Health Information and Quality Authority

An tÚdarás Um Fhaisnéis agus Cáilíocht Sláinte

A rapid health technology assessment of gene expression profiling tests for guiding the use of adjuvant chemotherapy in early-stage invasive breast cancer

Updated: 22 January 2024

Version history

Version	Date	Specific updates
V1.0	16 February 2023	Date of first publication
V1.1	22 January 2024	Minor update to: • The Prosigna indication • One finding of the Institute of Health Economics Alberta (2019) reported in Table 2.2 • Add text detailing the potential for product price negotiations to the key findings and executive summary

Table of contents

About the Health Information and Quality Authority

The Health Information and Quality Authority (HIQA) is an independent statutory authority established to promote safety and quality in the provision of health and social care services for the benefit of the health and welfare of the public.

HIQA's mandate to date extends across a wide range of public, private and voluntary sector services. Reporting to the Minister for Health and engaging with the Minister for Children, Equality, Disability, Integration and Youth, HIQA has responsibility for the following:

- **Setting standards for health and social care services** — Developing person-centred standards and guidance, based on evidence and international best practice, for health and social care services in Ireland.

- **Regulating social care services** — The Chief Inspector within HIQA is responsible for registering and inspecting residential services for older people and people with a disability, and children's special care units.

- **Regulating health services** — Regulating medical exposure to ionising radiation.

- **Monitoring services** — Monitoring the safety and quality of health services and children's social services, and investigating as necessary serious concerns about the health and welfare of people who use these services.

- **Health technology assessment** — Evaluating the clinical and cost-effectiveness of health programmes, policies, medicines, medical equipment, diagnostic and surgical techniques, health promotion and protection activities, and providing advice to enable the best use of resources and the best outcomes for people who use our health service.

- **Health information** — Advising on the efficient and secure collection and sharing of health information, setting standards, evaluating information resources and publishing information on the delivery and performance of Ireland's health and social care services.

- **National Care Experience Programme** — Carrying out national service-user experience surveys across a range of health services, in conjunction with the Department of Health and the HSE.

Foreword

Breast cancer is the most commonly diagnosed cancer in women in Ireland. The majority of newly diagnosed cases are hormone receptor-positive (HR+), human epidermal growth factor receptor 2-negative (HER2-), early-stage (stage I-IIIa) breast cancer. Surgery is considered the first-line treatment for these patients. Following surgery, further therapy, including chemotherapy, may be needed. Chemotherapy can reduce the risk of recurrence and has contributed to declining breast cancer mortality; however, some patients may not benefit from chemotherapy and so could be spared its side effects and complications. Historically, the choice of treatment was guided by clinical and pathological factors. In more recent years, gene expression profiling (GEP) tests have been developed to further inform decisions regarding the use of chemotherapy in breast cancer.

GEP tests are intended to provide information on disease prognosis and some may also be able to identify the patients who are most likely to benefit from chemotherapy. In this way they support clinical decision-making regarding a patient's need for adjuvant chemotherapy. Currently in Ireland, the HSE reimburses the Oncotype DX® GEP test. However, there are three other GEP tests available that are not reimbursed by the HSE: MammaPrint®, EndoPredict®, and Prosigna®. The aim of this rapid health technology assessment is to provide advice to the HSE on these alternatives to Oncotype DX® that may be used to inform decision-making in relation to the management of early-stage invasive breast cancer.

Work on this rapid health technology assessment was undertaken by an Evaluation Team from the Health Technology Assessment Directorate in HIQA. A multidisciplinary Expert Advisory Group was convened to advise the Evaluation Team during the course of the rapid health technology assessment. HIQA would like to thank its Evaluation Team, the members of the Expert Advisory Group and all who contributed to the preparation of this report.

Dr Máirín Ryan

Deputy Chief Executive & Director of Health Technology Assessment

Acknowledgements

HIQA would like to thank all of the individuals and organisations who provided their time, advice and information in support of this health technology assessment. Particular thanks are due to members of the Expert Advisory Group for their time and advice and to the National Cancer Registry Ireland for sharing and advising on data used in Chapter 3.

Expert advisory group membership

The membership of the EAG was as follows:

Ms Jennifer Feighan	Patient representative
Dr Tina Hickey	Patient representative, National Cancer Control Programme Survivorship and Patient Public Engagement Steering Group
Prof. Sean Hynes	Consultant Histopathologist, University Hospital Galway
Dr Claire Kilty	Acting Head of Research, Irish Cancer Society
Dr Cillian McDowell	Health Services Researcher, HIQA
Dr Clare Meaney	Chief II Pharmacist, HSE National Cancer Control Programme
Ms Michelle O'Neill	Deputy Director of Health Technology Assessment, HIQA
Prof. Seamus O'Reilly	Consultant Medical Oncologist, Cork University Hospital
Prof. Martin O'Sullivan	Consultant Surgeon, Cork University Hospital
Dr Máirín Ryan (Chair)	Director of Health Technology Assessment, HIQA
Dr Susan Spillane	Head of Assessment, Health Technology Assessment, HIQA
Dr Conor Teljeur	Chief Scientist, HIQA
Prof. Janice Walshe	Consultant Medical Oncologist, St. Vincent's University Hospital

Members of the Evaluation Team

Dr Cillian McDowell, Dr Paul Carty, Mr Fearghal Comaskey, Dr Patricia Harrington, Ms Jingjing Jiang, Ms Leah McManus, Dr Michelle Norris, Ms Michelle O'Neill, Ms Joan Quigley, Dr Máirín Ryan, Dr Susan Spillane, Dr Conor Teljeur.

Conflicts of interest

There were no reported potential conflicts of interest for members of the Expert Advisory Group or Evaluation Team.

Key Findings and Advice to the Health Service Executive

Gene expression profiling (GEP) tests are intended to provide information on disease prognosis (that is, distant recurrence and survival) and to predict the benefit of chemotherapy (that is, to identify the people who are most likely to benefit from chemotherapy). In 2011, following a recommendation by the Health Service Executive (HSE) National Cancer Control Programme (NCCP) Technology Review Committee, the HSE began to reimburse the Oncotype DX® GEP test to guide adjuvant chemotherapy decisions in patients with lymph node negative (LN-) early-stage breast cancer. In 2019, reimbursement of Oncotype DX® was extended to patients with hormone receptor-positive (HR+) human epidermal growth factor receptor 2-negative (HER2-) and lymph node positive (LN+; 1-3 nodes) early-stage breast cancer. There are currently three other commercially available GEP tests that are not reimbursed in Ireland: MammaPrint®, EndoPredict®, and Prosigna®.

Following a request from the HSE NCCP, HIQA agreed to undertake a rapid health technology assessment (HTA) on the use of commercially available GEP tests for the purpose of guiding adjuvant chemotherapy decisions in patients with HR+, HER2-, early-stage invasive breast cancer. The aim of the rapid HTA is to provide advice to the HSE on alternatives to Oncotype DX® that may be used to inform decision-making in relation to the management of early-stage invasive breast cancer.

The key findings of this rapid HTA, which informed HIQA's advice, were:

- In Ireland, an average of approximately 1,800 cases of HR+, HER2- stage I-IIIa breast cancer were diagnosed each year between 2015 and 2019, representing the majority of new breast cancer cases.
- Surgery is considered the first-line treatment for stage I-IIIa breast cancer. Following surgery, further ('adjuvant') therapy, including chemotherapy, may be needed to control disease in the breast, lymph nodes and surrounding areas to reduce the risk of recurrence and or metastasis.
- Of women with stage I-IIIa breast cancer diagnosed between 2010 and 2015 (that is, the most recent year with complete data on therapies, receptor status, stage, and survival), 92% underwent surgery, 76% received radiotherapy, 70% received hormone therapy, and 38% received chemotherapy within the first year of diagnosis.
- While chemotherapy can reduce the risk of cancer recurrence, it is associated with short- and long-term adverse events, and not all patients with early-stage

breast cancer will benefit from its use. From a health system perspective, there are substantial costs associated with the management of such adverse events, in addition to the direct costs of providing chemotherapy.

- Choices regarding the appropriate treatment strategy are based on a range of clinicopathological factors (such as tumour size, disease stage and age). GEP tests may also be used to provide information on disease prognosis and to predict the potential for benefit from adjuvant chemotherapy. This information supports clinical decision-making regarding the need for adjuvant chemotherapy.
- Four GEP tests (Oncotype DX®, Mammaprint®, EndoPredict® and Prosigna®) are assessed within this rapid HTA. The clinical indications vary between the tests:
  - Oncotype DX® is indicated for ER+, HER2-, lymph node-negative (LN-) or lymph node-positive (LN+), stage I, II or IIIa breast cancer.
  - MammaPrint® is indicated for LN- or LN+, stage I, II or operable stage III breast cancer, irrespective of ER or HER2 status.
  - EndoPredict® is indicated for ER+, HER2-, LN- or LN+, stage I or II breast cancer.
  - Prosigna® is indicated for post-menopausal women with HR+, LN- stage I or II breast cancer and post-menopausal women with HR+, LN+ (1 to 3 positive nodes, or 4 or more positive nodes) stage II or IIIa breast cancer.
- The four GEP tests use formalin-fixed paraffin-embedded samples, which are routinely prepared during diagnostic testing. EndoPredict® and Prosigna® can be performed in local laboratories with the relevant equipment (platforms, assays, kits and reagents), while analysis of Mammaprint® and Oncotype DX® is limited to centralised laboratories in the US and additionally the Netherlands in the case of Mammaprint®. The anticipated turnaround times range from 3 to 10 days following receipt of the sample at the relevant laboratory.
- The list prices (excluding VAT) for the four GEP tests, as provided by the respective manufacturers, are:
  - Oncotype DX® €3,180
  - MammaPrint® €3,042
  - EndoPredict® €1,975
  - Prosigna® €1,934.

  It is important to note, however, that the reimbursement of Oncotype DX® within the Irish public healthcare system occurred following confidential price negotiations. Therefore, the price listed here does not necessarily reflect the

current price to the HSE. Similarly, any reimbursement of the other GEP tests would follow negotiations between the manufacturers and the HSE.

- Oncotype DX® is the only GEP test currently reimbursed by the HSE. In Ireland, on average approximately 1,800 cases of HR+, HER2- stage I-IIIa breast cancer were diagnosed each year between 2015 and 2019. As per communication with the EAG, the majority of these patients receive Oncotype DX® testing (exact figure not provided as it is commercially sensitive). Therefore, the potential annual cost of GEP tests is substantial.
- International practice regarding the use of GEP tests varies widely.
  - Of six international clinical guidelines published between 2017 and 2022:
    - Four recommended Oncotype DX® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- patients, with two of these also recommending its use in LN+ patients and another recommending its use in postmenopausal LN+ patients.
    - Three recommended MammaPrint® for guiding chemotherapy decisions in HR+, HER2-, LN- or LN+ patients, but in two guidelines, this recommendation was limited to patients at high clinical risk and in one of these the recommendation was further restricted to patients aged 50 years or more. A fourth guideline reported that the predictive ability of MammaPrint® was inconclusive due to a lack of high-quality evidence.
    - One recommended both EndoPredict® and Prosigna® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- and LN+ patients. Another recommended EndoPredict® in postmenopausal LN- and LN+ patients and that Prosigna® could be used in postmenopausal LN- patients.
    - Two, which did not differentiate between the alternative GEP tests or specify the impact of lymph node involvement, recommended the use of GEP tests in informing adjuvant chemotherapy decisions.
- Nine HTAs published between 2016 and 2022 had contrasting conclusions regarding the use of GEP tests. Evidence supporting GEP test use was more consistent for LN- patients than LN+ patients. Two HTAs concluded that GEP tests have insufficient evidence of clinical utility and or predictive ability. Two HTAs conditionally recommended their use subject to the collection of prospective data on their clinical utility.
- A systematic review was undertaken to assess the prognostic ability, predictive ability and decision impact of the four GEP tests.

- Evidence from 49 studies suggests that each of the four GEP tests likely has modest prognostic value, with greater consistency of evidence among LN- populations than LN+.
  - There was considerable variation across study designs, analytic approaches, risk cut-off scores used within GEP tests, choice of outcomes, and study populations examined. Therefore, meaningful quantification of each test's ability to predict cancer recurrence and or survival was not possible.
  - Direct comparisons of tests were sparse, making it difficult to differentiate between the prognostic abilities of the tests.
  - Each test may add prognostic value beyond that of clinical and pathological information already available to clinicians and patients, although the extent to which such value is added is unclear.
- Regarding the predictive abilities of GEP tests:
  - Three tests, Oncotype DX®, Mammaprint®, and EndoPredict®, are indicated for predictive use. Two, Oncotype DX® and Mammaprint®, have had their predictive abilities assessed in randomised controlled trials (TAILORx and RxPONDER for Oncotype DX®, and MINDACT for Mammaprint®).
  - Among LN- patients, evidence from the MINDACT trial indicated that MammaPrint® does not offer predictive value beyond that of a modified Adjuvant! Online algorithm that used data on ER status, HER2 status, nodal status, tumour grade, and tumour size.
  - Evidence for the predictive ability of Oncotype DX® from the TAILORx trial indicated that LN- women with an Oncotype DX-derived recurrence score (RS) of 11-25 could be safely spared chemotherapy. However, this was uncertain due to major limitations. Specifically, the unbalanced participant flow and participant selection in TAILORx likely biased results and limited generalisability to the Irish setting. Additionally, unlike MINDACT, TAILORx had no comparator for Onctoype DX®, meaning that its relative predictive ability was not assessed.
  - Among LN+ patients, findings from the MINDACT trial suggest that there may be merit in using MammaPrint® among high clinical risk LN+ patients aged 50 years and over, as patients in this group with a low genomic risk score may be safely spared chemotherapy. However, this finding was uncertain and due to the clinical risk assessment tool used and the use of frozen rather than FFPE tissue samples, the generalisability of the findings to an Irish context is unclear.

