June 2023
21 records
Agendia https://agendia.com/
30 May 2023
45 records
Exact Sciences (aka Genomic Health) www.exactsciences.com/
26 May 2023
5 records
NanoString https://nanostring.com/
26 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 1946–3 May 2023

4 May 2023
77 records

#	Searches	Results
1	exp Breast Neoplasms/	339,611
2	exp mammary neoplasms/	23,370
3	exp breast/	52,667
4	exp neoplasms/	3,824,859
5	3 and 4	32,448
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.	471,982
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.	44,866
8	1 or 2 or 5 or 6 or 7	496,080
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”/	264,079
17	Economics/	27,499
18	exp Economics, Hospital/	25,708
19	exp Economics, Medical/	14,388
20	Economics, Nursing/	4013
21	exp models, economic/	16,199
22	Economics, Pharmaceutical/	3101

#	Searches	Results
23	exp "Fees and Charges"/	31,352
24	exp Budgets/	14,104
25	budget*.tw.	35,158
26	ec.fs.	442,581
27	cost*.ti.	142,156
28	(cost* adj2 (effective* or utilit* or benefit* or minimi*)).ab.	188,331
29	(economic* or pharmacoeconomic* or pharmaco-economic*).ti.	59,859
30	(price* or pricing*).tw.	51,979
31	(financial or finance or finances or financed).tw.	120,944
32	(fee or fees).tw.	21,211
33	(value adj2 (money or monetary)).tw.	2985
34	quality-adjusted life years/	15,581
35	(qaly or qalys).af.	14091
36	(quality adjusted life year or quality adjusted life years).af.	23,657
37	or/16-36	895,282
38	15 and 37	77

EMBASE 1974–2023 Week 17
4 May 2023
317 records

#	Searches	Results
1	breast tumor/	94,109
2	exp breast/	127,715
3	exp neoplasm/	5,486,388
4	2 and 3	82,947
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.	748,784
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.	43,946
7	1 or 4 or 5 or 6	766,964
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/	642,405

#	Searches	Results
16	exp breast/	127,715
17	exp neoplasm/	5,486,388
18	16 and 17	82,947
19	Socioeconomics/	159,524
20	Cost benefit analysis/	93,753
21	Cost effectiveness analysis/	179,610
22	Cost of illness/	21,158
23	Cost control/	75,866
24	Economic aspect/	123,726
25	Financial management/	120,747
26	Health care cost/	222,179
27	Health care financing/	13,847
28	Health economics/	35,524
29	Hospital cost/	25,189
30	(fiscal or financial or finance or funding).tw.	286,082
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	1,107,927
36	14 and 35	317

Web of Science Science Citation Index Expanded (1900–), Conference Proceedings Citation Index – Science (1990–)
4 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)	1778
5	(mammaprint OR 70-gene OR “70 gene”) (Topic)	4308
6	(oncotype OR “recurrence score” OR 21-gene OR “21 gene”) (Topic)	1682
7	(prosigna OR pam50 OR 50-gene OR “50 gene”) (Topic)	188
8	#7 OR #6 OR #5 OR #4	7430
9	#8 AND #3 and 2023 or 2022 or 2021 or 2020 or 2019 or 2018 or 2017 (Publication Years)	1854

#	Searches	Results
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

EuroQol-5 Dimensions 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 1946–3 May 2023
16 May 2023
139 records

#	Searches	Results
1	exp Breast Neoplasms/	340,146
2	exp mammary neoplasms/	23,375
3	exp breast/	52,745
4	exp neoplasms/	3,830,835
5	3 and 4	32,512
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.	256,329
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)).ti.	15,265
8	1 or 2 or 5 or 6 or 7	402,600
9	(euroqol or euro qol or eq5d or "eq 5d" or eq-5d).tw.	16,230
10	8 and 9	203
11	limit 10 to yr="2017 -Current"	139

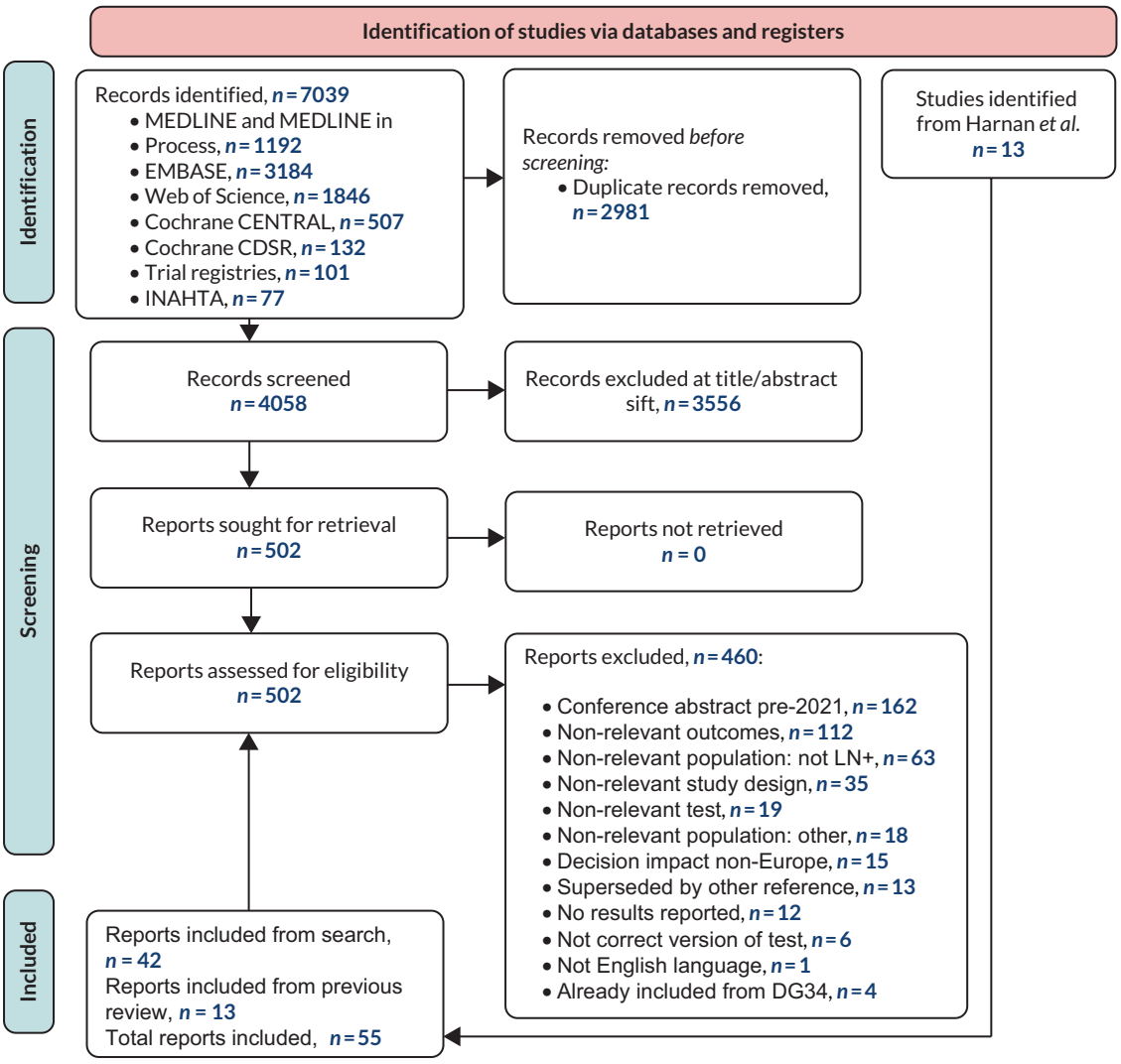
EMBASE 1974–2023 Week 19
16 May 2023
391 records

#	Searches	Results
1	exp breast tumor/	643,804
2	exp breast/	127,820
3	exp neoplasm/	5,497,634
4	2 and 3	83,006
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.	360,376
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.	16,188
7	1 or 4 or 5 or 6	687,386
8	(euroqol or euro qol or eq5d or "eq 5d" or eq-5d).tw.	29,905
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–)
16 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 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram for clinical studies



From: Page *et al.*[167]

For more information, visit: 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 RoB2.[27] Prognostic and prediction studies were assessed using the PROBAST;[28] 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 Prediction model study Risk Of Bias Assessment Tool 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 44*.

Results: risk of bias in prospective randomised controlled trials

The risk of bias in the two prospective RCTs, assessed using the Cochrane RoB2 tool,[27] is shown in *Table 45*. The two RCTs scored low risk of bias on all domains, and low risk of bias overall.

Results: risk of bias in prognostic studies

The risk of bias in prognostic studies, assessed using the PROBAST tool,[28] is presented in *Table 46* for RCT reanalyses and cohort reanalyses (within which the test was used retrospectively), and in *Table 47* 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 (see *Results: prognostic ability*). 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 (see *Results: prognostic ability*).