- The RxPONDER trial supported the predictive ability of Oncotype DX® among LN+ patients, finding that postmenopausal LN+ women with an RS 0-25 can be safely spared chemotherapy. However, it is unclear whether all participants would have received adjuvant chemotherapy in an Irish pathway in the absence of Oncotype testing as 12% had low clinical risk (that is, tumour size <2cm and Grade 1, using a modified version of Adjuvant! Online). It is likely that most would have received adjuvant chemotherapy. Therefore, Onctoype DX® may be used to spare some patients chemotherapy. However, similar to TAILORx, RxPONDER was limited by the lack of a comparator for Onctoype DX®, meaning that whether it offers predictive value beyond standard clinical and pathological information was not assessed. Additionally, these RxPONDER findings are derived from the first five years of data of a planned 15-year follow-up. It is possible that findings may change when longer-term results are available, as occurred with MINDACT.

- Regarding the impact of GEP tests on decision-making, several studies reported treatment recommendations prior to using a GEP test and the change to these recommendations after using a GEP test. Amongst these studies, all GEP tests were shown to impact treatment recommendations.

HIQA's advice to the HSE is as follows:

- In Ireland, on average approximately 1,800 cases of HR+, HER2- stage I-IIIa breast cancer were diagnosed each year between 2015 and 2019. The majority of these patients receive Oncotype DX® testing.
- Oncotype DX® is a gene expression profiling (GEP) test that is used to provide information on disease prognosis and to predict the potential for benefit from adjuvant chemotherapy. This information supports clinical decision-making regarding the need for adjuvant chemotherapy.
- There are three other commercially available GEP tests (MammaPrint®, EndoPredict®, and Prosigna®), two of which are indicated for predictive use (MammaPrint® and EndoPredict®). Currently, these three tests are not reimbursed by the HSE.
- To provide advice to the HSE on these alternative tests, a review of the clinical effectiveness evidence was undertaken. Taking into account the notable limitations of the evidence,

  Among LN- patients:

  - All four commercially available tests examined in the HTA provide prognostic information.
  - Considering predictive ability, although there are limited data to differentiate between the tests, the available evidence supports the continued use of Oncotype DX®.

  Among LN+ patients:

  - All four commercially available tests examined in the HTA were found to provide prognostic information.
  - Considering predictive ability, the evidence most strongly supports the continued use of Oncotype DX® in postmenopausal women, based on available five-year follow-up data.
- A decision to reimburse GEP tests other than Oncotype DX® should take account of differences in factors such as test indications, test costs and feasibility of use, particularly with respect to laboratory resources.

- In order to optimise the management and use of GEP tests in Ireland, consideration should be given to:
    - collecting data on GEP test use, linked to treatment and patient characteristics and outcomes, as part of a national database. These data could help clarify the clinical impact of these tests in Ireland.
    - developing guidance to outline the patient subgroups in which they should be used, the appropriate tumour sampling methods and preparation techniques, and interpretation of test results.

Executive Summary

The aim of this rapid health technology assessment (HTA) is to provide advice to the Health Service Executive (HSE) on alternative gene expression profiling (GEP) tests to Oncotype DX® that may be used to inform decision-making in relation to the management of early-stage invasive breast cancer. This rapid HTA considered the description of the technology, the epidemiology and burden of disease, and the clinical effectiveness domains of HTA.

Background

The Health Information and Quality Authority (HIQA) agreed to undertake a rapid HTA on the use of commercially available GEP tests for the purpose of guiding adjuvant chemotherapy decisions in patients with early-stage invasive breast cancer. Following a formal request from the National Cancer Control Programme, this topic was prioritised for inclusion in the HIQA HTA work plan.

Description of technology

Surgery is considered the first-line treatment for most types of breast cancer. Following surgery, adjuvant therapy, including chemotherapy, may be needed to control disease in the breast, lymph nodes and surrounding areas to reduce the risk of recurrence and or metastasis. While chemotherapy can reduce the risk of recurrence, it is associated with short- and long-term adverse events, and not all women with early-stage breast cancer will benefit from its use.

Choices regarding the appropriate treatment strategy are based on a range of clinicopathological factors (such as tumour size, disease stage and age). GEP tests may be used alongside these factors to inform adjuvant chemotherapy decisions.

GEP tests are intended to provide additional information on disease prognosis (that is, distant recurrence and survival) and the predicted benefit of chemotherapy (that is, identify the people who are most likely to benefit from chemotherapy). Four GEP tests (Oncotype DX®, Mammaprint®, EndoPredict® and Prosigna®) are assessed within this rapid HTA. Oncotype DX® is the only GEP test that is currently reimbursed by the HSE in Ireland.

Each of the four GEP tests are intended to inform adjuvant chemotherapy decisions in women with early-stage, hormone receptor positive (HR+), human epidermal growth factor receptor 2 negative (HER2-) breast cancer. However, the clinical indications vary between the tests:

- Oncotype DX® is indicated for ER+, HER2-, lymph node-negative (LN-) or lymph node-positive (LN+), stage I, II or IIIa breast cancer.
- MammaPrint® is indicated for LN- or LN+, stage I, II or operable stage III breast cancer, irrespective of ER or HER2 status.
- EndoPredict® is indicated for ER+, HER2-, LN- or LN+, stage I or II breast cancer.
- Prosigna® is indicated for post-menopausal women with HR+, LN- stage I or II breast cancer and post-menopausal women with HR+, LN+ (1 to 3 positive nodes, or 4 or more positive nodes) stage II or IIIa breast cancer.

All four GEP tests use formalin-fixed paraffin-embedded samples, which are routinely prepared during diagnostic testing. EndoPredict® and Prosigna® can be performed in local laboratories with the relevant equipment (platforms, assays, kits and reagents). In comparison, analysis of Mammaprint® and Oncotype DX® is limited to centralised laboratories in the US and also the Netherlands in the case of Mammaprint®. The anticipated turnaround times range from three to ten days following receipt of the sample at the relevant laboratory.

The list prices (excluding VAT) for the four GEP tests, as provided by the respective manufacturers, are:

- Oncotype DX® €3,180
- MammaPrint® €3,042
- EndoPredict® €1,975
- Prosigna® €1,934.

It is important to note, however, that the reimbursement of Oncotype DX® within the Irish public healthcare system occurred following confidential price negotiations. Therefore, the price listed here does not necessarily reflect the current price to the HSE. Similarly, any reimbursement of the other GEP tests would follow negotiations between the manufacturers and the HSE.

International practice regarding the use of GEP tests varies widely.

- Nine HTAs published between 2016 and 2022 had contrasting conclusions on the use of GEP tests. Several recommended their use in LN- patients, but there was less consistency with respect to LN+ patients. Two HTAs advised against GEP test use due to insufficient evidence of their clinical utility. Two HTAs conditionally recommended the use of GEP tests subject to the collection of prospective data on their clinical utility.

- Six relevant international clinical guidelines, published between 2017 and 2022, were identified. Of these:
    - Four recommended Oncotype DX® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- patients, with two of these also recommending its use in LN+ patients and another recommending its use in postmenopausal LN+ patients.
    - Three recommended MammaPrint® for guiding chemotherapy decisions in HR+, HER2-, LN- or LN+ patients, but in two guidelines this recommendation was limited to patients at high clinical risk and in one of these the recommendation was further restricted to patients aged 50 years or more. A fourth guideline reported that the predictive ability of MammaPrint® was inconclusive due to a lack of high-quality evidence.
    - One recommended both EndoPredict® and Prosigna® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- and LN+ patients. Another recommended EndoPredict® in postmenopausal LN- and LN+ patients and that Prosigna® could be used in postmenopausal LN- patients.
    - Two, which did not differentiate between the alternative GEP tests or specify the impact of lymph node involvement, recommended the use of GEP tests in informing adjuvant chemotherapy decisions.

Epidemiology

The population of interest for this rapid HTA, people with HR+, HER2-, early-stage (stage I-IIIa) invasive breast cancer, represents the majority of breast cancer cases in Ireland. Among this group, the risk of recurrence is highest in the second year post-diagnosis. Women may receive adjuvant chemotherapy to reduce this risk.

Breast cancer prognosis and response to treatment are influenced by a variety of patient- and tumour-related factors. These include the patient age, tumour size, tumour grade, histologic subtype, lymphovascular invasion of tumour cells, axillary lymph node status, and the presence of hormone and HER2 receptors.

The most common breast cancer classification system is the TNM staging system from the American Joint Committee on Cancer, which assigns patients to one of five breast cancer stages (0, I, II, III, or IV) based on the size of the tumour (T) and the degree of local infiltration, lymph node involvement (N) and the presence of metastasis (M) beyond the breast and regional lymph nodes.

In Ireland, between 2015 and 2019, there were an average of 1,641 new cases of HR+, HER2- stage I-II breast cancer and 1,806 cases of HR+, HER2- stage I-IIIa diagnosed each year.

Of all HR+, HER2- early-stage invasive breast cancers diagnosed between 2006 and 2015, approximately 42% (n=6,852) were diagnosed at stage I, 50% (n=8,225) at stage II and 8% (n=1,316) at stage IIIa. Considering lymph node status at diagnosis, 1% of stage I cases were LN+, 54% of stage II cases were LN+ and all stage IIIa cases were LN+.

Of women with stage I-IIIa breast cancer diagnosed between 2010 and 2015, within the first year of diagnosis, 92% (n=9,836) underwent surgery, 76% (n=8,106) received radiotherapy, 70% (n=7,497) received hormone therapy, and 38% (n=4,124) received chemotherapy. Of those who received chemotherapy, 99% (n=4,086) also underwent surgery.

For both stage I-II and stage I-IIIa HR+, HER2- breast cancer, estimated five- and nine-year net survival were 98% and 97%, respectively, in 2018.

In Ireland, there were an average of 178 deaths among HR+, HER2- stage I-II breast cancer patients and 216 deaths among HR+, HER2- stage I-IIIa patients each year between 2015 and 2019. The total number of deaths per five-year age band generally increased with increasing age. The age-standardised mortality rate for breast cancer cases overall has decreased over time, as reported by the National Cancer Registry Ireland.

Clinical effectiveness

A systematic review was undertaken to assess the prognostic ability (that is, the ability to accurately predict breast cancer outcomes), predictive ability (that is, the ability to identify people who will benefit most from chemotherapy) and decision impact of Oncotype DX®, MammaPrint®, Prosigna®, and EndoPredict®.

This review updates a review by the government agency Ontario Health which included 53 relevant studies. The updated search identified an additional 34 studies, resulting in a total of 87 relevant studies included in the current review. These considered the prognostic ability (n=49), predictive ability (n=24), and decision impact (n=24) of GEP tests (note that some studies reported data in more than one category).

Most prognostic and predictive studies were not designed to address whether an individual GEP test, or GEP tests in comparison to each other, can offer additional prognostic and or predictive information beyond that provided by routinely assessed clinicopathologic factors.

GEP test prognostic abilities

Evidence from 49 studies suggests that each of the four GEP tests likely has modest prognostic value for providing an estimate of a patient's likely future risk of cancer recurrence and or survival, with greater consistency of evidence among LN- populations than LN+.

There was considerable variation across study designs, analytic approaches, risk cut-off scores used within GEP tests, choice of outcomes, and study populations examined. Therefore, meaningful quantification of each test's ability to predict cancer recurrence and or survival was not possible.

Direct comparisons of tests were sparse (LN- populations: n=6; LN+ populations: n=4), making it difficult to differentiate between the prognostic abilities of the tests.

Each test may add prognostic value beyond that of other prognostic information available to clinicians and patients (that is, clinical and pathological information), although to what extent is unclear.

GEP test predictive abilities

Three tests (Oncotype DX®, MammaPrint®, and EndoPredict®) are indicated for predictive use. Of these, RCT evidence for predictive ability is available for two tests: Oncotype DX® and MammaPrint®. The associated trials are MINDACT for MammaPrint® and TAILORx and RxPONDER for Oncotype DX®.

Among LN- patients, evidence from the MINDACT trial indicated that MammaPrint® does not offer predictive value beyond that of a modified Adjuvant! Online algorithm that used data on ER status, HER2 status, nodal status, tumour grade, and tumour size.

Evidence for the predictive ability of Oncotype DX® from the TAILORx trial indicated that LN- women with an Oncotype DX-derived recurrence score (RS) of 11-25 could be safely spared chemotherapy, although this was uncertain due to major limitations. Specifically, these limitations included unbalanced participant flow and participant selection in TAILORx, which are likely to have biased results and limited the generalisability of the results to the Irish setting. Additionally, unlike MINDACT, TAILORx had no comparator for Onctoype DX®, meaning that its relative predictive ability was not assessed.

Among LN+ patients, findings from the MINDACT trial suggest that high clinical risk LN+ patients aged 50 years and over with a low MammaPrint® genomic risk score may be safely spared chemotherapy. However, this finding was uncertain and its generalisability to an Irish context is unclear due to the clinical risk assessment tool used and the use of frozen rather than FFPE tissue samples.

The RxPONDER trial supported the predictive ability of Oncotype DX® among LN+ patients, finding that postmenopausal LN+ women with an RS 0-25 can be safely spared chemotherapy. Considering the implications of these findings in the Irish setting, clinical opinion indicated that most LN+ patients would likely receive chemotherapy in the absence of GEP testing. Therefore, Onctoype DX® may be used to spare some patients chemotherapy. However, these findings are derived from the first five years of data of a planned 15-year follow-up. It is possible that these findings may change when longer-term results are available (as occurred with MINDACT). Additionally, similar to TAILORx, RxPONDER was limited by the lack of a comparator for Onctoype DX®; therefore, the predictive value of the test beyond standard clinical and pathological information was not assessed.

No trials assessed the predictive abilities of EndoPredict®, the only other test indicated for predictive use.

There were no direct comparisons of the GEP tests; therefore, differentiating between the predictive abilities of the tests was not feasible.

GEP test decision impact

Across all GEP tests, the 24 studies evaluating the impact of GEP test results on treatment recommendations found that between approximately 20% and 50% of treatment decisions were observed to have changed as a result of test administration. This suggests that the use of GEP tests impacts treatment recommendations. It is important to note that these studies did not assess whether these changes in treatment recommendations led to improved patient outcomes.

Concordance between tests

Large differences in the categorisation of patients across tests have been observed at an individual patient level. This discordance in risk group assignment, and the minimal overlap in the genes assessed across tests, suggests that there may be a number of ways of genetically predicting risk. However, despite differences in the individual level categorisation, the overall proportions of patients identified as low, intermediate, or high risk have been found to be comparable across tests.

Conclusions

This rapid HTA examined the ability of the four commercially available GEP tests (that is, Oncotype DX®, MammaPrint®, EndoPredict®, and Prosigna®) to guide adjuvant chemotherapy use among patients with HR+, HER2-, and LN- or LN+ (1-3 nodes) early-stage (stages I to IIIa) invasive breast cancer. GEP tests are intended to provide information on disease prognosis (that is, distant recurrence and survival) and the predicted benefit of chemotherapy (that is, identify the people who are most

likely to benefit from chemotherapy). Oncotype DX® is currently the only GEP test that is reimbursed by the Health Service Executive in Ireland. The majority of women with early-stage, HR+, HER2- breast cancer treated in Irish public hospitals receive Oncotype DX®.

Advice relating to GEP tests from previous HTAs and international guidelines varies substantially, despite being grounded broadly in the same evidence base. This variation in advice is likely to be influenced by issues such as study heterogeneity and flawed study designs that generally do not assess the relative prognostic and predictive value of GEP tests, resulting in a complicated and unclear evidence base. The current review found that the prognostic accuracy evidence is comparable across the four GEP tests. Three tests are indicated for predictive use (Oncotype DX®, MammaPrint®, and EndoPredict®), of which two have had their predictive ability assessed in RCTs (Oncotype DX® and MammaPrint®). Taking into account the notable limitations of the evidence, the predictive accuracy evidence is comparable between Oncotype DX® and MammaPrint® among LN- patients and strongest for Oncotype DX® among LN+ patients.

Several steps could be taken to help optimise the management and use of GEP tests. These may include the collection, as part of a national registry, of clinical and pathological characteristics of all patients in Ireland whose breast cancer specimens are sent for GEP testing. The development of a real-world database may aid in the assessment of the clinical impact and cost-effectiveness of GEP tests in Ireland. Also, there may be a role for the development of guidance to outline the patient subgroups in which gene expression profiling testing should be used, the appropriate tumour sampling methods and preparation techniques, and the interpretation of test results.

Plain language summary

The Health Information and Quality Authority (HIQA) is researching gene expression profiling tests for breast cancer at the request of the National Cancer Control Programme. HIQA focused on patients with a type of early-stage breast cancer that has spread into surrounding breast tissue and that has specific characteristics (that is, hormone receptor-positive and human epidermal growth factor receptor 2-negative breast cancer). This is the most common type of breast cancer in Ireland. The report does mention breast cancer in men, but much of the research available is on testing in women. Generally, surgery is the first treatment for these patients and it may be followed by further treatment, such as chemotherapy. Chemotherapy can reduce the risk of the cancer coming back but it can also cause some short- and long-term side-effects and is not always necessary.

After surgery to remove a tumour, gene expression profiling (GEP) tests can be performed to help decide if a patient would benefit from chemotherapy. The tests examine samples of a patient's tumour that were removed during surgery. The test results give a score which can be used to estimate whether a person has a high or low risk of the disease returning. Some tests may also be able to indicate whether a patient is more or less likely to benefit from chemotherapy. Currently in Ireland, the HSE covers the cost of one GEP test Oncotype DX®. There are three other tests the HSE does not cover the cost for: MammaPrint®, EndoPredict®, and Prosigna®. This report looked at the evidence for all four tests to provide advice to the HSE on alternatives to Oncotype DX®.

This report considered the results and quality of 87 studies that looked at how well these tests perform. The results suggested that all four GEP tests can give information on whether a person has a high or low risk of the cancer returning. However, their ability to indicate if a person is more or less likely to benefit from chemotherapy was limited. The studies examined had some limitations. Considering this, among the four tests examined, the test with the strongest evidence supporting its use was Oncotype DX®, particularly for patients whose disease has spread to the lymph nodes. Among patients whose cancer has not spread to the lymph nodes, there is currently not enough evidence to know which of the four GEP tests performs best.

We note that any decision on whether to cover the cost of GEP tests other than Oncotype DX® should take into account differences in factors such as:

- which patient groups a test can be used in
- the cost
- practicality of their use in an Irish setting.

Consideration should be given to collecting data on GEP test use in Ireland and to developing guidance on using and understanding the test results so that their management and use can be improved in the future.

List of abbreviations used in this report

aHR	adjusted hazard ratio
AJCC	American Joint Committee on Cancer
ASCO	American Society of Clinical Oncology
BRCA	BReast CAncer
CE	Conformité Européenne
EAG	Expert Advisory Group
ECIBC	European Commission Initiative on Breast Cancer
EGTM	European Group on Tumor Markers
EPclin	EndoPredict® clinical score
ER+	oestrogen receptor positive
ESMO	European Society of Medical Oncology
EUnetHTA	European Network for Health Technology Assessment
FFPE	formalin fixed paraffin embedded
GEP	gene expression profiling
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HER2-	human epidermal growth factor receptor 2-negative
HER2+	human epidermal growth factor receptor 2-positive
HIQA	Health Information and Quality Authority
HR	hazard ratio
HR+	hormone receptor positive
HSE	Health Service Executive
HTA	health technology assessment
IHC4+C	immunohistochemical 4 + Clinical score

LN-	lymph node-negative
LN+	lymph node-positive (the cancer has spread to 1 to 3 nodes)
LR	Likelihood ratio
MINDACT	Microarray in Node-Negative and 1 to 3 Positive Lymph Node Disease May Avoid Chemotherapy
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
MSAC	Medical Services Advisory Committee
NCCN	National Comprehensive Cancer Network
NCCP	National Cancer Control Programme
NCRI	National Cancer Registry Ireland
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
PAM50	Prediction Analysis of Microarray 50
PR	progesterone receptor
PROBAST	Prediction Model Risk Of Bias Assessment Tool
RCT	randomised controlled trial
RNA	ribonucleic acid
ROB2	Cochrane Risk of Bias tool 2
ROBANS	Risk of Bias Assessment tool for Non-randomized Studies
ROR	Risk of Recurrence
RS	Recurrence Score
RT-PCR	reverse transcription polymerase chain reaction
RxPONDER	A Clinical Trial RX for Positive Node, Endocrine Responsive Breast Cancer
SEER	The Surveillance, Epidemiology, and End Results Program

St. Gallen	St. Gallen International Expert Consensus
TAILORx	Trial Assigning Individualised Options for Treatment
TNM	tumour, node, metastasis
WHO	World Health Organization

1 Introduction

1.1 Background to the request

Breast cancer is the most commonly diagnosed female cancer in Ireland with over 3,000 cases registered with the National Cancer Registry Ireland (NCRI) annually; this is projected to increase by 63% by 2045.[1] Approximately 80% of all breast cancer cases in women in Ireland are diagnosed at an early stage (stage I or II).[2, 3] In Ireland, surgery is usually the first type of treatment for breast cancer.[4] As per the National Clinical Effectiveness Committee National Clinical Guideline "*Diagnosis, staging and treatment of patients with breast cancer*" published in 2015, patients with early and locally advanced breast cancer that undergo surgery may benefit from further ('adjuvant') systemic therapies.[5] Such therapies may include a combination of chemotherapy, endocrine therapy, radiotherapy, and or monoclonal antibodies (in the case of human epidermal growth factor receptor 2-positive (HER2+) breast cancer).[5] While chemotherapy can reduce the risk of recurrence and has contributed to declining breast cancer mortality,[6, 7] not all women with early-stage breast cancer benefit from chemotherapy.

Historically, the choice of treatment in breast cancer was guided by clinical and pathologic factors.[8] In more recent years, gene expression profiling (GEP) tests have been developed to further inform decisions regarding the use of adjuvant chemotherapy in breast cancer. These tests are intended to improve the categorisation of patients with respect to the risk of recurrence or death and to identify patients most likely to benefit from chemotherapy. In 2011, following a recommendation by the Health Service Executive (HSE) National Cancer Control Programme (NCCP) Technology Review Committee, the HSE began to reimburse the Oncotype DX® GEP test to guide adjuvant chemotherapy decisions in patients with lymph node negative (LN-) early-stage breast cancer.[9] In 2019, reimbursement of Oncotype DX® was extended to patients with HR+, HER2-, and LN+ (1-3 nodes) breast cancer.[10] At the time of writing, there are a number of other commercially available GEP tests that are not currently reimbursed in Ireland:

- EndoPredict®
- MammaPrint®
- Prosigna®.

Following a request from the HSE NCCP, HIQA agreed to undertake a rapid health technology assessment (HTA) on the use of commercially available GEP tests for the purpose of guiding adjuvant chemotherapy decisions in patients with early-stage invasive breast cancer. The primary aim of the rapid HTA is to provide advice

to the HSE on alternatives to Oncotype DX® that may be used to inform decision-making in relation to the management of early-stage invasive breast cancer.

1.2 Terms of reference

The terms of reference of the rapid HTA, agreed with the HSE NCCP, are to:

- describe the commercially available gene expression profiling (GEP) tests used to inform decision making in patients with early-stage invasive breast cancer and consider the organisational implications associated with their use
- describe the burden of disease associated with breast cancer in Ireland, with a particular focus on hormone receptor-positive (HR+), human epidermal growth factor receptor 2-negative (HER2-) early-stage (I, II, or IIIa) invasive breast cancer
- review the prognostic and predictive accuracy of commercially available GEP tests, along with their impact on clinical decision-making, in patients with early-stage invasive breast cancer
- based on the evidence in this assessment, provide advice to the HSE on alternative GEP tests to Oncotype Dx® to inform decision-making in relation to the management of early-stage invasive breast cancer.

1.3 Overall approach

Following an initial scoping of the available evidence, the terms of reference of this rapid HTA were agreed between HIQA and the HSE NCCP. HIQA appointed an Evaluation Team to carry out the assessment.

HIQA convened an Expert Advisory Group comprising representation from relevant stakeholders including the HSE NCCP, clinicians with specialist expertise in medical oncology, breast cancer surgery and histopathology, representation from people affected by cancer, and representation from the Irish Cancer Society research department.

The role of the Expert Advisory Group is to inform and guide the process, provide expert advice and information, and to provide access to data where appropriate. A full list of the membership of the Expert Advisory Group will be available in the acknowledgements section of this report.

The Terms of Reference of the Expert Advisory Group are to:

- contribute to the provision of high quality and considered advice by HIQA to the Health Service Executive
- contribute fully to the work, debate and decision-making processes of the group by providing expert guidance, as appropriate
- be prepared to provide expert advice on relevant issues outside of group meetings, as requested
- provide advice to HIQA regarding the scope of the analysis
- support the Evaluation Team led by HIQA during the assessment process by providing expert opinion and access to pertinent data, as appropriate
- review the project plan outline and advise on priorities, as required
- review the draft report from the Evaluation Team and recommend amendments, as appropriate
- contribute to HIQA's development of its approach to HTA by participating in an evaluation of the process on the conclusion of the assessment
- notify the project lead if a nominee can no longer participate or contribute to the process as non-participation may require alternative EAG membership to be sought.

The Terms of Reference of the rapid HTA will be reviewed by the Expert Advisory Group at its meeting. Draft chapters will be circulated to the Expert Advisory Group for review and will be discussed at one formal meeting of the group, with amendments made, where appropriate. The completed assessment will be submitted to the HSE NCCP as advice, and published on the HIQA website.

2 Description of technology

Key points

- Surgery is considered the first-line treatment for most types of breast cancer. Following surgery, further ('adjuvant') therapy, including chemotherapy, may be needed to control disease in the breast, lymph nodes and surrounding areas to reduce the risk of recurrence and or metastasis. While chemotherapy can reduce the risk of recurrence, it is associated with short- and long-term adverse events, and not all women with early-stage breast cancer will benefit from its use.
- Choices regarding the appropriate treatment strategy are based on a wide range of clinicopathological factors (such as tumour size, disease stage and age). Gene expression profiling (GEP) tests may also be used, alongside these factors, to inform decisions regarding the use of adjuvant chemotherapy.
- GEP tests are intended to provide additional information on disease prognosis (that is, distant recurrence and survival) and the predicted benefit of (that is, identify the people who are most likely to benefit from) chemotherapy. Four GEP tests (Oncotype DX®, EndoPredict®, Mammaprint® and Prosigna®) are assessed within this rapid HTA. Oncotype DX® is the only GEP test that is currently reimbursed by the Health Service Executive in Ireland.
- Each of the four GEP tests are intended for informing decisions regarding the use of adjuvant chemotherapy in women with early-stage, hormone receptor positive (HR+), human epidermal growth factor receptor 2 negative (HER2-) breast cancer. However, the clinical indications vary between the tests.
  - Oncotype DX® is indicated for ER+, HER2-, lymph node-negative (LN-) or lymph node-positive (LN+), stage I, II or IIIa breast cancer.
  - MammaPrint® is indicated for LN- or LN+, stage I, II or operable stage III breast cancer, irrespective of ER or HER2 status.
  - EndoPredict® is indicated for ER+, HER2-, LN- or LN+, stage I or II breast cancer.
  - Prosigna® is indicated for post-menopausal women with HR+, LN- stage I or II breast cancer and post-menopausal women with HR+, LN+ (1 to 3 positive nodes, or 4 or more positive nodes) stage II or IIIa breast cancer.
- All four GEP tests use formalin-fixed paraffin-embedded samples, which are routinely prepared during diagnostic testing. EndoPredict® and Prosigna® can be performed in local laboratories with the relevant equipment (platforms, assays, kits and reagents), while analysis of Mammaprint® and Oncotype DX®

is limited to centralised laboratories in the US and additionally the Netherlands in the case of Mammaprint®. The anticipated turnaround times vary by test, ranging from 3 to 10 days following receipt of the sample at the relevant laboratory.