The following factors either were 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

TABLE 44 Risk of bias and applicability (adapted from PROBAST)

Number	Criterion	Scoring for this review
Risk of bias		
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 • Yes (predicting chemotherapy benefit): RCT or reanalysis of RCT • No (prognostic): non-nested case control or case series AND/OR some/all patients had chemotherapy • 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 • No: some eligible patients excluded (e.g. not sent for testing, insufficient tissue, test failures, missing data, AND/OR non-prospective registry) • 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 a similar way for all participants (most/all studies in this review likely to score Yes as uses standardised test) • No: test not assessed in a 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 • No: if not blinded • 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 • No: all outcomes non-standardised and not defined a priori • Unclear: if unclear
Domain 3 outcomes	Were the outcomes determined without knowledge of test (predictor) information?	• Yes: if outcome assessors blinded to test results • No: if not blinded • Unclear: if unclear
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) • No: test result may have influenced use of chemotherapy (No for observational studies of prospective use of test) • (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 • No: < 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 • No: if some enrolled patients not analysed

continued

TABLE 44 Risk of bias and applicability (adapted from PROBAST) (*continued*)

Number	Criterion	Scoring for this review
Applicability		
Domain 1 participants	Did the included participants match the review question?	• Yes: all patients in scope (HR+, HER2−, LN1–3) • Mostly: < 20% out of scope • No: > 20% out of scope • 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 • 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 • No: no outcomes matched the review question

TABLE 45 Risk of bias in prospective RCTs (using Cochrane RoB2)

	Risk of bias due to …					
RCT	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[29]	Low	Low	Low	Low	Low	Low
MINDACT Piccart 2021[30]	Low	Low	Low	Low	Low	Low

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.

Results: risk of bias in prediction studies

The risk of bias in prediction studies, assessed using the PROBAST tool,[28] is presented in *Table 48*.

The following factors may have affected results to some extent. For Domain 1 (participants), only the SWOG-8814 study[32] was a reanalysis of a 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 (see *Results: prediction of chemotherapy benefit*). 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

TABLE 46 Risk of bias in prognostic studies (retrospective reanalyses of RCTs and cohorts)

			Risk of bias									Applicability		
			Domain 1 participants		Domain 2 predictors		Domain 3 outcomes			Domain 4 analysis		Participants	Predictors	Outcomes
Reference	Cohort	Design 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[32]	SWOG-8814	RCT-R V	Y (RCT-R, ET only)	N (InT, TF)	Y	Y	Y	Y	Y	Y	N	N (> 20% LN4 +)	Y	Y
Constantinidou 2022[65]	Cyprus + Notts	Cohort-R V	Y (cohort-R, ET only)	N (InT, MD)	Y	Y	Y	Y	Y	N	N	Y	Y	Y
Drukker 2014[53]	VdV cohort, the Netherlands	Cohort-R 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[63]	ABCSG-6/8	RCT-R V	Y (RCT-R, ET only)	UC	Y	UC	Y	Y	Y	Y	Y	Y	Y	Y
Gnant 2014;[56] Filipits 2014[57]	ABCSG-8, Austria	RCT-R 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)[55]	Germany, PATH	Cohort-R V	N (cohort-R, some CT)	N (reason NR)	Y	UC	Y	Y	Y	N	N	UC	Y	Y
Laenkholm 2018[58]	DBCG, Denmark	Cohort-R V	Y (cohort-R, ET only)	N (FT, MD)	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
Lundgren 2022[62]	SBII:pre trial	RCT-R V	Y (RCT-R, ET only)	N (InT, FT, MD)	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
Mamounas 2018[49]	NSABP-28	RCT-R V	N (RCT-R, all CT)	N (InT, MS)	Y	UC	Y	Y	Y	Y	Y	UC (HER2 NR)	Y	Y
Martin 2016;[59] Martin 2014[60]	GEICAM 9906, Spain	RCT-R V	N (RCT-R, adjuvant CT)	N (MD)	Y	Y	Y	Y	Y	Y	N	N (> 20% LN4 +)	Y	Y
Mook 2009[33]	NKI and Italy	Cohort-R 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[50]	PACS01	RCT-R V	N (RCT-R, some CT)	N (FT, InT, MS)	Y	Y	Y	Y	Y	Y	Y	N (> 20% LN4 +)	Y	Y
Pu 2020[61]	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

continued

TABLE 46 Risk of bias in prognostic studies (retrospective reanalyses of RCTs and cohorts) (*continued*)

			Risk of bias									Applicability		
			Domain 1 participants		Domain 2 predictors		Domain 3 outcomes			Domain 4 analysis		Participants	Predictors	Outcomes
Reference	Cohort	Design 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?
Sestak 2018;[20] 2017[48]	TransATAC	RCT-R V	Y (RCT-R, ET only)	N (InT; FT)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sestak 2020[64]	Lobular subgroup (TransATAC + ABCSG-6/8)	RCT-R V	Y (RCT-R, ET only)	UC	Y	Y	Y	Y	Y	Y	N	Mostly (20% LN4 +)	Y	Y
Vliek 2017[54]	RASTER	Cohort-R 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; FT, failed test; InT, insufficient tissue; LN, number of positive lymph nodes; MD, missing data; MS, missing samples; N, no; NR, not reported; RCT-R, reanalysis of RCT; TF, test failure; UC, unclear; V, validation study; WHEL, Women's Healthy Eating and Living; Y, yes.

TABLE 47 Risk of bias in prognostic studies (observational studies of prospective use of test)

			Risk of bias									Applicability		
			Domain 1 participants		Domain 2 predictors		Domain 3 outcomes			Domain 4 analysis		Participants	Predictors	Outcomes
Reference	Cohort	Design 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[77]	Red Cross Hospital, Munich, Germany	Observational V	N (prospective use of test, some CT)	Y	Y	Y	Y	UC	N	Y	N	Mostly (20% LNmic)	Y	Y
Ibraheem 2020[70]	NCDB	Observational V	N (prospective use of test, CT)	N (MD, SFT)	Y	Y	Y	UC	N	Y	Y	Y	Y	Y
Massarweh 2018[69] Petkov 2016[67] Roberts 2017[68]	SEER	Observational 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[76]	WSG PlanB	Observational V	N (prospective use of test, some CT)	N (dropout, screening failure)	Y	Y	Y	UC	N	Y	Y	Y	Y	Y
Poorvu 2020[75]	Young Women's Breast Cancer Study	Observational 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; InT, insufficient tissue; LNmic, lymph node micrometastases; MD, missing data; MS, missing samples; N, no; NR, not reported; SFT, only those sent for test included; UC, unclear; V, validation study; WSG, West German Study Group; Y, yes.

TABLE 48 Risk of bias in prediction studies

			Risk of bias									Applicability		
			Domain 1 participants		Domain 2 predictors		Domain 3 outcomes			Domain 4 analysis		Participants	Predictors	Outcomes
Reference	Cohort	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 protocol?	Tests match review question?	Outcomes match review question?
Albain 2010[32]	SWOG-8814	RCT-R V	Y (RCT-R)	N (InT, TF)	Y	Y	Y	Y	Y	Y	N	N (> 20% LN4 +)	Y	Y
Mook 2009[33]	NKI and Italy	Cohort-R V	N (not RCT)	N (InT, RNA qual)	Y	Y	Y	Y	Y	Y	Y	N (> 20% ER–, 16% HER2 +)	Y	Y
Abel 2022[81]	NCDB	Observational V	N (not RCT)	N (MD, SFT)	Y	Y	Y	Y	N	Y	N	Y	Y	Y
Cao 2022 (abst)[82]														
Ibraheem 2019[71]														
Iorgulescu 2019[83]														
Kumar 2023 (abst)[84]														
Nash 2023[72]														
Weiser 2021[74]														
Weiser 2022[73]														
Petkov 2020 (abst)[80]	SEER	Observational 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)[79]	Clalit, Israel	Observational V	N (not RCT)	N (SFT)	Y	Y	Y	Y	N	Y	N	N (> 20% LNmic)	Y	Y
Stemmer 2017[66]														

Cohort-R, reanalysis of cohort study; InT, insufficient tissue; MD, missing data; MS, missing samples; 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.

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 (see *Results: prediction of chemotherapy benefit*).

The following factors either were 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.