- International practice regarding the use of GEP tests varies widely:
  - There were contrasting conclusions from nine HTAs, published between 2016 and 2022, regarding the use of GEP tests. While four HTAs explicitly recommended their use in LN- patients, there is less consistency with respect to LN+ patients. Two HTAs advised against the use of GEP tests due to insufficient evidence of clinical utility and or predictive ability. Two HTAs conditionally recommended the use of GEP tests subject to prospective data collection on the clinical utility of these tests.
  - Six relevant international clinical guidelines, published between 2017 and 2022, were identified. Of these:
    - Four recommended Oncotype DX® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- patients, with two of these also recommending its use in LN+ patients and another recommending its use in postmenopausal LN+ patients.
    - Three recommended MammaPrint® for guiding chemotherapy decisions in HR+, HER2-, LN- or LN+ patients, but in two guidelines this recommendation was limited to patients at high clinical risk and in one of these the recommendation was further restricted to patients aged 50 years or more. A fourth guideline reported that the predictive ability of MammaPrint® was inconclusive due to a lack of high-quality evidence.
    - One recommended both EndoPredict® and Prosigna® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- and LN+ patients. Another recommended EndoPredict® in postmenopausal LN- and LN+ patients and that Prosigna® could be used in postmenopausal LN- patients.
    - Two, which did not differentiate between the alternative GEP tests or specify the impact of lymph node involvement, recommended the use of GEP tests in informing adjuvant chemotherapy decisions.

2.1 Introduction

The purpose of this chapter is to describe the role and use of the four gene expression profiling (GEP) tests being assessed in this rapid HTA. To facilitate understanding, a brief description of breast cancer is provided in section 2.2. Section 2.3 provides an overview of the existing clinical pathway for patients with breast cancer and the current use of Oncotype DX® in Ireland. Section 2.4 details the technical characteristics of the four GEP tests assessed within this rapid HTA (that is, Oncotype DX®, EndoPredict®, MammaPrint® and Prosigna®) and the logistical considerations associated with their use. The use of GEP tests internationally to assist in clinical decision-making regarding the use of adjuvant chemotherapy is summarised in section 2.5.

2.2 Breast cancer

Breast cancer is a disease in which abnormal cells in the breast begin to grow and divide uncontrollably, eventually forming a growth or abnormal mass of tissue (that is, a tumour). Breast cancer begins in the tissue of the breast, most commonly originating in the cells that line the milk ducts.

Classifications of breast cancer can be based on cancer stage, histological grade, the presence (positivity) or absence (negativity) of certain receptors, molecular subtype, or by the pattern of gene expression. The most commonly used breast cancer classification system is the TNM staging system from the American Joint Committee on Cancer (AJCC).[11] This assigns patients to one of five breast cancer stages (0, I, II, III, or IV) based on the size of the tumour (T), the degree of local infiltration and lymph node involvement (N), and the presence of metastasis (M) beyond the breast and regional lymph nodes. Early-stage breast cancer is variably defined, but is generally categorised as stage I–II or I–IIIa.

The presence of hormone and or human epidermal growth factor receptor 2 (HER2) receptors can affect the treatment course and potentially breast cancer outcomes. Common receptors tested for in breast cancer cells include the oestrogen receptor (ER), the progesterone receptor (PR), and the human epidermal growth factor receptor 2 (HER2). The target population for the current rapid HTA is people with hormone receptor-positive (HR+), HER2 negative (HER2-), early-stage invasive breast cancer.

The natural history of breast cancer, and the morbidity and mortality associated with HR+, HER2- early-stage breast cancer in Ireland is described in Chapter 3.

2.3 Overview of existing clinical pathway

This section provides an overview of the clinical pathway for the screening, diagnosis and treatment of breast cancer in Ireland, including a description of the role of GEP tests and the current use of Oncotype DX®. This section is informed by both national and international literature describing best practice for the screening, diagnosis and treatment of breast cancer.[3, 9, 10, 12-20]

2.3.1 Screening

BreastCheck is the national population-based screening programme for breast cancer in Ireland, run as part of the National Screening Service.[21] It was first introduced in the East of the country in 2000 before being rolled out to other regions in 2007.[2, 22, 23] BreastCheck offers breast cancer screening to all women aged 50 to 69 years every two years, with the aim of reducing breast cancer-related mortality in the population through the detection of early cancerous or pre-cancerous cells and treatment of the disease at the earliest possible stage.[13, 24]

Screening involves a mammogram of the breast (that is, an X-ray of the breast tissue) at a BreastCheck clinic or mobile screening unit. Women who receive an abnormal screening result are recalled for further tests and evaluation. This may include undergoing a biopsy where a tissue sample is removed from the breast to undergo laboratory testing to see if the cells are cancerous.

2.3.2 Diagnosis and pathology

Breast cancer can be diagnosed following an abnormal BreastCheck screening result, or following referral for further investigation due to the presence of signs or symptoms that could be associated with breast cancer. When breast cancer is suspected, a diagnosis is made by 'triple assessment'; this involves clinical examination in combination with imaging and confirmation by pathological assessment.[12, 18]

Clinical examination includes bimanual palpation of the breasts and regional lymph nodes along with collection of patient information such as full medical history, family history of cancer, and menopausal status. Imaging includes bilateral mammography and or ultrasonography of the breast and regional lymph nodes. Magnetic resonance imaging (MRI), may also be considered in particular circumstances (such as if there are discrepancies between conventional imaging and clinical examination or in the case of familial breast cancer associated with BRCA mutations).[18]

Pathological evaluation is based on a core needle biopsy of the primary tumour and or fine-needle aspiration cytology of the axillary lymph nodes (if involvement is suspected). The pathological report generally includes the presence or absence of

ductal carcinoma in situ, the histological type, grade and immunohistochemistry (IHC), and evaluation of ER status. For invasive cancer, the pathological report may also include IHC evaluation of PR and HER2 expression or HER2 gene amplification. Final pathological diagnosis is made according to the World Health Organization (WHO) classification and the eighth edition of the AJCC TNM staging system.[11, 25] Tumours are generally grouped according to subtypes (described in Chapter 3), defined by routine histology and IHC data, to inform prognosis and treatment decisions. The four main breast cancer tumour subytpes are Luminal A, Luminal B, HER2-enriched and basal-like.[26]

Staging describes how extensive the breast cancer is, including the size of the tumour, whether the cancer has spread to lymph nodes or distant parts of the body, and what its biomarkers are. Clinical staging is based on the results of tests done before surgery, including physical examination, biopsy and imaging tests (for example, mammography, ultrasonography and MRI scans). Pathological staging is based on the examination of tissue from the breast or lymph nodes removed during surgery. In newly diagnosed asymptomatic early-stage breast cancer, distant metastases are rare and most patients will not benefit from additional laboratory testing; therefore, routine staging evaluations are directed at locoregional disease (that is, limited to the region surrounding the tumour in the breast).[12, 18] Further imaging tests (such as a chest or bone scan and abdominal imaging) may be considered in some instances (for example, in patients with large tumours (>5 cm)).

Classifications of breast cancer can be based on cancer stage, histological grade, the presence (positivity) or absence (negativity) of certain receptors, molecular subtype, or by the pattern of gene expression. As noted in section 2.2, the classification system most commonly used for staging of breast cancer is the AJCC TNM staging system, which considers the size of the tumour and its extension, the degree of lymph node involvement, and the presence of metastasis.[11] This classification system is described further in Chapter 3.

Laboratory work-up occurs before surgery (see section 2.3.3) and systemic (neo)adjuvant therapy. In early-stage breast cancer, the most important prognostic factors are the expression of HR and HER2, the number of involved regional lymph nodes, tumour histology, size, grade and the presence of lymphovascular invasion.[18] GEP tests, as described in section 2.3.4, may also be used to complement pathological assessment by providing additional prognostic and or predictive information that may help guide treatment decisions.

2.3.3 Treatment

The main treatments available for breast cancer are:

- surgery
- radiotherapy
- chemotherapy
- endocrine (hormone) therapy
- biological therapy or immunotherapy (that is, targeted monoclonal antibody therapy).[4]

Treatment of breast cancer is complex and may involve the combination of local modalities (such as radiotherapy, chemotherapy, hormone therapy and immunotherapy) and supportive measures (such as management of treatment-related adverse effects and toxicities, and psychosocial effects),[27] delivered in diverse sequences, depending on the type and stage of the breast cancer. Breast cancer surgery, considered the first-line treatment for most types of breast cancer,[4, 14] can entail lumpectomy or mastectomy followed by mammographic surveillance. Lumpectomy (also referred to as a partial mastectomy or wide excision) is a breast-conserving surgery in which the area of the breast containing the tumour is removed, with the remaining breast tissue preserved. In Western Europe, 60-80% of all newly diagnosed breast cancers can be treated by lumpectomy.[18] Mastectomy involves the removal of the whole breast. For invasive cancers in which the cancer has spread to nearby lymph nodes, the affected lymph nodes can also be removed during the surgery or as a separate operation (that is, sentinel lymph node biopsy or axillary lymph node dissection).[28, 29]

Systemic therapies, including chemotherapy and endocrine therapy, may be given prior to surgery (referred to as 'neoadjuvant' therapy), with the aim of downstaging the tumour. The National Clinical Effectiveness Committee National Clinical Guideline, published in 2015, made recommendations regarding the use of neoadjuvant and adjuvant systematic therapy.[5] The guideline recommends that all breast cancer patients at moderate or high risk of recurrence should be considered for adjuvant chemotherapy, and that any patient who is a candidate for adjuvant systemic therapy can be considered for neoadjuvant systemic therapy. Guidelines by the European Society of Medical Oncology (ESMO) recommend that, where clinically indicated, neoadjuvant therapy should begin within two to four weeks of diagnosis and staging.[18] Following surgery, women with early and locally advanced breast cancer may need further (adjuvant) therapy including cytotoxic chemotherapy, radiotherapy, monoclonal antibodies and or endocrine therapy, again depending on the molecular subtype of the cancer.

These adjuvant therapies can assist in controlling disease in the breast, lymph nodes and surrounding areas and in reducing the risk of the cancer spreading (metastasis) and or recurrence.[5, 20] The ESMO guidelines recommend the initiation of adjuvant systemic therapy within three to six weeks after surgery,[18] as efficacy may reduce

with delays beyond this.[30] Different types of breast cancer respond differently to alternative types of adjuvant therapy. HR+ cancers are likely to respond to oral hormone therapies, such as tamoxifen and aromatase inhibitors, taken for between five and 10 years.[28, 29] A 2012 meta-analysis of randomised trials showed that 20% of HR+ patients will experience long-term recurrence if treated with hormone therapy only; thus, chemotherapy may also be necessary within this population.[31, 32]

Choices regarding the appropriate treatment strategy (including the use of neoadjuvant and adjuvant therapy) are based on a wide range of factors, including the:

- molecular characteristics of the tumour (size, location, number of lesions and extent of lymph node involvement)
- stage and grade of disease
- pathology (subtype, biomarkers and gene expression)
- patient's characteristics (age, menopausal status, comorbidities and general health status)
- predicted response to alternative treatment types and the associated short- and long-term toxicities
- person's risk of relapse or likely course of disease
- person's individual preferences.[8, 18]

Approximately 38% of women with HR+, HER2-, stage I-IIIa breast cancer received chemotherapy between 2010 and 2015 (see section 3.4.2 for further detail). While chemotherapy can reduce the risk of recurrence and has contributed to declining breast cancer mortality,[6, 7] not all women with early-stage breast cancer benefit from chemotherapy.

As chemotherapy affects both cancerous and normal healthy cells, patients are likely to experience chemotherapy-related side effects. These side effects can vary from patient to patient and according to the chemotherapy agents used, the dose and the duration of treatment.[33] Short-term side effects that may occur during the treatment period include loss of appetite, nausea or vomiting, mouth soreness, constipation or diarrhoea, tiredness, reduced resistance to infections, and hair loss. Longer-term side effects that persist beyond the treatment period can include infertility, development of bone thinning conditions (for example, osteopenia and osteoporosis), hypertension, neuropathy, problems with cognitive function, lung damage and heart problems. Although some of these side effects, which can have a significant impact on quality of life,[34-36] can be prevented or managed with appropriate medications,[4] others may impose a long-term burden and or increase risk of other negative health outcomes (for example, heart failure).[37, 38]

Chemotherapy is generally administered by intravenous infusion as part of an outpatient appointment every two to three weeks over a period of four to eight months.[4] Public patients without a medical card who receive outpatient care or who are required to stay overnight in hospital have been subject to a statutory charge of €80 per visit up to a maximum of €800 for 10 visits in a consecutive 12-month period, although these charges will be abolished from April 2023.[39, 40] Other out-of-pocket costs incurred by the patient as part of chemotherapy include the potential loss of earnings, caregiver expenses, and transport and travel expenses.[41, 42] From a health system perspective, there are substantial costs associated with the management of adverse events, in addition to the direct costs of providing chemotherapy.[43] As such, preventing unnecessary chemotherapy is beneficial to both the patient and the health system.