Appendix 4 Additional tables for prognostic ability

TABLE 49 Prognostic data (Oncotype DX)

Reference Study/ cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years/other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
Oncotype DX: distant recurrence, ET monotherapy																
Sestak 2018,[20] 2017[48] 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	0–10 years 80.6 5–10 years 82.1	0–10 years 70.9 5–10 years 80.5	0–10 years 62.0 5–10 years 72.5	0–5 years: int vs. low: HR 3.84 (1.31 to 11.23) 0–5 years: high vs. low: HR 4.45 (1.19 to 16.58) 0–10 years: int vs. low: HR 1.66 (0.86 to 3.23) 0–10 years: high vs. low: HR 2.35 (0.99 to 5.60) 0–10 years: per 1 SD change: 1.39 (1.05 to 1.85) *Adj: LR vs. CTS (p = 0.06) and NPI (p = 0.1)	Y Y N N Y N*
Oncotype DX: distant recurrence, variable ET/CT																
Mamounas 2018[49] 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	0–10 years: $p < 0.001$ *0–10 years: adj HR per 50-RS: 2.42 (NR); $p < 0.001$	Y Y*
Penault-Llorca 2018[50] 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	–	–	–	0–5 years: HR per 50-RS: 4.14 (2.67 to 6.43); $p < 0.001$ *0–5 years: adj HR 3.36 (1.88 to 6.00), $p < 0.001$	Y Y*

Reference Study/ cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years/other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
Oncotype DX: DFS																
Albain 2010[32] 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	0–5 years: HR 5.55 (2.32 to 3.28); $p = 0.0002$ 0–10 years: between risk groups: $p = 0.017$ 0–10 years: HR per 50-RS: 2.64 (1.33 to 5.27); $p = 0.006$ 5–10 years: HR 0.86 (0.27 to 2.74); $p = 0.80$	Y Y Y N
Mamounas 2018[49] 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	0–10 years: $p < 0.001$ *0–5 years: adj HR per 50-RS 3.81 (2.67 to 5.43); $p < 0.001$ *0–10 years: adj HR per 50-RS 2.53 (1.90 to 3.38); $p < 0.001$ *5–10 years: adj HR per 50-RS 1.39 (0.88 to 2.19); $p = 0.16$	Y Y* Y* N*
Penault-Llorca 2018[50] 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	–	–	–	0–5 years: HR per 50-RS: 3.28 (2.18 to 4.94); $p < 0.001$ *0–5 years: adj HR 2.66 (1.62 to 4.37), $p < 0.001$	Y Y*
Kalinsky 2021[29] RxPONDER	IDFS	*n* = 5018 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	–	–	–	*0–5 years: HR per unit-RS (within RS 0–25): 1.05 (1.04 to 1.07), $p < 0.001$ (adj meno and CT)	Y*		

continued

TABLE 49 Prognostic data (Oncotype DX) (*continued*)

Reference Study/ cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years/other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
				Post-meno	All ≤ 25	–	–	–	–	–	–	–	–	–	*0–5 years: 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	–	–	–	–	–	–	–	–	–	*0–5 years: 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[51] RxPONDER	IDFS	*n* = 4015 CT + ET vs. ET Prosp RCT	LN1–3 100% HR+ 100% HER2–	White *n* = 2833	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[32] 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	0–10 years: between risk groups: $p = 0.003$ 0–10 years: HR per RS-50 : 4.42 (1.96 to 9.97), $p = 0.0006$	Y Y
Penault-Llorca 2018[50] 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	–	–	–	0–5 years: 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 0–5 years %			Risk 0–10 years/other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
Mamounas 2018[49] 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	0–10 years: $p < 0.001$ *0–10 years: adj HR per 50-RS: 3.09 (CI NR); $p < 0.001$	Y Y*
Mamounas 2018[49] 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	0–10 years: $p < 0.001$ *0–10 years: adj HR per 50-RS: 3.38 (CI NR); $p < 0.001$	Y Y*`

Adj, adjusted; CT, chemotherapy; CTS, Clinical Treatment Score (set of clinical factors); int, intermediate; LN, lymph nodes (number positive); LR, likelihood ratio; meno, menopausal; N, no; NR, not reported; prosp, prospective; RCT-R, RCT reanalysis; SD, standard deviation; sig, significant; Y, yes.

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.

TABLE 50 Prognostic data (MammaPrint)

Reference Study/cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years /other %			HR between test groups (95% CI)	Sig?[a] *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[30]	DMFS	*n* = 1176	LN1–3	High mAOL (*n* = 989)	> 0 low, ≤ 0 high	69	–	31	95.7 (50% CT)	–	89.0 (all CT)	8 years 91.0 (50% CT)	–	8 years 79.1 (all CT)	–	–
MINDACT (not on prognostics summary table since CT use per risk group was influenced by test result)		CT + ET vs. ET	100% HR+	Low mAOL (*n* = 187)	> 0 low, ≤ 0 high	92	–	8	96.3 (no CT)	–	–	8 years 94.0 (no CT)	–	–	–	–
		Prosp-RCT	100% HER2–													
	DMFI	*n* = 1176 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)	8 years 92.3 (50% CT)	–	8 years 80.9 (all CT)	–	–
				Low mAOL (*n* = 187)	> 0 low, ≤ 0 high	92	–	8	97.5 (no CT)	–	–	8 years 95.2 (no CT)	–	–	–	–
Lopes Cardozo 2022[52]	DMFI	*N* = 201 (ultra-low)	LN1–3	–	> 0.355 ultra-low	Ultra-low: 15	–	–	Ultra-low: 97.4	–	–	8 years Ultra-low: 95.2	–	–	–	–
MINDACT		Var ET/CT	99% ER+													
		Prosp-RCT	97% HER2–													
Drukker 2014[53]	DMFS	*n* = 144	LN1–3 : 74%	Age < 53 years	0.4	38	–	62	94.5	–	64.7	10 years 78.6	–	10 years 54.3 25 years 44.5	0–25 years: HR 2.24 (1.25 to 4.00); *p* = 0.01	Y
VdV cohort, the Netherlands		Var ET/CT	LN4 +: 26%									25 years				
		Cohort-R	77% ER+									NE				
			NR HER2													
Mook 2009[33] 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 years: HR 4.13 (1.72 to 9.96); *p* = 0.002 *0–10 years: adj HR: 2.99 (0.996 to 8.99); *p* = 0.051	Y N*
Vliek 2017[54] 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 years: low vs. high: HR 4.7 (1.3 to 16.2); *p* = 0.008	Y

Reference Study/cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years /other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
				All ages High mAOL (*n* = 109)	NR	40	–	60	97.7	–	86.1	95.2	–	79.5	0–10 years: low vs. high: HR 4.8 (1.1 to 21.4), *p* = 0.022	Y
MammaPrint: DFS																
Piccart 2021[30] MINDACT	DFS	*n* = 1176 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)	8 years 84.5 (50% CT)	–	8 years 74.5 (all CT)	–	–
				Low mAOL (*n* = 187)	> 0 low, ≤ 0 high	92	–	8	92.6 (no CT)	–	–	8 years 85.6 (no CT)	–	–	–	–
MammaPrint: OS and BCSS																
Piccart 2021[30] MINDACT	OS	*n* = 1176 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)	8 years 95.1 (50% CT)	–	8 years 89.1 (all CT)	–	–
				Low mAOL (*n* = 187)	> 0 low, ≤ 0 high	92	–	8	98.1 (no CT)	–	–	8 years 98.1 (no CT)	–	–	–	–
Drukker 2014[53] VdV cohort, the Netherlands	OS	*n* = 144 Var ET/CT Cohort-R	LN1–3 : 74% LN4 +: 26% 77% ER+ NR HER2	Age < 53 years	0.4	38	–	62	98.2	–	76.9	10 years 92.5 25 years 42.2	–	10 years 58.7 25 years 47.1	0–25 years: HR 1.83 (1.07 to 3.11), *p* = 0.03	Y
Jackisch 2022 (abst)[55] Germany, PATH	OS	*n* = 38 Var ET/CT Cohort-R	LN+ NR	All meno (assumed)	NR	53	–	47	–	–	–	93.3	–	40.4	–	–
Mook 2009[33] 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 years: HR 5.70 (2.01 to 16.23); *p* = 0.001 *0–10 years: adj HR 7.17 (1.81 to 28.43); *p* = 0.005	Y Y*
				All meno High AOL (*n* = 209)	NR	–	–	–	–	–	–	94	–	76	0–10 years: HR 4.12 (1.45 to 11.76); *p* = 0.008	Y

Adj, adjusted; cohort-R, cohort reanalysis; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; N, no; NR, not reported; prosp, prospective; sig, significant; var, variable; Y, yes.

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.

TABLE 51 Prognostic data (Prosigna)

Reference Study/cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution % Low	Int	High	Risk 0–5 years % Low	Int	High	Risk 0–10 years/other % Low	Int	High	HR between test groups (95% CI)	Sig?[a] *Adj
Prosigna: distant recurrence, ET monotherapy																
Sestak 2018,[20] 2017[48] 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	0–10 years 100 5–10 years 100	0–10 years 79.3 5–10 years 87.0	0–10 years 69.3 5–10 years 75.0	0–5 years: int vs. high: HR 1.30 (0.47 to 3.60) 0–10 years: int vs. high: HR 1.37 (0.69 to 2.72) HR per 1 SD change: 1.58 (1.16 to 2.15) *LR vs. CTS (p = 0.04) and NPI (p = 0.09)	N N Y Y, N*
Gnant 2014,[56] Filipits 2014[57] 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	–	–	–	0–10 years 100 5–15 years 100	0–10 years 93.6 5–15 years 87.0	0–10 years 76.1 5–15 years 75.0	5–15 years: low risk: no events 5–15 years: int vs. high: HR 3.15 (1.20 to 8.24); p = 0.020 *0–10 years: prognostic over clinical factors (p < 0.0001) *5–15 years: prognostic over clinical factors (p = 0.003)	– Y Y* Y*
Laenkholm 2018[58] DBCG, Denmark	DRFS	*n* = 1395 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	0–10 years: unadj: p < 0.001 *0–10 years: low vs. int: adj HR 0.39 (0.20 to 0.77) *0–10 years: high vs. int: adj HR 1.54 (1.04 to 2.26); p < 0.001	Y Y* Y*
						40 only	–	–	–	–	–	–	95.2 (low to int)	78.1	–	–
Prosigna: distant recurrence, variable ET/CT																
Martin 2016,[59] Martin 2014[60] 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	0–10 years: low vs. int: HR 4.4 (NR) 0–10 years: low vs. high: HR 5.8 (NR), p < 0.0001 *Prosigna vs. EPclin + clinical factors (p = 0.567)	– Y N*
Prosigna: DFS																
Pu 2020[61] WHEL Study	DFS	*n* = 344 Var ET/CT RCT-R	LN+ 100% ER+ 100% HER2–	All meno	NR	26	53	21	–	–	–	81	64	56	0–10 years: p = 0.02	Y

Reference Study/cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno status	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years/other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
Prosigna: OS and BCSS																
Lundgren 2022[62] SBII:pre trial	OS	*n* = 123 ET/none RCT-R	LN1–3 100% ER+ 100% HER2–	Pre- meno	16, 40	2	42	57	–	–	–	–	–	–	0–10 years: int vs. high: HR 1.84 (0.91 to 3.74); *p* = 0.09 *0–10 years: int vs. high: adj HR 1.32 (0.61 to 2.88); *p* = 0.48 > 10 years: int vs. high: HR 1.02 (0.54 to 1.93); *p* = 0.96 *> 10 years: int vs. high: adj HR 1.29 (0.66 to 2.53); *p* = 0.46	N N* N N*
Lundgren 2022[62] SBII:pre trial	BCFI	*n* = 123 ET/none RCT-R	LN1–3 100% ER+ 100% HER2–	Pre- meno	16, 40	2	42	57	–	–	–	–	–	–	0–10 years: int vs. high: HR 1.99 (1.08 to 3.66); *p* = 0.03 *0–10 years: int vs. high: adj HR 1.85 (0.95 to 3.58); *p* = 0.07 > 10 years: int vs. high: HR 1.19 (0.50 to 2.80); *p* = 0.70 *> 10 years: int vs. high: adj HR 1.13 (0.43 to 2.95); *p* = 0.81	Y N* N N*
Pu 2020[61] WHEL Study	BCSS	*n* = 344 Var ET/ CT RCT-R	LN+ 100% ER+ 100% HER2–	All meno	NR	26	53	21	–	–	–	90	84	77	0–10 years: *p* = 0.003	Y

Adj, adjusted; BCFI, breast cancer-free interval; cohort-R, cohort reanalysis; CT, chemotherapy; CTS, Clinical Treatment Score (set of clinical factors); int, intermediate; LN, lymph nodes (number positive); LR, likelihood ratio; meno, menopausal; N, no; NPI, Nottingham Prognostic Index; NR, not reported; RCT-R, RCT reanalysis; SD, standard deviation; sig, significant; Y, yes.

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.

TABLE 52 Prognostic data (EPclin)

Reference Study/cohort	Outcome	*N*, ET/CT Design	Nodal status HR, HER2	Meno Clin risk	Test cut-offs	Distribution %			Risk 0–5 years %			Risk 0–10 years/other %			HR between test groups (95% CI)	Sig?[a] *Adj
						Low	Int	High	Low	Int	High	Low	Int	High		
EPclin: distant recurrence, ET monotherapy																
Sestak 2018,[20] 2017[48] TransATAC	DRFI	*n* = 183 ET mono RCT-R	LN1–3 100% HR+ 100% HER2–	Post-meno	3.3	23	–	77	97.9	–	87.6	0–10 years 94.4 5–10 years 96.7	0–10 years - 5–10 years -	0–10 years 69.7 5–10 years 76.4	0–5 years: high vs. low: HR 6.00 (0.80 to 44.93) 0–10 years: high vs. low: HR 6.77 (1.63 to 28.07) 0–10 years: per 1 SD change: 1.69 (1.29 to 2.22) *LR vs. CTS (p = 0.20) or NPI (p = 0.02)	N Y Y N, Y*
Filipits 2019[63] ABCSG-6/8	DRFR	*n* = 453 ET mono RCT-R	LN1–3 100% ER+ 100% HER2–	Post-meno	3.3	35	–	65	–	–	–	0–10 years 95.6 0–15 years 84.7 5–10 years 98.2 5–15 years 87.0	–	0–10 years 80.9 0–15 years 75.1 5–10 years 90.5 5–15 years 84.0	0–10 years: HR 3.65 (1.73 to 7.68); p = 0.0003 *0–10 years: adj HR 2.68 (1.77 to 4.08); p < 0.0001 5–15 years: HR 3.00 (1.03 to 8.71); p = 0.034 *5–15 years: adj HR 3.43 (1.74 to 6.76); p = 0.0005	Y Y* Y Y*
Sestak 2020[64] 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 *EPclin vs. clinical factors (p = 0.0026)	Y Y*
Constantinidou 2022[65] 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 *Adj HR (cont score): 2.91 (1.70 to 4.97), p < 0.001	N Y*
EPclin: distant recurrence, variable ET/CT																
Martin 2016,[59] Martin 2014[60] 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 *EPclin vs. clinical factors (p = 0.0018)	Y 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

Adj, adjusted; cohort-R, cohort reanalysis; cont, continuous; CT, chemotherapy; CTS, Clinical Treatment Score (set of clinical factors); int, intermediate; LN, lymph nodes (number positive); LR, likelihood ratio; meno, menopausal; N, no; NR, not reported; RCT-R, RCT reanalysis; SD, standard deviation; sig, significant; Y, yes.
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.

DOI: 10.3310/KGFD4040

Appendix 5 Additional tables for observational data

TABLE 53 Observational data for Oncotype DX (all outcomes and analyses)

Cohort	Reference	Nodal status HR, HER2	Outcome	*N* ET/CT	Meno Age Clin	Test cut-offs	Distribution % Low	Int	High	% risk of outcome Low	Int	High	HR between test risk groups (95% CI)	Sig?[a] *Adj
Oncotype DX: distant recurrence														
Clalit, Israel	Stemmer 2017[66]	LN1mic: 42% LN1–3 : 58% 100% ER+ 100% HER2–	DRFI (0–5 years)	*n* = 709 Var ET/CT	All meno	18, 30	53	36	10	96.8 (7% CT)	93.7 (40% CT)	83.1 (86% CT)	0–5 years: low vs. high: HR 0.19 (0.09 to 0.40) 0–5 years: int vs. high: HR 0.39 (0.20 to 0.79); $p < 0.001$ *0–5 years: adj HR: low vs. high: HR 0.23 (0.11 to 0.50) *0–5 years: adj HR: int vs. high: HR 0.42 (0.20 to 0.86); $p = 0.001$	Y Y Y* Y*
						11, 25	≤ 25: 81	19	95.7 (5% CT)	96.0 (18% CT)	86.9 (77% CT)	0–5 years: $p < 0.001$	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)	0–5 years: $p < 0.001$	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)	0–5 years: $p = 0.017$	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)	0–5 years: $p = 0.458$	N
Young Women's Breast Cancer Study	Poorvu 2020[75]	LNmic, LN1–3 100% ER+ 100% HER2–	DRFS (0–6 years)	*n* = 163 Var ET/CT	Age ≤ 40	18, 30	33	42	25	0–6 years: 85.9 (83% CT)	0–6 years: 87.3 (97% CT)	0–6 years: 62.8 (98% CT)	0–6 years: $p = 0.004$	Y
						11, 25	9	54	37	0–6 years: 92.3 (79% CT)	0–6 years: 85.2 (92% CT)	0–6 years: 71.3 (97% CT)	0–6 years: $p = 0.10$	N

continued

TABLE 53 Observational data for Oncotype DX (all outcomes and analyses) (*continued*)

Cohort	Reference	Nodal status HR, HER2	Outcome	*N* ET/CT	Meno Age Clin	Test cut-offs	Distribution % Low	Int	High	% risk of outcome Low	Int	High	HR between test risk groups (95% CI)	Sig?[a] *Adj
Oncotype DX: DFS														
WSG PlanB	Nitz 2017[76]	LN1–3 100% HR+ 100% HER2–	DFS (0–5 years)	*n* = 110 Var ET/CT	All meno	0–10				94.4 (No CT)	–	–	–	–
Red Cross Hospital, Munich, Germany	Braun 2022[77]	LNmic: 20% LN1–3 : 80% 100% HR+ 100% HER2–	DFS (0–5 years)	*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[66]	LN1mic: 42% LN1–3 : 58% 100% ER+ 100% HER2–	BCSS (0–5 years)	*n* = 709 Var ET/CT	All meno	18, 30	53	36	10	99.5 (7% CT)	96.6 (40% CT)	94.3 (86% CT)	0–5 years: $p < 0.001$	Y
						11, 25	RS≤ 25:81		19	99.1 (5% CT)	98.8 (18% CT)	93.5 (77% CT)	0–5 years: $p < 0.001$	Y
						≤ 25, 26–30					98.9 (15% CT)	92.6 (67% CT)	–	–
						18–25					97.8 (31% CT)		–	–
SEER registry	Petkov 2016[67]	LN1mic, LN1–3 100% HR+ 100% HER2–	BCSS (< 5 years)	*n* = 4691 Var ET/CT	All	18, 30	57	36	7	99.0 (23% CT)	97.7 (47% CT)	85.7 (75% CT)	< 5 years: high vs. low: HR 11.0 (7.8 to 15.5) < 5 years: int vs. low: HR 3.1 (2.3 to 4.3), $p < 0.001$ *< 5 years: adj HR: high vs. low: HR 7.8 (5.3 to 11.6) *< 5 years: adj HR: int vs. low: HR 3.0 (2.1 to 4.2); $p < 0.001$	Y Y Y* Y*
				n = 328 Var ET/CT	Black ethnicity	18, 30	54	36	9	99.4 (CT NR)	98.9	91.3	< 5 years: $p = 0.4117$	N
				n = 4021 Var ET/CT	White ethnicity	18, 30	58	36	7	99 (CT NR)	97.6	84.1	< 5 years: $p < 0.0001$	Y
				n = 320 Var ET/CT	Other ethnicity	18, 30	57	34	8	98.5 (CT NR)	99.1	100	< 5 years: $p = 0.8427$	N
SEER registry	Roberts 2017[68]	LN1mic, LN1–3 100% HR+ 100% HER2–	BCSS (0–5 years)	*n* = 6483 Var ET/CT	All	18, 30	58	35	7	98.8 (CT NR)	97.3	88.5	0–5 years: $p < 0.001$ *0–5 years: adj: $p < 0.001$	Y Y*