2.3.4 Use of GEP tests

In addition to the factors outlined above, decisions regarding the use of adjuvant chemotherapy in breast cancer treatment can be further informed by the use of GEP tests. These tests are intended to provide additional information on disease prognosis (for example, the risk of distant metastasis) and the predictive benefit (that is, identify the people who are most likely to benefit from) of chemotherapy. It is claimed that the prognostic information provided by these tests, described in more detail in section 2.4, is more accurate than that of other non-genetic clinicopathological criteria.[29]

GEP testing is generally performed following surgery with the results used in conjunction with other clinicopathological factors to guide treatment decisions. Improved information on a patient's prognostic risk and or their likely response to adjuvant chemotherapy may help to target chemotherapy to the patients who will benefit most from it.[44, 45] Considerable uncertainty has been noted in several systematic reviews regarding the relative effectiveness of GEP tests,[29, 44-46] due to low-quality evidence and scarce head-to-head comparisons of the tests. This uncertainty in the evidence base is reflected in differing international policies regarding the reimbursement of GEP tests, as described in section 2.5.

Other clinical tools have been developed to provide prognostic information to help clinicians and patients with early-stage breast cancer better understand the potential benefit of adjuvant therapy. Tools that have been used commonly include Adjuvant! Online (currently unavailable for use), PREDICT, and the Nottingham Prognostic Index.[9, 47-49] The Breast Cancer Index is an additional tool available to inform decisions regarding whether endocrine therapy should be extended from five to 10 years in particular patients.[50] The use of these clinical tools is not assessed within this rapid HTA.

Clinical practice in Ireland

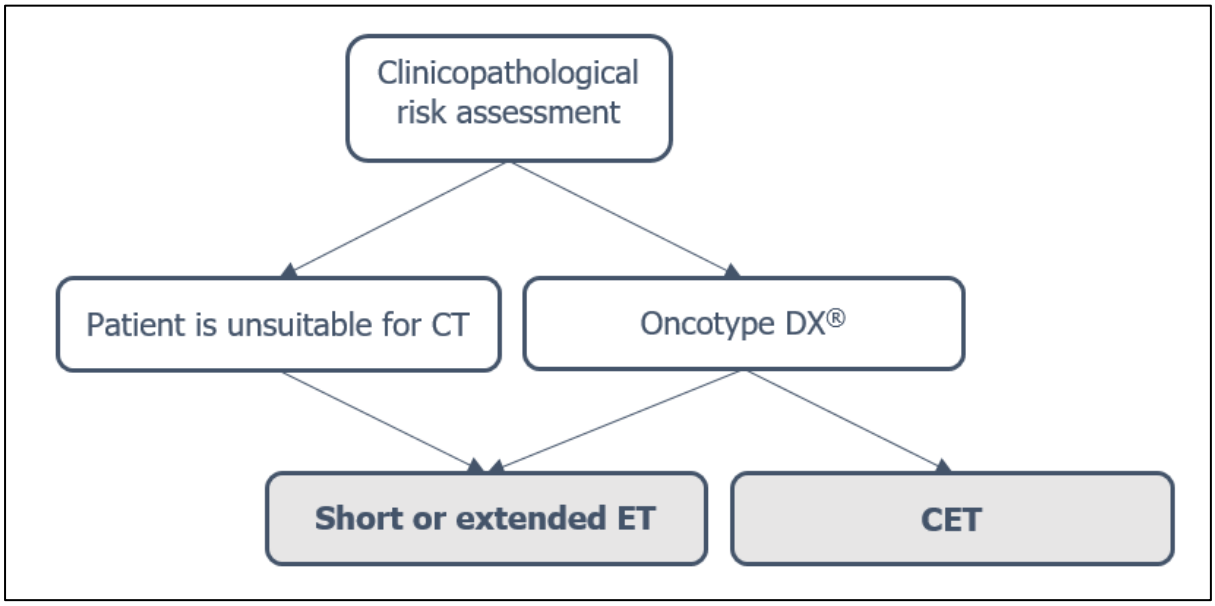
In 2011, following advice from the Irish Society of Medical Oncology,[9] the HSE approved reimbursement of Oncotype DX® to support decision-making regarding the use of adjuvant chemotherapy in patients with HR+, HER2-, LN- early stage breast cancer.[10, 15] Prior to 2011, Oncotype DX® had been available for use on a trial basis only, as part of the Trial Assigning Individualised Options for Treatment (TAILORx trial).[16, 51] In their advice to the HSE in 2011,[9] The Irish Society of Medical Oncology recommended the establishment of a National Registry to collect clinical and pathological characteristics of patients whose breast cancer specimens were sent for Oncotype DX® testing. To date, such a registry has not been established. In 2019, reimbursement of Oncotype DX® was extended to patients with HR+, HER2-, LN+ (1 to 3 nodes) breast cancer, in line with recommendations by the National Cancer Control Programme (NCCP) Systemic Anti-Cancer Therapy Breast Clinical Advisory Group.[10] No other GEP tests are currently reimbursed by the HSE to support decision-making regarding the use of adjuvant chemotherapy in patients with early-stage breast cancer.

Following breast cancer surgery, pathological information (such as nodal status and tumour size) is available to the treating medical oncologist. In conjunction with a multidisciplinary team, the medical oncologist uses this pathological information along with clinical information (such as patient age and comorbidities) to classify the patient on the basis of their risk of recurrence. As per the 2015 National Clinical Effectiveness Committee National Clinical Guideline, all breast cancer patients at moderate or high risk of recurrence should be considered for adjuvant chemotherapy.[5]

Chemoendocrine therapy is generally indicated in those at high risk of recurrence while treatment for those at low risk of recurrence is typically limited to endocrine therapy. In consultation with the patient, the Oncotype DX® test score may be used to guide the decision regarding the appropriate choice of adjuvant therapy. The information provided by the Oncotype DX® test score is intended to supplement, rather than replace, standard clinicopathological criteria. In Ireland, the Oncotype DX® test is not undertaken in patients who are not considered candidates for adjuvant chemotherapy for clinical reasons (for example, due to the patient's age, health status or menopausal status), and may not be undertaken in some patients for pathological reasons (for example, low tumour grade or small tumour size).[52]

Figure 1.1 provides an indicative outline of the role of Oncotype Dx® in the clinical management algorithm based on the information provided in the eligibility criteria.[9] However, it should be acknowledged that the precise use of Oncotype DX® could be subject to variation in clinical practice.

Figure 1.1. Use of Oncotype DX® within clinical management algorithm following breast cancer surgery



Key: CET – chemoendocrine therapy; CT – chemotherapy; ET – endocrine therapy.
* Clinical risk assessment is based on clinicopathological factors (such as age, ER and HER2 status, tumour size, nodal status, tumour grade, etc.) conducted according to national and international clinical guidelines.

Potential chemotherapy drug cost savings from use of Oncotype DX®

Three studies, published between 2014 and 2021, reviewed changes in chemotherapy use in Ireland since the introduction of Oncotype DX®.[3, 16, 17] The first study was a retrospective cohort study of 471 consecutive female ER+, LN- patients diagnosed between January 2006 and May 2013.[16] Of these, 51% (n=240) underwent Oncotype DX® testing (including 97 patients enrolled in the TAILORx trial) and 46% (n=216) received chemotherapy. The study showed a stepwise reduction in chemotherapy use over time from 63% (n=34) of all patients treated between 2006 and 2007 to 32% (n=47) of those treated between October 2011 and May 2013. The authors reported (in the study abstract only) that 138 of the 240 (58%) patients who underwent Oncotype DX® testing were spared chemotherapy.

The second study was a retrospective, cross-sectional study of 583 patients with ER+, LN- breast cancer for whom chemotherapy treatment status was available; all underwent Oncotype DX® testing between October 2011 and February 2013.[17] The authors assigned each patient a theoretical recommendation on whether or not they would receive chemotherapy; this recommendation was based on tumour grade and was informed by a survey of Irish breast oncologists. The proportion of patients who might have received chemotherapy on the basis of this recommendation was compared with the proportion that actually received chemotherapy. The authors reported that Oncotype DX® test results led to a change in the theoretical

chemotherapy recommendation in 59% (n=345) of patients; for 339 of these patients, the change was from a recommendation of chemotherapy to no chemotherapy. Compared with the theoretical pre-Oncotype DX® treatment recommendations, the study concluded that the use of Oncotype DX® would have led to a net saving of approximately €1,361 per patient during the study period. This finding was based on the total cost of using chemotherapy in all patients with an assumed pre-Oncotype DX® treatment recommendation for chemotherapy compared with the cost of testing all patients with Oncotype DX® and the resulting total cost of chemotherapy when following the Oncotype DX® treatment recommendations.

An updated analysis was performed on the records of 955 patients with HR+ LN- breast cancer that underwent Oncotype DX® testing between October 2011 and February 2019, and for whom an Oncotype DX® score was available.[3] When compared with the theoretical recommendation, the authors reported that the Oncotype DX® test results led to a reduction in the proportion of patients receiving chemotherapy (from 88% (n=839) to 27% (n=262)). The authors estimated that this represented a net saving of approximately €1,238 per patient in chemotherapy costs avoided over the study period following the use of Oncotype DX®, these patients would also have been spared the negative short- and long-term side effects of chemotherapy.

List prices

The price of the GEP test is an important consideration in decision-making regarding its potential use. The current list prices (excluding VAT), as provided by the respective manufacturers, for each of the GEP tests assessed in this rapid HTA are as follows:

- Oncotype DX® is €3,180
- MammaPrint® is €3,042
- EndoPredict® is €1,975
- Prosigna® is €1,934.

It is important to note that the reimbursement of Oncotype DX® within the Irish public healthcare system occurred following confidential price negotiations.[15] Therefore, the price listed here does not necessarily reflect the current price to the HSE. Similarly, any reimbursement of the other GEP tests would follow negotiations between the manufacturers and the HSE. Nonetheless, an average of approximately 1,800 cases of HR+, HER2- stage I-IIIa breast cancer were diagnosed each year between 2015 and 2019 and the majority of these patients received Oncotype DX® testing, as per communication with the EAG (exact figure is commercially sensitive). Therefore, the potential annual cost of GEP tests is substantial.

2.4 Technical characteristics of GEP tests

This section provides an overview of the technical characteristics of the GEP tests. To facilitate understanding, sections 2.4.1 to 2.4.3 provide an overview of relevant terminology, a description of gene expression profiling, and the main laboratory-based analysis techniques employed by the GEP tests. The information specific to each individual GEP test was sourced from each of the relevant manufacturer's websites and the international literature.

2.4.1 Terminology

Prognosis refers to the likely outcome or course of a disease, including the likelihood of recovery or recurrence. Accordingly, a prognostic factor provides information on the likely clinical outcome at the time of diagnosis independent of or in the absence of further therapy. In the context of breast cancer, prognostic information may be provided by relevant clinicopathological factors (such as tumour size, histologic grade, patient age, etc.) which act as disease markers for cancer growth, invasion and metastatic potential.[53, 54]

In contrast, a **predictive** factor provides information on the likelihood of response to a given therapy (for example, chemotherapy) such that this information can be used to tailor individualised treatment plans. The most established predictive markers in early-stage breast cancer are the expression of ER (for predicting response to hormone therapy) and HER2 (for predicting response to HER2-directed therapy).[55] Although prognostic and predictive factors can be categorised separately, factors such as the presence of HR and HER2 are considered both prognostic and predictive.

To be considered either prognostic or predictive for application in clinical practice, a test for a disease marker should demonstrate the following:

- Analytical validity – this refers to the technical aspects of a test, such as accuracy, reproducibility and reliability.
- Clinical validity – the ability of a test to categorise a target population into subgroups that differ according to biological or clinical outcomes. However, this does not imply that a prognostic or predictive factor should be used in patient care.
- Clinical utility – this implies that a prognostic or predictive factor is useful in the care of patients, based on evidence that it improves patient outcomes.[56, 57]

In addition to these three factors, a test used for prognostic or predictive purposes should be able to provide significant and independent value, be determined feasible, reproducible, and interpretable, and should not use tissue required for other tests.[55]

High-quality prospective randomised clinical trials (RCTs), in which the proposed prognostic or predictive factor is defined as the primary endpoint, are required to determine the clinical utility of a proposed prognostic or predictive factor.[58] However, such evidence is commonly obtained from studies which adopt a 'prospective-retrospective' design (that is, a retrospective design using prospectively collected specimens), ideally from RCTs.[59] The design of studies assessing the clinical utility of GEP tests for prognostic and predictive purposes in early-stage breast cancer are described further in Chapter 4.

2.4.2 Gene Expression

Gene expression is the process by which the information encoded within DNA is converted into functional products in the cell, such as proteins. DNA is first transcribed to produce messenger ribonucleic acid (mRNA), and mRNA carries the genetic message through the cell to the location of protein formation within the cell.[60]

Gene expression profiling measures mRNA levels, thereby determining the patterns of genes that are expressed by a cell population at a given time. In doing so, the functions of a cell can be observed. Specific patterns of gene expression (gene signatures) may be found in studies to correlate with known clinical outcomes, such as prognosis or response to treatment. Such patterns may be validated in independent groups of tumours, ideally by different techniques and teams. If validated and found to have clinical utility, such signatures could be applied prospectively to help guide treatment decisions in newly diagnosed patients.

Laboratory-based GEP assays investigate the expression of specific gene panels by measuring their RNA levels in breast cancer specimens using various analytical techniques, these including reverse transcription polymerase chain reaction (RT-PCR) and DNA microarrays (summarised in section 2.4.3). In conjunction with other available clinical and pathological investigations, the information provided by GEP assays can be used to inform prognosis and treatment decisions in early-stage invasive breast cancer.

2.4.3 Analysis Techniques

RT-PCR and microarray technologies, described below, are examples of analysis techniques that can be used for diagnostic purposes across a variety of disease areas.