Cohort	Reference	Nodal status HR, HER2	Outcome	N ET/CT	Meno Age Clin	Test cut-offs	Distribution %			% risk of outcome			HR between test risk groups (95% CI)	Sig?[a] *Adj
							Low	Int	High	Low	Int	High		
			OS (0–5 years)	*n* = 6483 Var ET/CT	All	18, 30	58	35	7	92.1 (CT NR)	90.9	81.7	0–5 years: $p < 0.001$ *0–5 years: adj: $p < 0.001$	Y Y*
SEER registry	Massarweh 2018[69]	LN1mic, LN1–3 100% HR+ 100% HER2–	BCSS (0–5 years)	*n* = 6437 Var ET/CT	Women	18, 30	59	35	7	98.8 (23% CT)	97.3 (48% CT)	89.2 (77% CT)	0–5 years: $p < 0.001$	Y
				n = 46 Var ET/CT	Men	18, 30	52	26	22	100 (33% CT)	100 (50% CT)	N/A (60% CT)	0–5 years: $p = 0.02$	Y
			OS (0–5 years)	*n* = 6437 Var ET/CT	Women	18, 30	59	35	7	92.2 (23% CT)	90.8 (48% CT)	83.2 (77% CT)	0–5 years: $p < 0.001$	Y
				n = 46 Var ET/CT	Men	18, 30	52	26	22	78.9 (33% CT)	100 (50% CT)	N/A (60% CT)	0–5 years: $p = 0.002$	Y
NCDB	Ibraheem 2020[70]	LN1–3 100% HR+ 100% HER2–	OS (0–5 years)	*n* = 25,029 Var ET/CT	All meno	11, 25	24	64	13	–	–	–	0–5 years: int vs. low: HR 1.15 (0.97 to 1.36) High vs. low: HR 2.94 (2.43 to 3.56) Per 10-RS: HR 1.38 (1.31 to 1.44)	N Y Y
NCDB	Ibraheem 2019[71]	LN1–3 : 97% LN4-9 : 3%	OS (0–5 years)	*n* = 13,163 Var ET/CT	All meno	11, 25	–	–	–	–	–	–	0–5 years: RS 18–25 vs. 11–17: HR 1.20 (1.07 to 1.35); $p < 0.001$ *RS 18–25 vs. 11–17: adj HR 1.15 (1.03 to 1.29); $p < 0.001$ RS 26–30 vs. 11–17: HR 1.91 (1.65 to 2.22); $p < 0.001$ *RS 26–30 vs. 11–17: adj HR 1.62 (1.38 to 1.89); $p < 0.001$	Y Y* Y Y*
NCDB	Nash 2023[72]	LN1–3 100% HR+ 100% HER2–	OS (NR, med FU 5.5 years)	*n* = 4124 Var ET/CT	Age 40–50	11, 25	–	–	–	–	–	–	*RS 26–30 vs. 0–25: adj HR 2.29 (1.49 to 4.86) *RS 31–50 vs. 0–25: adj HR 3.70 (2.03 to 6.75) *RS 51–100 vs. 0–25: adj HR 2.31 (0.78–6.86); $p < 0.001$	Y* Y* N*
NCDB	Weiser 2022[73]	LN1–3 100% HR+ 100% HER2–	OS (0–5 years)	*n* = 2691 Var ET/CT	Lobular	11, 25	–	–	–	95.5	95.5	83.8	0–5 years: $p = 0.0004$ *Adj: sig	Y Y*
NCDB	Weiser 2021[74]	LN1–3 100% HR+ 100% HER2–	OS (0–5 years)	*n* = 28,591 Var ET/CT	All	≤ 25	–	–	–	–	–	–	*0–5 years: RS 18–25 vs. RS 12–17: adj HR 1.30 (1.00 to 1.68)	Y*

Adj, adjusted; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; N, no; NR, not reported; sig, significant; var, variable; Y, yes.
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.

Appendix 6 Additional tables for chemotherapy effect within risk groups

TABLE 54 Chemotherapy effect within risk groups: Registry data for Oncotype DX (all outcomes)

Cohort	Reference	Nodal status HR, HER2	Outcome	*N*	Meno Age Clin	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)			Inter-action	Pred[a] *Adj
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High		
Oncotype DX: observational: distant recurrence																				
Clalit, Israel	Stemmer 2017[66]	LN1mic: 42% LN1–3 : 58% 100% ER+ 100% HER2–	DRFI 0–5 years	*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)[79]	LN+ 100% ER+ 100% HER2–	DRFS 0–7 years	*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[66]	LN1mic: 42% LN1–3 : 58% 100% ER+ 100% HER2–	BCSS 0–5 years	*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)[79]	LN+ 100% ER+ 100% HER2–	BCSS 0–7 years	*n* = 140	All meno	26–30	–	–	–	–	98.7	93.8	–	–	4.9	–	–	*p* = 0.024	–	–
SEER	Petkov 2020 (abst)[80]	LN1mic-LN3 100% HR+ 100% HER2–	BCSS 0–5 years	*n* = 2588	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	–	–	–	–	–	–

Cohort	Reference	Nodal status HR, HER2	Outcome	*N*	Meno Age Clin	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)			Inter-action	Pred[a] *Adj
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High		
						26–100	–	–	–	–	93.9	95.6	–	–	−1.7	–	–	–		
Clalit, Israel	Rotem 2022 (abst)[79]	LN+ 100% ER+ 100% HER2−	OS 0–7 years	*n* = 140	All meno	26–30	–	–	–	–	96.3	93.8	–	–	2.5	–	–	Not sig	–	–
NCDB	Abel 2022[81]	LN1–3 100% HR+ 100% HER2−	OS 0–5 years	*n* = 21,370	Ductal	All ≤ 25	–	–	–	–	–	–	-	–	–	*p* = 0.278	–	–	–	
				n = 6356	Lobular	All ≤ 25	–	–	–	–	–	–	–	–	–	*p* = 0.532	–	–	–	
				n = 4251	Age < 50 Ductal	All ≤ 25	–	–	–	–	–	–	–	–	–	Unadj: 0.44 (0.22 to 0.86), *p* = 0.016	–	–	–	
				n = 1062	Age < 50 Lobular	All ≤ 25	–	–	–	–	–	–	–	–	–	Unadj: 0.54 (0.14 to 2.18), *p* = 0.39	–	–	–	
NCDB (cont)	Cao 2022 (abst)[82]	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[71]	LN1–3 : 97% LN4–9 : 3% 100% HR+ 100% HER2−	OS 0–5 years	*n* = 13,163	All meno	11–17	–	–	97.7	96.5	–	–	–	1.2	-	–	Adj: 0.63 (0.40 to 0.99), *p* = 0.044 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		
				n = 3101	Age ≤ 50	All 11–25	–	–	–	–	–	–	–	–	–	–	Adj: 0.68 (0.35 to 1.32), *p* = 0.25	–	–	–

continued

TABLE 54 Chemotherapy effect within risk groups: Registry data for Oncotype DX (all outcomes) (*continued*)

Cohort	Reference	Nodal status HR, HER2	Outcome	N	Meno Age Clin	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)			Inter-action	Pred[a] *Adj
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High		
				n = 8886	Age > 50	All 11–25	–	–	–	–	–	–	–	–	–	–	Adj: 0.64 (0.47 to 0.86), *p* = 0.004	–	–	–
NCDB (cont)	Iorgulescu 2019[83]	LN1–3 100% ER+ 100% HER2–	OS 0–5 years	*n* = 2735	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)[84]	LN1–3: > 90% LN4 +: < 10% 100% HR+ 100% HER2–	OS 0–10 years	*n* = 8628	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 = 8628	Age 18–40	All ≤ 25	–	–	86.0	82.8	–	–	3.2	–		Adj: 0.43 (0.22 to 0.85)		–	–	–
				n = 8628	Age 40–50	All ≤ 25	–	–	94.7	92.2	–	–	2.5	–		Adj: 0.59 (0.39 to 0.87)		–	–	–
NCDB (cont)	Nash 2023[72]	LN1–3 100% HR+ 100% HER2–	OS NR, med FU 5.5 years	*N* = 4124	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	Reference	Nodal status HR, HER2	Outcome	*N*	Meno Age Clin	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)			Inter-action	Pred[a] *Adj
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High		
						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[73]	LN1–3 100% HR+ 100% HER2–	OS 0–5 years	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 = 2691	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[74]	LN1–3 100% HR+ 100% HER2–	OS 0–5 years	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)	–	–	–
						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	–	–	–

continued

TABLE 54 Chemotherapy effect within risk groups: Registry data for Oncotype DX (all outcomes) (*continued*)

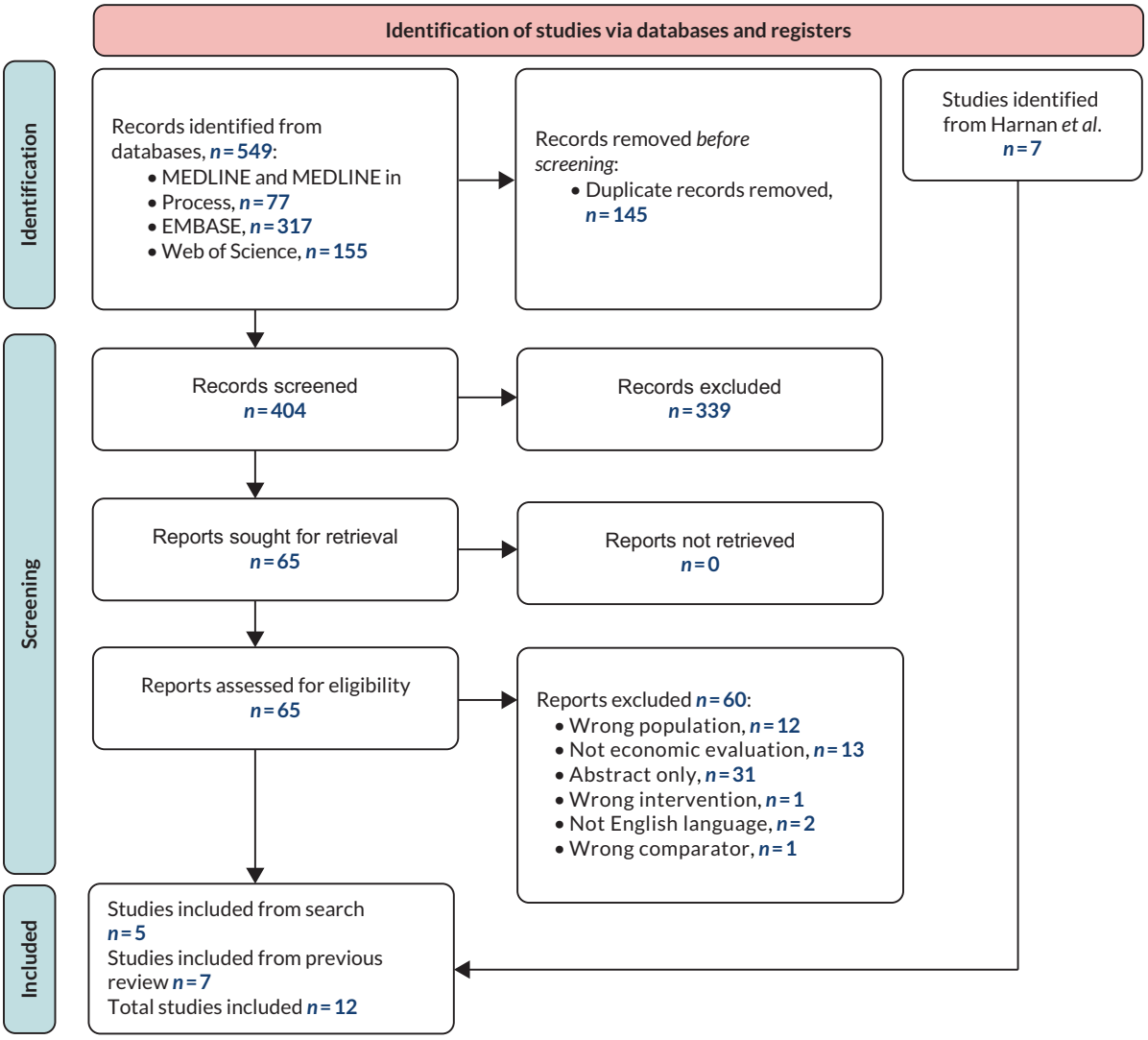
Cohort	Reference	Nodal status HR, HER2	Outcome	*N*	Meno Age Clin	Test cut-offs	% risk of outcome						Abs diff CT vs. no CT			HR for CT vs. no CT (95% CI)				
							Low		Int		High									
							CT	No	CT	No	CT	No	Low	Int	High	Low	Int	High	Inter-action	Pred[a] *Adj
						12–17	–	–	–	–	–	–	–	3.6	–	–	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), $p < 0.001$	–	–	–

Abs diff, absolute difference; adj, adjusted; CT, chemotherapy; int, intermediate; LN, lymph nodes (number positive); meno, menopausal; NR, not reported; prosp, prospective; pred, predictive of CT benefit; RS, Recurrence Score (Oncotype DX); sig, significant; unadj, unadjusted.

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.

Appendix 7 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagrams for published economic evaluations and health-related quality-of-life studies

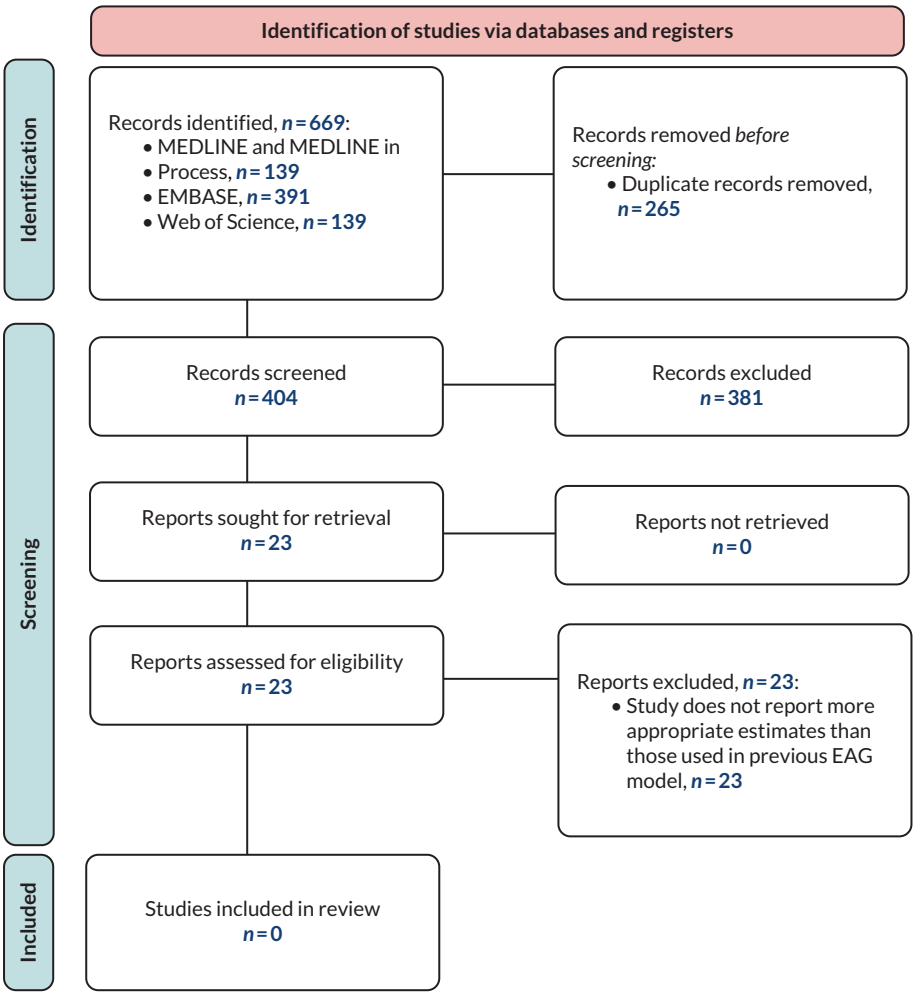
PRISMA 2020 flow diagram for economic evaluations of tumour profiling tests



From: Page *et al.*[167]

For more information, visit: www.prisma-statement.org/

PRISMA 2020 flow diagram for HRQoL associated with different health states for women with breast cancer



From: Page *et al.*[167]

For more information, visit: www.prisma-statement.org/

Appendix 8 Adjuvant chemotherapy infusion time by regimen

TABLE 55 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 600 mg/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 500 mg/m²	3	0.50	1.50
Docetaxel 100 mg/m²	3	1.00	3.00
TC (4 cycles)			
Cyclophosphamide 600 mg/m²	4	0.08	0.33
Docetaxel 75 mg/m²	4	1.00	4.00
EC90/T75 (4 + 4 cycles)			
Epirubicin 90 mg/m²	4	0.08	0.33
Cyclophosphamide 600 mg/m²	4	0.08	0.33
Docetaxel 75 mg/m²	4	0.50	2.00
EC90 (4 cycles)			
Epirubicin 90 mg/m²	4	0.08	0.33
Cyclophosphamide 600 mg/m²	4	0.08	0.33
Accelerated EC90/P (4 + 4 cycles)			
Epirubicin 90 mg/m²	4	0.08	0.33
Cyclophosphamide 600 mg/m²	4	0.08	0.33
Paclitaxel 175 mg/m²	4	1.00	4.00
C-D (6 cycles)			
Carboplatin AUC 6 (assumed 600 mg)	6	1.00	6.00
Docetaxel 75 mg/m²	6	1.00	6.00
TAC (6 cycles)			
Docetaxel 75 mg/m²	6	1.00	6.00
Doxorubicin 50 mg/m²	6	0.08	0.50

continued

TABLE 55 Infusion time for each chemotherapy regimen included in the EAG model (*continued*)