RT-PCR is commonly used for the analysis of gene expression and quantification of RNA in research and clinical settings. It comprises a genetic amplification technique that measures RNA expression levels. In RT-PCR, the RNA genome is converted into

its complementary DNA molecule, which enables amplification using PCR. Quantification of the DNA produced after each round of amplification is based on fluorescent dyes that are inserted into double-stranded DNA or by modified short single stranded DNA probes that fluoresce when bound with complementary DNA. In RT-PCR, this process is automated and measurements are made after each round of amplification.[61]

A **microarray** is a miniaturised device comprising a small flat surface (usually a glass slide or 'chip') onto which ordered arrangements of individual samples are positioned, allowing simultaneous detection of thousands of genes.[62, 63] In this technique, a collection of DNA sequences (called 'probes') are arranged in a grid on a microarray such that the identity of each fragment is known through its location on the array. These DNA sequences are then used to detect the concentration of the corresponding complementary RNA sequences (called 'targets') present in a sample. Microarray-based methods allow the use of small sample volumes, efficient analyses and high throughput.[64]

RT-PCR is used for the Oncotype DX® and EndoPredict® tests, while a microarray is used in the MammaPrint® test. There are also proprietary systems that combine these techniques such as the nCounter Dx Analysis System® which is used in the Prosigna® test (described in section 2.4.7).

2.4.4 Oncotype DX Breast Recurrence Score®

Oncotype DX®, developed by Genomic Health (acquired by Exact Sciences), is a Conformité Européenne (CE)-marked test intended to estimate the 10-year risk of distant recurrence and predict the likelihood of benefit from adjuvant chemotherapy. It is intended for use in women with HR+, HER2-, LN-, micrometastases or LN+ (1 to 3 positive nodes) early-stage (stage I, II or IIIa) invasive breast cancer.

The Oncotype DX® test uses RNA extracted from a formalin fixed paraffin embedded (FFPE) tissue sample which is analysed using RT-PCR. The molecular diagnostic test reveals the underlying tumour biology by quantifying the expression of 21 genes, stratified into two gene categories:

- 16 cancer-related genes correlated with distant recurrence-free survival
- five reference (normalisation) genes.[9]

A proprietary algorithm is used to generate a Recurrence Score® (RS) result, on a numeric scale from 0 to 100, based on the expression of cancer-related genes present in a FFPE tissue sample. On the basis of this RS result, the risk of distant metastasis at 10 years, if treated with endocrine therapy only for five years, is quantified along with the likely benefit from the addition of chemotherapy.

Various cut-offs have been proposed based on the RS.[31] The standard cut-offs were originally defined by the manufacturer as low (0 to 10), intermediate (11 to 25) and high (26 to 100) risk according to the results of the Trial Assigning IndividuaLised Options for Treatment (TAILORx) published in 2015.[65] These cut-offs have been updated according to the latest evidence from the TAILORx, RxPONDER (A Clinical Trial RX for Positive Node, Endocrine Responsive Breast Cancer), and the NSABP B-20 and SWOG-8814 trials.[66-71] In women aged over 50 years with LN- or LN+ disease the RS categories are now interpreted as follows:

- RS of ≤25 indicates a low risk of distant recurrence
- RS of ≥26 indicates a high risk of distant recurrence.[68, 72]

Based on exploratory analyses of the TAILORx trial, the manufacturer proposed alternative risk categories for patients aged 50 years or younger.[66, 73] In women aged 50 years or younger with LN- or LN+ disease, an RS of ≥26 can be considered high risk of distant recurrence. In women aged 50 years or younger with LN- disease, an RS of ≤15 can be considered low risk, while the manufacturer has defined two categories (RS of 16 to 20 and RS of 21 to 25) in which there is an intermediate risk of distant recurrence and in whom there may be benefit from adjuvant chemotherapy. In women aged 50 years or younger with LN+ disease, the manufacturer has defined two further categories (RS of 0 to 13 and RS of 14 to 25) in which there is an intermediate risk of distant recurrence and in whom there may be benefit from adjuvant chemotherapy.

The RS result is calculated independently of clinical and pathological factors (for example, age, tumour size and grade), but can be combined with these factors using the Recurrence Score® Pathology Clinical calculator. This calculator has not been independently validated and is not CE-marked.[44, 45, 74]

Samples for the Oncotype DX® test are analysed in a centralised laboratory in the US (the licensed Genomic Health® laboratory). Results are returned via a secure online platform,[75] with an anticipated return time between 7 to 10 calendar days following receipt of the specimen in the US laboratory. Turnaround times of up to 12 days have been reported in the UK.[76] Further details on the Oncotype DX® test are presented in Table 2.1. As noted in section 2.3.5, the Oncotype DX® test has been reimbursed in Ireland since 2011.

2.4.5 MammaPrint®

MammaPrint® is a CE-marked assay developed using untreated patient samples by Agendia. It is indicated as a prognostic (to predict the likelihood of distant metastases developing within 5 to 10 years of an initial diagnosis of breast cancer) and predictive (whether the person with breast cancer would benefit from adjuvant

chemotherapy) test. MammaPrint® is intended for use in women with LN- or LN+ (1 to 3 positive nodes) early-stage (stage I, II and operable stage III) breast cancer, irrespective of ER or HER2 status. MammaPrint® uses a microarray platform to measure the expression of 465 reference genes and 70 cancer-related genes. These include genes associated with the following segments of the metastatic pathway:

- growth and proliferation
- angiogenesis
- local invasion
- entering the circulation
- survival in the circulation
- entering organs from the circulation
- adaption to the microenvironment at a secondary site.

Based on the combined expression of the 70 genes, the MammaPrint® Index is calculated using a proprietary algorithm which indicates the risk of developing distant metastases within 10 years in the absence of any adjuvant endocrine therapy or chemotherapy. The MammaPrint® test score is provided on a continuous scale ranging from −1 to +1, which corresponds to a binary risk classification system (that is, high or low risk). A MammaPrint® Index score of -1 to 0 indicates a high risk of distant metastasis within five to 10 years, while a MammaPrint® Index score of 0 to 1 indicates a low risk of distant metastasis within five to 10 years.(77)

Samples for Mammaprint® are analysed in Agendia's centralised laboratories in the US. Turnaround times for MammaPrint®, as reported by the manufacturer, are approximately four to five days from receipt of the sample.(78) Further details on MammaPrint® are presented in Table 2.1.

2.4.6 Endopredict®

EndoPredict® is a CE-marked assay developed by Myriad Genetics. It is indicated as a prognostic (that is, informs the likelihood of distant metastasis) and predictive (that is, predicts the likelihood of benefit from adjuvant chemotherapy) test. Specifically, assuming five years of endocrine therapy only, the test estimates the risk of distant recurrence within 0 to 10 years and within 5 to 15 years, and estimates the absolute benefit from chemotherapy at 10 years. EndoPredict® is suggested for use in women with ER+, HER2-, LN- or LN+ (1 to 3 positive nodes) early-stage (stage I or II) breast cancer.(79) The test is based on analysis of RNA, extracted from a FFPE, by RT-PCR. To generate a molecular score, the test measures the expression of 12 genes:

- three proliferative-associated genes
- five oestrogen receptor signalling-associated genes

- three normalisation (reference) genes
- one control gene.[80]

After RT-PCR, the raw data are exported to an online evaluation CE-marked software (EndoPredict® Report Generator) which calculates the 12-gene molecular score (also referred to as EP score in the scientific literature) and an EndoPredict® clinical risk score (EPclin risk score).

The 12-gene molecular score categorises patients' risk of distant disease recurrence on a scale of 0 to 15. A 12-gene molecular score of <5 indicates low risk of distant recurrence in the next 10 years. A 12-gene molecular score of ≥5 indicates a high risk of distant recurrence in the next 10 years. The 12-gene molecular score comprises the molecular score only and is not the final test result.

The EPclin risk score categorises the risk of distant recurrence within 10 years (assuming five years of endocrine therapy), by incorporating information on nodal status and tumour size in addition to the 12-gene molecular score. The EPclin risk score is on a scale from 1 to 6. A patient with an EPclin score of ≥3.32867 is categorised as having a high risk of distant recurrence, while an EPclin score of <3.32867 indicates low risk.[81]

In the EU, EndoPredict® can be performed by a local provider with the relevant instruments and kits or a partner laboratory. The turnaround time for test results varies depending on the laboratory, with a turnaround time of two days reported from receipt of sample to results from a local laboratory.[45] Outside of the EU, the centralised laboratory in the US reports a seven day turnaround time from receipt of sample and a regional laboratory in Australia reports four to five days.[79, 82] Further details on EndoPredict® are presented in Table 2.1.

2.4.7 Prosigna®

Prosigna®, manufactured by Veracyte, is a CE-marked in vitro diagnostic assay based on the Prediction Analysis of Microarray 50 (PAM50) gene signature. Prosigna® provides information on the breast cancer subtype and is intended to predict the risk of distant recurrence, assuming five years of endocrine therapy. The test is intended for use in post-menopausal women with either HR+, LN- stage I or II breast cancer or with HR+, LN+ (1 to 3 positive nodes) stage II or IIIa breast cancer.

Prosigna® analyses the activity of 50 genes which can be used for subtype classification of breast cancer. An additional 22 genes are analysed concordantly, including:

- eight housekeeping genes used for signal normalisation

- six positive controls
- eight negative controls.

Prosigna® is based on direct messenger RNA counting using fluorescent probes and an nCounter® Digital Analyser. The gene signature visualised by nCounter® platform is used to calculate a proliferation score, determined by evaluating multiple genes associated with the proliferation pathway.[83] The nCounter® Dx Analysis System uses a novel digital barcoding chemistry to deliver high precision multiplexed assays.[84, 85] The Prosigna® Breast Cancer Prognostic Gene Signature Assay is supported using this system.

The risk of distant recurrence is calculated based on the PAM50 gene signature, breast cancer subtype, tumour size, nodal status and proliferation score. The test provides a risk of recurrence (ROR) score, ranging on an integer scale from 0 to 100, which is taken into consideration along with nodal status to classify risk categories. In LN- patients, the risk categories are defined as follows:

- ROR of 0-40 indicates low risk of distant recurrence within 10 years
- ROR of 41-60 indicates intermediate risk of distant recurrence within 10 years
- ROR of 61-100 indicates high risk of distant recurrence within 10 years.[44]

In LN+ patients, the risk categories are defined as follows:

- ROR of 0-15 indicates low risk of distant recurrence within 10 years
- ROR of 16-40 indicates intermediate risk of distant recurrence within 10 years
- ROR of 41-100 indicates high risk of distant recurrence within 10 years.

Prosigna® is performed in local laboratories that have the relevant technologies (that is, NanoString's nCounter® platform), assays (Prosigna® Breast Cancer Prognostic Gene Signature Assay including the RNA isolation kit manufactured by Roche or any other RNA isolation kit that has been validated) and reagents (including reference sample, codeset, prepack and cartridges). The turnaround time for Prosigna®, from specimen receipt to communication of result, is three days according to the Prosigna® instructions for use.[86] Further details on Prosigna® are presented in Table 2.1.

Table 2.1 Summary of included GEP tests

	Oncotype DX®	MammaPrint®	EndoPredict®	Prosigna®
Manufacturer	Genomic Health, Inc. (a wholly owned subsidiary of Exact Sciences Corporation)	Agendia	Myriad International GmbH	Veracyte
CE mark accreditation	CE marked	CE-marked	CE-marked	CE-marked
Sample type	FFPE	FFPE	FFPE	FFPE
Test method	RT-PCR	Microarray-based	RT-PCR	nCounter® Dx Analysis System
Description (genes)	21-gene assay (16 cancer-related, 5 reference)	70-gene assay (70 cancer-related, 465 reference)	12-gene assay (8 cancer-related, 4 reference)	50-gene assay (50 cancer-related, 22 reference)
Indication(s)	Stage I-IIIa, HR+, HER2-, LN-, micrometastases or LN+ for prognostic and predictive use	Stage I-II (and operable stage III), ER+ or ER-, HER2-, LN- or LN+ for prognostic and predictive use	Stage I-II, ER+, HER2−, LN- or LN+ for prognostic and predictive use	Stage I-II, LN- and stage II-IIIa, LN+ (both in HR+, post-menopausal women only) for prognostic use
Clinicopathological factors considered	Not specified	Not specified	Tumour size, LN status	Breast cancer subtype, tumour size, LN status, proliferation score
Result measurement	Recurrence Score® result (RS): 0 to 100	MammaPrint® Index (MPI): -1 to 1	12-gene molecular score: 0 to 15 EPclin score: 1 to 6	Risk of Recurrence (ROR): 0 to 100
Categories for risk measurement	Age >50 years Low risk: 0 to 25 High risk: 26 to 100 Age ≤50 years and LN- Low risk: 0 to 15 Intermediate risk: 16 to 20 and 21 to 25 High risk: 26 to 100 Age ≤50 years and LN+ Intermediate risk: 0 to 13 and 14 to 25 High risk: 26 to 100	Ultra-low risk: 0.355 to 1 Low risk: 0 to 0.354 High risk: −1 to 0	EPclin risk score Low risk: < 3.32867 High risk: ≥ 3.32867	LN- patients Low risk: 0 to 40 Intermediate risk: 41 to 60 High risk: 61 to 100 LN+ patients Low risk: 0 to 15 Intermediate risk: 16 to 40 High risk: 41 to 100
Testing location	Test service only (US)	Test service only (US)	Local laboratory or test service (US; only for orders outside of the EU)	Local laboratory
List price (excluding VAT)	€3,180	€3,042	€1,975	€1,934

Key: CE – Conformité Européenne; ER – oestrogen receptor; FFPE – formalin-fixed paraffin-embedded; GEP – gene expression profiling; HER2 – human epidermal factor receptor 2; HR – hormone receptor; LN – lymph mode; RT-PCR – reverse transcription polymerase chain reaction.®

2.5 International practice

This section provides an overview of international practice in relation to the use of GEP tests. In particular, it describes the findings of relevant HTAs and the recommendations of international clinical guidelines.