Regimen	Doses per course	Infusion time per dose (hours)	Infusion time per course
Cyclophosphamide 500 mg/m²	6	0.08	0.50
Weekly P (12 weeks)			
Paclitaxel 80 mg/m²	12	1.00	12.00
EC90/weekly P (4 cycles, 12 weeks)			
Epirubicin 90 mg/m²	4	0.08	0.33
Cyclophosphamide 600 mg/m²	4	0.08	0.33
Paclitaxel 80 mg/m²	12	1.00	12.00

Appendix 9 Cost-effectiveness acceptability curves for External Assessment Group base-case scenarios

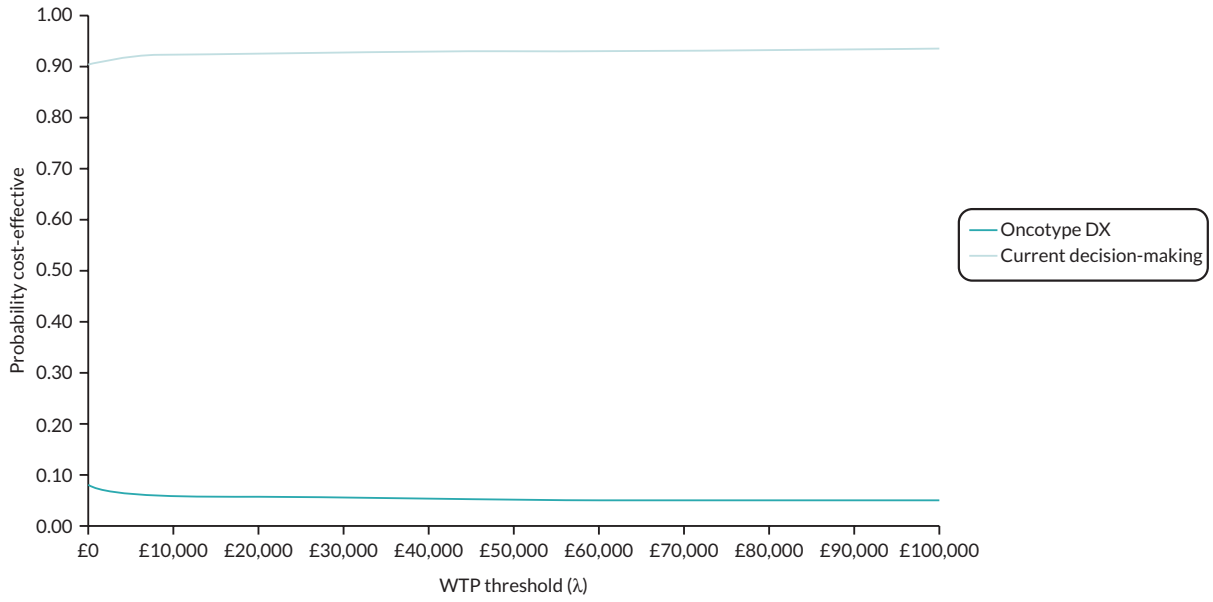


FIGURE 7 Cost-effectiveness acceptability curves, BC1 – Oncotype DX, RxPONDER pre-menopausal (predictive benefit).

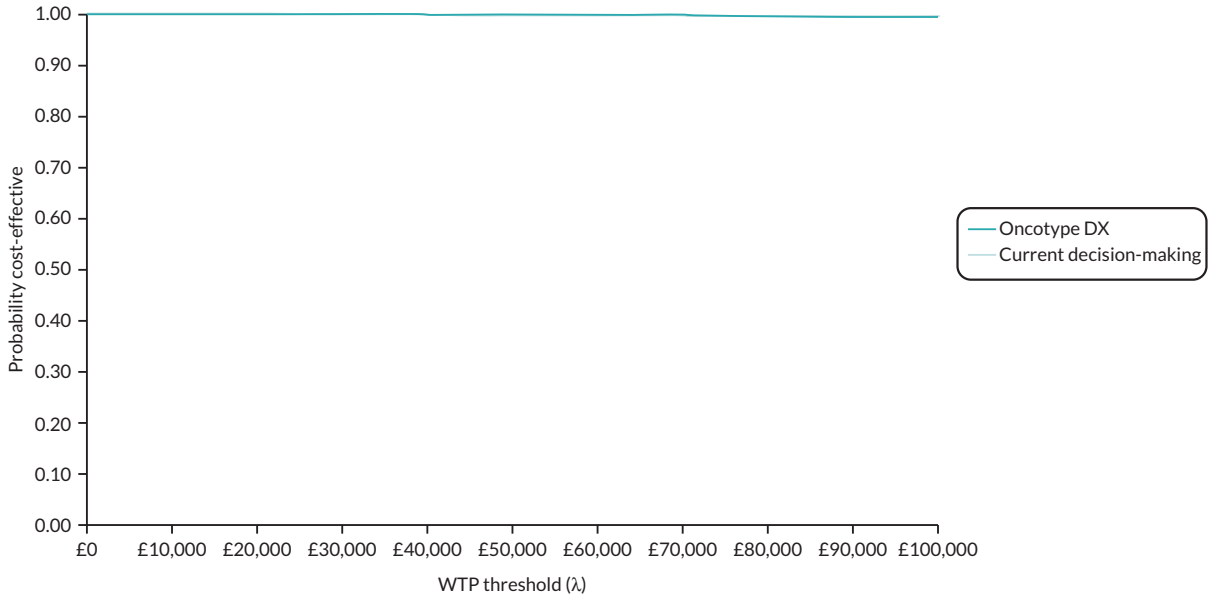


FIGURE 8 Cost-effectiveness acceptability curves, BC2 – Oncotype DX, RxPONDER post-menopausal (predictive benefit).

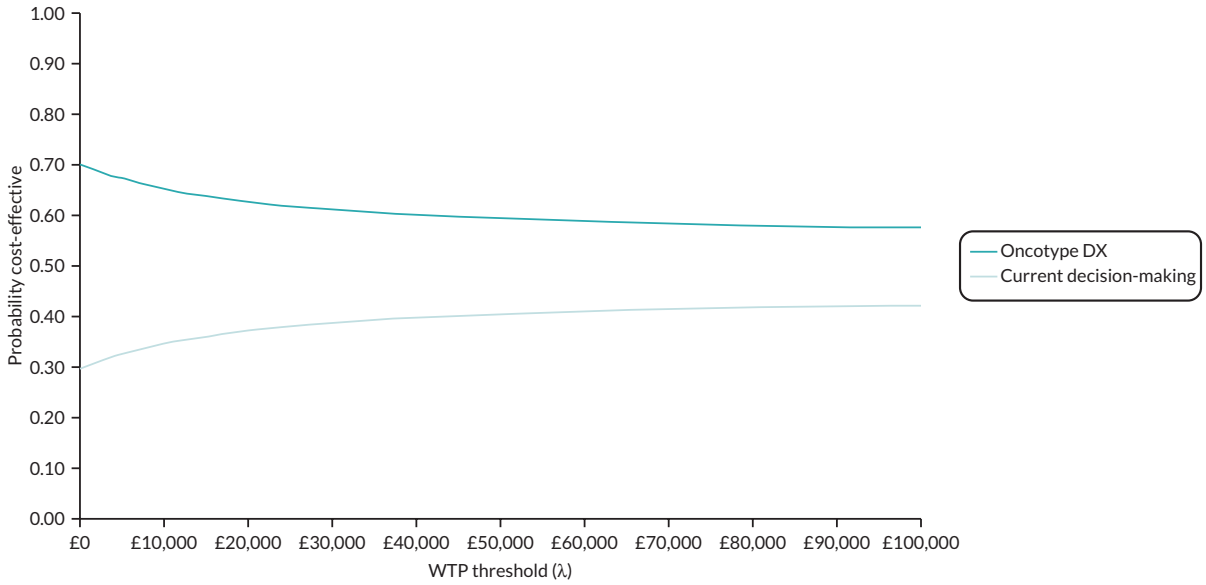


FIGURE 9 Cost-effectiveness acceptability curves, BC3 – Oncotype DX, TransATAC, post-menopausal (predictive benefit).

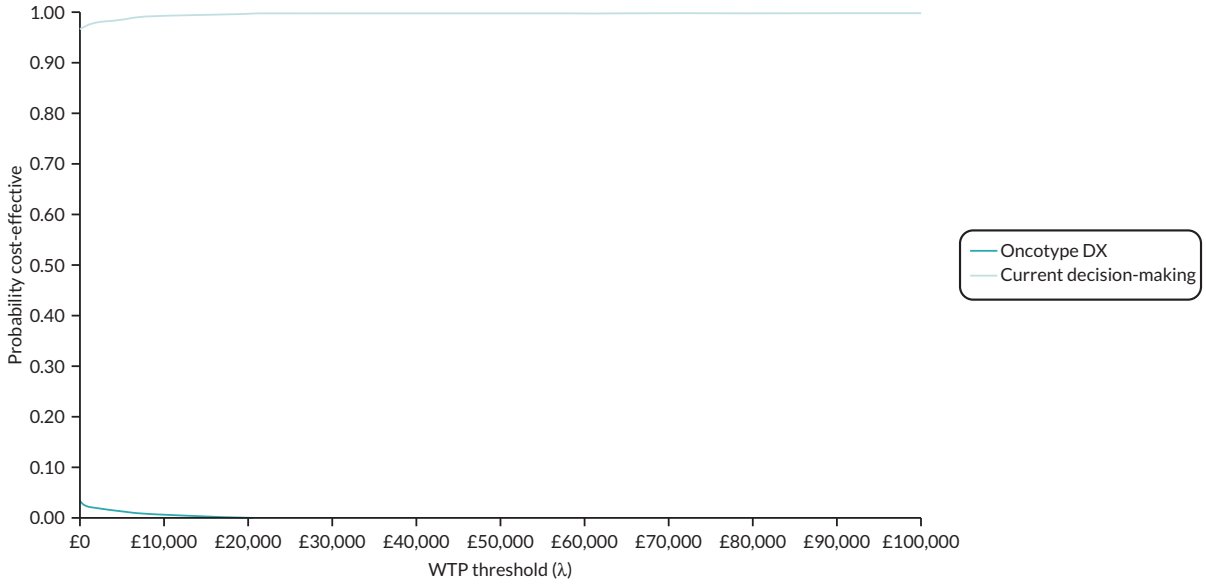


FIGURE 10 Cost-effectiveness acceptability curves, BC4 – Oncotype DX, TransATAC, post-menopausal (non-predictive benefit).

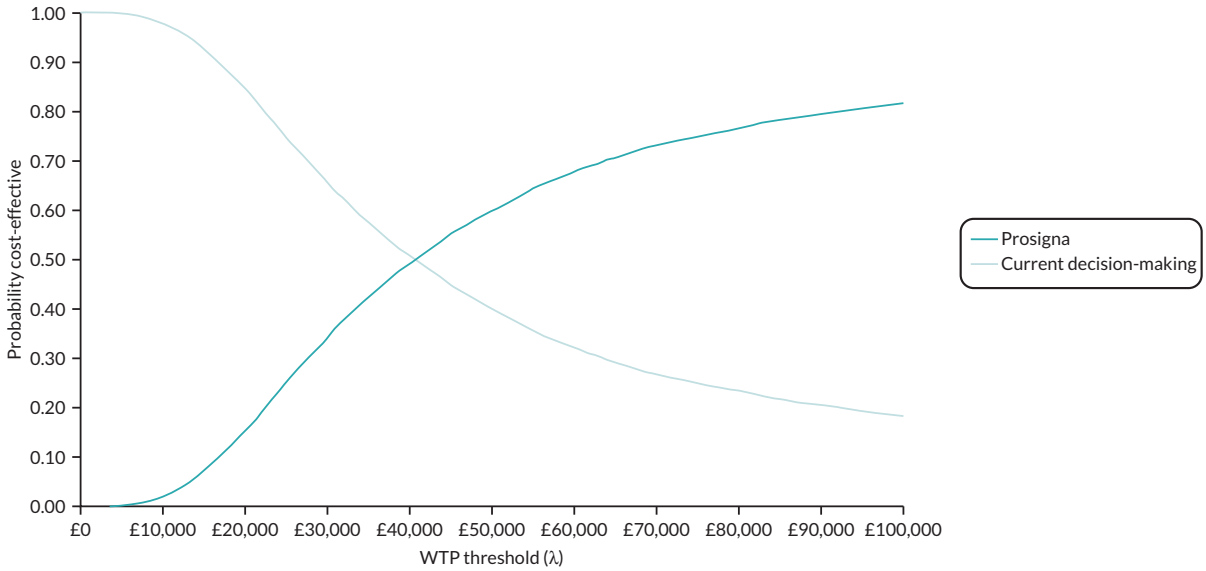


FIGURE 11 Cost-effectiveness acceptability curves, BC5 – Prosigna, TransATAC, post-menopausal (non-predictive benefit).

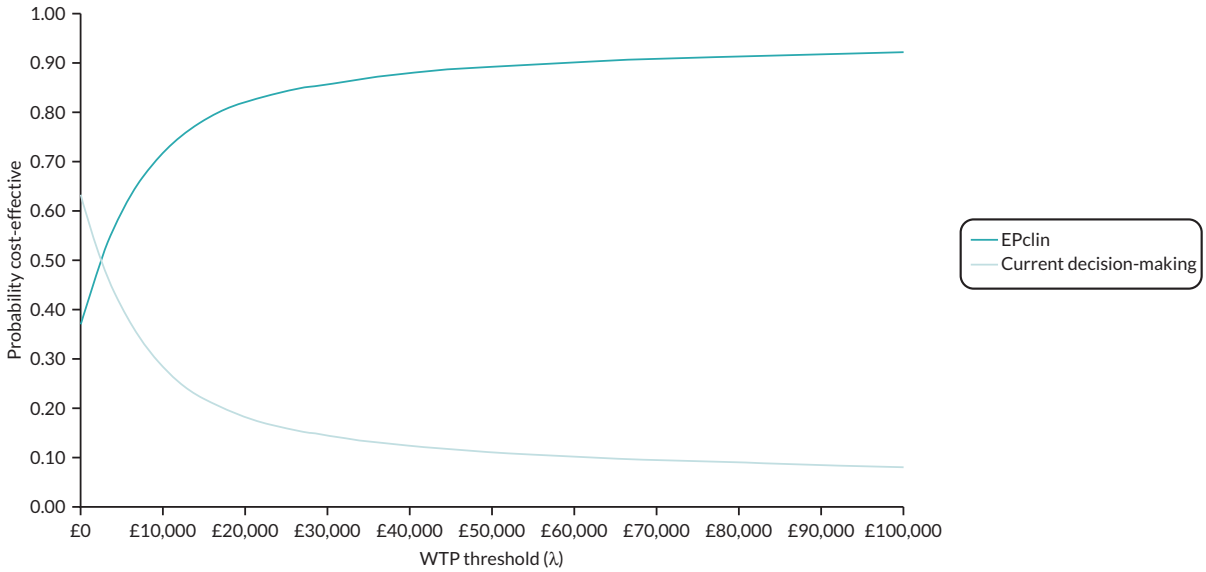


FIGURE 12 Cost-effectiveness acceptability curves, BC6 – EPclin, TransATAC, post-menopausal (non-predictive benefit).

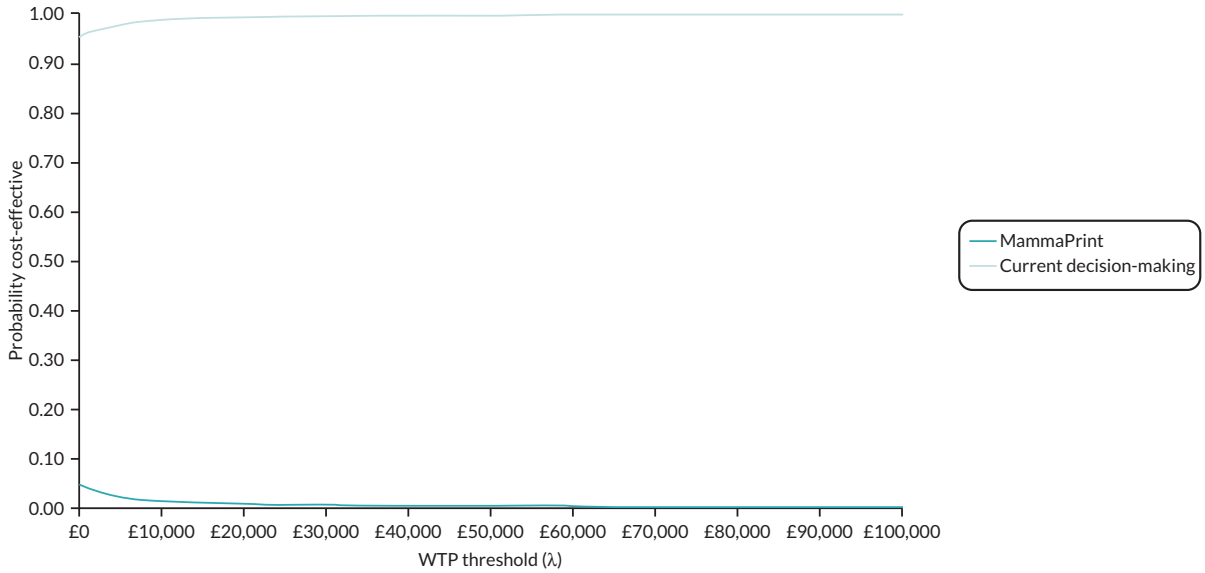


FIGURE 13 Cost-effectiveness acceptability curves, BC7 – MammaPrint, MINDACT, LN+ subgroup (non-predictive benefit).

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library