2.5.1 Health technology assessments

Nine relevant HTAs, published between 2016 and 2022, were identified.[29, 31, 44, 45, 87-92] Two of these HTAs were from Canada,[29, 31] two were from the US,[88, 89] with one each from Australia,[90, 93] France,[92] Sweden,[94] and the UK,[44, 45] and the final HTA was a pan-European rapid assessment conducted by the European Network for Health Technology Assessment (EUnetHTA).[87]

Four were full HTAs, explicitly linked to jurisdiction-specific decision-making processes, which assessed the clinical- and cost-effectiveness associated with the use of GEP tests to support treatment decisions in patients with HR+, HER2- early stage breast cancer.[29, 31, 44, 45, 90] Four HTAs mainly comprised (systematic and non-systematic) reviews assessing the benefit associated with the use of GEP tests to assist treatment decisions in patients with HR+, HER2- early stage breast cancer.[87-89, 91, 92] The other HTA was mainly based on the findings of a health economic evaluation.[94] The findings of these HTAs are summarised in Table 2.2.

Six other HTAs were identified.[61, 91, 95-99] These are excluded from the below summary given that they were conducted prior to the publication of pivotal trials (such as the updated TAILORx and MINDACT trials),[61, 66, 95, 96, 99-101] were subsequently updated by a more recent HTA,[97, 98] or did not provide clear conclusions based on the HTA's findings.[91]

Prognostic ability of GEP tests

Four of the nine HTAs reported findings in relation to the ability of GEP tests to accurately predict the risk of disease recurrence.[29, 31, 44, 45, 90] The other five HTAs did not report conclusions regarding the prognostic ability of the included GEP tests.[87-89, 92, 94]

One HTA from the UK, undertaken by the National Institute for Health Research and commissioned by the National Institute for Health and Care Excellence (NICE), assessed the prognostic ability of four GEP tests.[44, 45] The HTA found that, compared with commonly used clinicopathological factors, EndoPredict®, MammaPrint®, Oncotype DX® and Prosigna® provided additional prognostic information in patients with HR+, HER2- tumours with either LN- or LN+ disease. This finding was based on unadjusted analyses which demonstrated significant

prognostic accuracy for the risk of recurrence. However, they found these results were more varied in LN+ patients than in LN- patients.

Both HTAs from Canada assessed the prognostic ability of GEP tests.[29, 31] The HTA by the Institute of Health Economics Alberta, published in 2019, found that the clinical evidence generally supported the additional prognostic ability of Oncotype DX® and Prosigna®.[31] For both tests, the evidence of additional prognostic ability was most pronounced for those classed at low and high risk of recurrence, but uncertainty remained in terms of those at intermediate risk. This HTA also found that Prosigna® more accurately predicted freedom from 10-year distant recurrence than Oncotype DX® in LN-, postmenopausal patients receiving endocrine therapy only. The Ontario HTA, published in 2020, reported that in LN- patients EndoPredict®, MammaPrint®, Oncotype DX® and Prosigna® were likely prognostic for freedom from distant recurrence, disease-free survival and overall survival, and that these GEP tests may be prognostic in LN+ patients for freedom from distant recurrence, disease-free survival and overall survival.[29]

The HTA from Australia, published by the Medical Services Advisory Committee (MSAC) in 2021, also assessed the prognostic ability of GEP tests.[90] Following the submission of HTA applications by the respective manufacturers, the MSAC found that it is likely MammaPrint®, EndoPredict® and Prosigna® have some modest prognostic value, in terms of estimating the risk of cancer recurrence when used alongside other available information (for example, commonly used clinicopathological factors). However, the MSAC concluded that overall there was insufficient evidence of added prognostic value beyond that of other available information.

Predictive ability of GEP tests

All nine HTAs reported findings in relation to the predictive ability of GEP tests to inform decisions regarding the use of adjuvant chemotherapy.[29, 31, 44, 45, 87-91]

The UK HTA found that there was limited evidence of the predictive ability for Oncotype DX® and MammaPrint®, citing uncertainties in the evidence base regarding whether or not the use of these tests was associated with predicted benefit of chemotherapy.[44, 45] Evidence of predictive ability was not identified for EndoPredict® or Prosigna®. Following the HTA, NICE conditionally recommended EndoPredict®, Oncotype DX® and Prosigna® for reimbursement in the National Health Service (NHS) to guide adjuvant chemotherapy decisions in ER+ HER2- LN- patients. This conditional recommendation included the collection of data, using the National Cancer Registration and Analysis Service, on the use of the tests within the NHS to address remaining uncertainty on the clinical impact (that is, chemotherapy use, recurrence and survival outcomes). MammaPrint® was not recommended

because it was not considered to be cost-effective.[44, 45] No recommendations were reported regarding the use of these GEP tests in ER+ HER2- LN+ patients.

The EUnetHTA HTA, published in 2018, concluded that there was insufficient evidence demonstrating the clinical utility of withholding adjuvant chemotherapy based on MammaPrint® testing.[87] This finding was based on the evidence provided by one randomised controlled trial (RCT), the Microarray in Node-Negative and 1 to 3 Positive Lymph Node Disease May Avoid Chemotherapy (MINDACT) trial,[100] only. No other GEP tests were included in the EUnetHTA review.

The HTA published by the Washington State Health Authority in 2018 found that women with early-stage invasive breast cancer, considered to be at high clinical risk of recurrence, may forego adjuvant chemotherapy on the basis of a low MammaPrint® genomic risk score.[88] This finding was based only on the outcomes of the MINDACT trial. The HTA also found that the clinical evidence supported the use of Oncotype DX® to inform decisions regarding the use of adjuvant chemotherapy in women with early-stage invasive breast cancer, particularly in identifying women at low risk who would not benefit from adjuvant chemotherapy. The HTA's conclusions did not specifically address the impact of lymph node involvement in the HR+, HER- population.

The HTA published by the Oregon Health Authority in 2018 recommended the use of Oncotype DX® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- and LN+ patients.[89] The HTA concluded that Oncotype DX® had the largest and best-established evidence base. The HTA also recommended EndoPredict® and Prosigna® for guiding adjuvant chemotherapy decisions in HR+, HER2-, LN- patients, but not in LN+ patients.[89] MammaPrint® was recommended for guiding adjuvant chemotherapy decisions only in patients who are HR+, HER2-, LN- that have been categorised as at high clinical risk.[89] The HTA classed the recommendations for EndoPredict®, Prosigna® and MammaPrint® as 'weak' on the basis that additional studies were necessary to better establish clinical utility and predictive value.

The HTA by the Institute of Health Economics Alberta (published in 2019) also reported that the evidence base supported the use of Oncotype DX® in LN- patients to inform chemotherapy decisions.[31] However, due to conflicting findings in the clinical evidence it reported that no conclusions could be drawn regarding the predictive ability of Oncotype DX® in LN+ patients. Similarly, due to a lack of published evidence, no clear findings were presented in the HTA regarding the predictive ability of Prosigna® in LN- or LN+ patients. However, the Ontario HTA reported that MammaPrint® and Oncotype DX® may be predictive of chemotherapy benefit in LN- and LN+ patients, with weaker evidence available for LN+ patients.[29]

The HTA from France, published by the Haute Autorité de Santé in 2019, reported that there was insufficient evidence of clinical utility to recommend routine use of GEP tests, while also noting that GEP tests are not intended to replace standard clinicopathlogical criteria.[92] However, the Haute Autorité de Santé recommended conditional funding of EndoPredict®, MammaPrint®, Oncotype DX® and Prosigna® with a view to collecting comparative prospective data in LN- patients who are at intermediate risk of recurrence (based on clinicopathological factors) but have no major clinicopathological discordance (which was not clearly defined) and uncertain predicted benefit from chemotherapy. The Haute Autorité de Santé reported that these data could be combined with updated results from the RxPONDER, OPTIMA and WSG ADAPT trials.[102-104] The Haute Autorité de Santé also recommended the validation of a risk prediction model in the French clinical setting in order to optimise decision-making regarding the use of adjuvant chemotherapy in circumstances where there is uncertainty arising from discordant clinicopathological factors.

The Australian MSAC found that there was insufficient evidence supporting the use of MammaPrint®, EndoPredict® and Prosigna® for determining which patients could safely avoid chemotherapy.[90] Accordingly, MSAC did not support public funding of these GEP tests, noting that none of them had been shown to be safe, effective or cost-effective. Prior to this, MSAC recommended against public reimbursement of Oncotype DX®, based on a HTA submitted by the manufacturer in 2019.[93] The HTA incorporated the 2018 TAILORx trial results,[66, 68, 72] and the analysis included intermediate risk patients using the revised RS thresholds.[105] MSAC reported that the TAILORx trial had not demonstrated the ability of Oncotype DX® to identify patients who could safely avoid the use of adjuvant chemotherapy, nor had the analysis demonstrated that Oncotype DX® could identify intermediate risk patients who would benefit from adjuvant chemotherapy.

Finally, the HTA published by the Swedish Medical Technologies Product (MTP) Council in 2021 reported that Oncotye DX® and Prosigna® could be used to inform adjuvant chemotherapy decisions in patients in which there is uncertainty regarding the potential benefit of adjuvant chemotherapy.[94] However, the Swedish MTP Council reported that this recommendation, based primarily on a health economic evaluation, was subject to uncertainty. In particular, there was a lack of evidence demonstrating direct patient benefit from Prosigna® testing, and the Oncotype DX® study population lacked generalisability to the Swedish patient population.

Table 2.2. Summary of findings from previously published HTAs

Agency (year)	EndoPredict®	MammaPrint®	Oncotype DX®	Prosigna®
EUnetHTA (2018)[87]	Not assessed	Insufficient evidence to demonstrate clinical utility improved patient outcomes	Not assessed	Not assessed
HAS (2019)[92]	Evidence of clinical utility was insufficient to recommend routine use of GEP tests. However, recommended that temporary and conditional research and innovation program funding be extended (no date specified) for future review of evidence.			
Institute of Health Economics Alberta (2019)[31]	Not assessed	Not assessed	Evidence supports use in HR+, HER2-, LN- population, but uncertainty remains in LN+	Evidence supports use in HR+, HER2-, LN- population, but uncertainty remains in LN+
MSAC (2019, 2022)[90, 93]	Insufficient evidence of predictive ability	Insufficient evidence of predictive ability	Insufficient evidence of predictive ability	Insufficient evidence of predictive ability
Swedish MTP Council (2021)[94]	Not assessed	Not assessed	Can be used to inform adjuvant chemotherapy decisions in patients where there is uncertainty	Can be used to inform adjuvant chemotherapy decisions in patients where there is uncertainty
NICE (2018)[44, 45]	Can be used in ER+, HER2-, LN- *	Not recommended	Can be used in ER+, HER2-, LN-	Can be used in ER+, HER2-, LN-
Ontario Health (Quality) (2020)[29]	May be predictive in HR+, HER2-, LN- and LN+ populations, but evidence is weak in LN+ population. No evidence assessing the predictive benefit of EndoPredict® or Prosigna® was identified.			
Oregon Health Authority (2018)[89]	Evidence supports use in HR+, HER2-, LN- but not LN+	Not assessed	Evidence supports use in HR+, HER2-, LN- and LN+ populations	Evidence supports use in HR+, HER2-, LN- but not LN+
Washington State Health Authority (2018)[88]	Not assessed	Women at high clinical risk that receive a low MammaPrint® risk score may forego chemotherapy	Evidence supports use (non-specific)	Not assessed

Key: ER – oestrogen receptor; EUnetHTA – European Network for Health Technology Assessment; GEP – gene expression profiling; HAS – Haute Authorité de Santé; HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; HTA – health technology assessment; LN – lymph node; MSAC – Medical Services Advisory Committee; MTP – Medical Technologies Product; NICE – National Institute for Health and Care Excellence.

* Conditional recommendation subject to the following criteria:

- the patient has an intermediate risk of recurrence using a validated tool such as PREDICT or the Nottingham Prognostic Index
- information provided by the test would help the patient choose, with their clinician, whether or not to have adjuvant chemotherapy
- the respective manufacturers provide the tests to the NHS at the discounted prices agreed in the proposals
- both clinicians and manufacturers make timely, complete and linkable record-level test data available to the National Cancer Registration and Analysis System.

2.5.2 International clinical guidelines

Six relevant international (or regional) clinical guidelines, published between 2017 and 2022, were identified.[18, 19, 101, 106-110] These guidelines were developed by the:

- American Society of Clinical Oncology (ASCO)
- European Commission Initiative on Breast Cancer (ECIBC)
- European Group on Tumor Markers (EGTM)
- European Society of Medical Oncology (ESMO)
- National Comprehensive Cancer Network (NCCN)
- St. Gallen International Expert Consensus (St. Gallen).

These guidelines were generally informed by (systematic and non-systematic) reviews of existing literature and finalised by expert consensus. A diverse range of approaches was adopted with regards to rating the quality of the evidence underpinning guideline recommendations and rating the strength of each recommendation. Given the variation in approaches, these ratings are not reported in the below summary. The guidelines often did not explicitly specify in which patients the GEP tests should be used (that is, whether all patients should be offered GEP testing or only those for whom there is uncertainty in clinical decision-making). The clinical subgroup is reported when this information was provided in the guideline. The guidelines' recommendations are summarised in Table 2.3.

A guideline by the American Joint Committee on Cancer (AJCC) regarding incorporation of GEP tests in the AJCC Breast TNM staging system was excluded from the summary given its focus on breast cancer stage classification.[11, 111]

Prognostic ability of GEP tests

Two guidelines recommended the use of EndoPredict®, MammaPrint®, Oncotype DX® and Prosigna®, in combination with established clinicopathological factors, for prognostic purposes in patients with HR+, HER2- LN- and LN+ early-stage invasive breast cancer.[19, 106]

The ESMO guidelines, which did not differentiate between the alternative GEP tests, recommended the use of validated GEP tests, in combination with other clinicopathological factors, to gain additional prognostic information.[18]

Predictive ability of GEP tests

Four of the six guidelines recommended the use of Oncotype DX® for guiding adjuvant chemotherapy decisions in patients with HR+, HER2- LN- patients early-stage breast cancer.[19, 101, 106-108, 112] Two of these four guidelines (by EGTM and NCCN) also recommended Oncotype DX® for guiding adjuvant chemotherapy

decisions in patients with HR+, HER2- LN+ patients.[19, 106] However, the NCCN guideline reported that the evidence underpinning this recommendation in LN+ premenopausal women was of lower quality.[19] Finally, the ASCO guideline noted that Oncotype DX® could be used in postmenopausal LN+ patients, but not in premenopausal LN+ patients.[112]

The EGTM guideline recommended that MammaPrint® could be used in patients who are HR+, HER2-, LN- or LN+.[106] In particular, the EGTM guideline highlighted that patients at high risk based on clinicopathological factors and low-risk MammaPrint® score could be candidates for safely avoiding adjuvant chemotherapy.[106] The ECIBC guideline recommended MammaPrint® for guiding chemotherapy decisions in women with HR+, HER2-, LN- or LN+ early-stage breast cancer that are at high clinical risk.[107] The ASCO guideline recommended the use of MammaPrint® in HR+, HER2-, LN- or LN+ patients who are aged 50 years or over and at high clinical risk only.[108, 109, 112] Due to a lack of meaningful benefit and the associated large cost, the ECIBC guideline recommended against the use of MammaPrint® in HR+, HER2-, LN- or LN+ patients who are at low clinical risk.[107] The NCCN guideline published in 2022 reported that the predictive ability of MammaPrint® was currently not known due to a lack of high-quality evidence.[19]

Two of the six guidelines, EGTM and ASCO, recommended both EndoPredict® and Prosigna®, in combination with established clinicopathological factors, for guiding adjuvant chemotherapy decisions.[106, 108, 112] The EGTM guideline recommended the use of EndoPredict® and Prosigna® in HR+, HER2-, LN- and LN+ patients.[106] The ASCO guideline noted that EndoPredict® could be used in postmenopausal LN- and LN+ patients, while Prosigna® could be used in postmenopausal LN- patients but that the evidence was inconclusive for postmenopausal LN+ patients.[108] Both guidelines recommended against the use of EndoPredict® and Prosigna® in premenopausal patients, irrespective of LN involvement.[106, 108, 112] The NCCN guideline (published in 2022) reported that the predictive ability of EndoPredict® and Prosigna® was currently not known due to a lack of high-quality evidence.[19]

The ESMO guidelines, which did not differentiate between the alternative GEP tests, recommended using validated GEP tests, combined with other clinicopathological factors, to gain additional predictive information.[18] Similarly, the St. Gallen expert consensus panel recommended that GEP tests should be used in the vast majority of instances to inform treatment decisions for women with ER+, HER2- breast cancers and limited lymph node involvement irrespective of tumour grade or menopausal status.[110] The panel's recommendations did not differentiate between the alternative GEP tests.

Table 2.3. Recommendations from international clinical guidelines*

Group (year)	EndoPredict®	MammaPrint®	Oncotype DX®	Prosigna®
ASCO (2017, 2019, 2022)[108, 109, 112]	May be used in postmenopausal (LN- or LN+) patients. Should not be used in premenopausal (LN- or LN+) patients.	May be used in patients aged ≥50 years (LN- or LN+) that have high clinical risk.** Should not be used in patients aged ≤50 years (LN- or LN+) that have high clinical risk.** Should not be used in patients who have low clinical risk.**	Recommended in LN- patients: • patients with RS ≥26 should be offered CET • patients aged ≤50 years with a RS of 16-25 can be offered CET. May be used in postmenopausal patients who are LN+: • patients with a RS ≥26 should be offered CET. Should not be offered to premenopausal patients who are LN+.	May be used in postmenopausal, LN- patients. Evidence is inconclusive for postmenopausal, LN+ patients. Should not be used in premenopausal (LN- or LN+) patients.
ECIBC (2021)[107]	No recommendation	Recommended in LN- patients at high clinical risk and LN+ patients at high clinical risk, but not in those at low clinical risk	Recommended in LN-	No recommendation
EGTM (2017)[106]	Recommended in LN- and LN+	Recommended in LN- and LN+	Recommended in LN- and LN+	Recommended in LN- and LN+
ESMO (2019)[18]	Recommends use of validated GEP tests (non-specific) in combination with other clinicopathological factors			
NCCN (2022)[19]	Not currently known due to lack of high-quality evidence	Not currently known due to lack of high-quality evidence	Recommended in LN- and LN+	Not currently known due to lack of high-quality evidence
St. Gallen (2021)[110]	Use of GEP tests (non-specific) should be considered in the vast majority of cases when chemotherapy is being considered for people with ER+, HER2- breast cancers with limited LN involvement (1-3 nodes), irrespective of tumour grade or menopausal status			

Key: ASCO – American Society of Clinical Oncology; CET – chemoendocrine therapy; ECIBC – European Commission Initiative on Breast Cancer; EGTM – European Group on Tumor Markers; ER – oestrogen receptor; ESMO – European Society of Medical Oncology; ET – endocrine therapy; LN – lymph node; NCCN – National Comprehensive Cancer Network; RS – recurrence score; St. Gallen – St. Gallen International Expert Consensus.

* All recommendations are for patients with early-stage HR+, HER2- breast cancer. LN+ refers to patients with 1-3 positive nodes only.

** ASCO 2022 reported clinical risk based on a modified version of Adjuvant! Online.[112]

2.6 Discussion

GEP tests in breast cancer are intended to provide information on disease prognosis and to predict whether or not a patient is likely to benefit from adjuvant chemotherapy. GEP tests are intended to supplement clinical judgement in cases where there is uncertainty in the decision-making process in light of clinical, pathological, and or patient-related factors. In Ireland, the Oncotype DX® test is the only GEP test reimbursed through the publicly funded healthcare system (that is, the HSE). The Oncotype DX® test is available for use in patients with HR+, HER2-, LN- and LN+ early-stage breast cancer. The four GEP tests assessed within this rapid HTA (that is, Oncotype DX®, Mammaprint®, EndoPredict® and Prosigna®) differ according to the clinical indication for which they are approved or marketed, analytical technique, genes analysed, risk scoring approach and risk score interpretation.

The regulatory status in the EU of the four GEP tests is worthy of noting. Conformité Européenne accreditation (CE marking) is required for all in-vitro diagnostic devices sold in Europe. This indicates that the in-vitro diagnostic device complies with the European In-Vitro Diagnostic Devices Directive (IVDD 98/79/EC) and that it can be placed on the single market.[113] Each of the four GEP tests assessed within this rapid HTA are self-declared and CE marked under IVDD 98/79/EC. This Directive does not require assessment or appraisal of clinical evidence by a notified body before devices are placed on the market.

In May 2022, the new In Vitro Diagnostic Regulation (IVDR 2017/746) came into effect. The IVDR introduced four classes of IVDs which take into account their risk and intended purpose (A, B, C and D). Clinical evidence and post-market surveillance requirements vary according to the class of the IVD, however notified body certification will be required for class B, C and D devices. Transitional provisions in the IVDR allow devices self-declared under IVDD 98/79/EC to continue to be placed on the market and made available according to various deadlines that are dependent on the class of the device. In order to continue to place these self-declared devices on the market, manufacturers must seek certification from a notified body before the expiration of the transitional provisions set out in Regulation 2022/112.[114]

The turnaround time (that is, the time taken between requesting and receiving test results), which could potentially delay patient access to appropriate therapy and contribute to patient harms (in terms of anxiety and stress), is a relevant consideration in the use of alternative GEP tests. The turnaround time may be impacted by the extent of sample preparation work required within local laboratories prior to submitting the sample for GEP testing. As noted in sections 2.4.4 to 2.4.7, the turnaround time and costs of transportation can vary by test and testing

location. It may be reasonable to assume that the turnaround time and transportation costs could be higher for GEP tests that need to be sent to centralised laboratories outside of Ireland for analysis. However, it would also be challenging to accommodate GEP testing for early-stage breast cancer within current cancer laboratory workflows and resource constraints in Ireland. The impact of these logistical issues and the costs arising from sample handling, sample transportation and additional laboratory resourcing (for example, from labour and equipment) are not assessed in this rapid HTA. Any decision on alternative GEP test use should ensure the availability of a reliable test that can, in a timely manner, accurately identify breast cancer patients who can forego adjuvant chemotherapy, and support informed decision-making by the patient and the multidisciplinary clinical team. Barriers to the use of alternative GEP tests, arising from logistical and resourcing constraints and the potential impact of these issues on patient outcomes are important considerations for decision-making.

As reported in section 2.5, international guidance and practice regarding the use of GEP tests varies widely, particularly in terms of the use of GEP tests in LN+ populations. It should be noted, however, that the summaries of international HTAs and clinical guidelines detailed within this chapter were not underpinned by a systematic search. Therefore, all relevant HTAs and international guidelines may not have been captured. In addition, a number of the included guidance documents were available in non-English languages only and therefore interpretation may be subject to translation error. It is also worth noting that international practice evolves over time, often coinciding with the publication of important trials, and that this guidance is therefore likely to change. In Ireland, available guidance which refers to the use of Oncotype DX® includes the 2015 National Clinical Effectiveness Committee National Clinical Guideline and the eligibility criteria outlined by NCCP Systemic Anti-Cancer Therapy Breast Clinical Advisory Group.[5, 10] Given the volume of international evidence and variation in international practice, further best practice guidance (in terms of tissue sampling and the use and interpretation of GEP tests) may be warranted in the Irish context.

3 Epidemiology

Key points

- The target population of interest for the current rapid HTA is people with hormone receptor-positive (HR+) human epidermal growth factor receptor 2-negative (HER2-) early-stage (stage I-IIIa) invasive breast cancer. This grouping represents the majority of breast cancer cases in Ireland.
- For HR+, HER2-, early-stage breast cancer, the risk of recurrence is highest in the second year post-diagnosis. Women may receive adjuvant chemotherapy to reduce this risk.
- Breast cancer prognosis and response to treatment are influenced by a variety of factors, including those related to the patient, for example, age, and those related to the tumour. Relevant tumour factors include tumour size, tumour grade, histologic subtype, lymphovascular invasion of tumour cells and axillary lymph node status, and the presence of hormone and HER2 receptors.
- The most common breast cancer classification system is the TNM staging system from the American Joint Committee on Cancer, which assigns patients to one of five breast cancer stages (0, I, II, III, or IV) based on the size of the tumour (T) and the degree of local infiltration, lymph node involvement (N) and the presence of metastasis (M) beyond the breast and regional lymph nodes.
- Based on the five-year average from 2015 to 2019, there were 1,641 new cases of HR+, HER2- stage I-II breast cancer and 1,806 cases of HR+, HER2- stage I-IIIa diagnosed in Ireland. Cases were most frequently diagnosed in those aged 50 to 64 years, with 44% of stage I-II and 43% of stage I-IIIa of cases diagnosed in this age group between 2010 and 2019.
- Of all HR+, HER2- early-stage invasive breast cancers diagnosed between 2006 and 2015, 42% (n=6,852) were diagnosed at stage I, 50% (n=8,225) at stage II and 8% (n=1,316) at stage IIIa. Considering lymph node status at diagnosis, 1% of stage I cases were LN+, 54% of stage II cases were LN+ and all stage IIIa cases were LN+.
- Of women with stage I-IIIa breast cancer diagnosed between 2010 and 2015, within the first year of diagnosis, 92% (n=9,836) underwent surgery, 76% (n=8,106) received radiotherapy, 70% (n=7,497) received hormone therapy,

and 38% (n=4,124) received chemotherapy. Of those who received chemotherapy, 99% (n=4,086) also underwent surgery.

- Five-year net survival was estimated at 98% for both stage I-II and stage I-IIIa HR+, HER2- breast cancer, in 2018. Nine-year net survival was estimated at 97% for both stage I-II and stage I-IIIa HR+, HER2- breast cancer.
- The five-year rolling averages of deaths between 2015 and 2019 in people with HR+, HER2- stage I-II and stage I-IIIa deaths in Ireland were 178 and 216, respectively. The total number of deaths per five-year age band generally increased with increasing age. The age-standardised mortality rate for all breast cancer cases has decreased over time, as reported by the National Cancer Registry Ireland.

3.1 Introduction

This chapter describes the target population for this rapid HTA of gene expression profiling (GEP) tests in early stage invasive breast cancer (section 3.2), summarises the natural history of breast cancer (section 3.3), and describes the associated burden of disease in Ireland (section 3.4) based on data collected and made available by the National Cancer Registry Ireland (NCRI).

3.2 Target population

Breast cancer is the most common invasive cancer diagnosed in women in Ireland, accounting for 30% of all invasive cancers and affecting one in eight women over their lifetime.[115, 116] GEP tests are indicated for both men and women, however breast cancer in men is rare.[117, 118] Therefore, while the scope of this rapid HTA includes men, its focus is mainly on the use of GEP tests in women.

The presence of hormone receptors and or human epidermal growth factor receptor 2 (HER2) can affect the treatment course and potentially the outcome of breast cancer. Hormone receptor positive (HR+) breast cancers are those with receptors for oestrogen and or progesterone, commonly termed oestrogen-receptor positive (ER+) and progesterone-receptor positive (PR+).[2] Breast cancers that do not have receptors for the protein HER2 are referred to as a HER2- breast cancer. HR+, HER2- breast cancers represent approximately 70% of all breast cancers diagnosed in western countries, with an estimated 40% of these having spread to the lymph nodes by the time of diagnosis.[107, 119] The specific target population for the current rapid HTA is people with HR+, HER2- early-stage invasive breast cancer.

From 2003 to 2015, 79% of all breast cancers in women were recorded as being HR+, with 81% of these being in women aged 50 to 64 years.[2] Information on hormone and HER2 status of breast cancers diagnosed in Ireland has been routinely collected by the NCRI since 2006.[120] From 2006 to 2015, approximately 77% of breast cancer cases in women were recorded as being HER2-, with 42% of those cases being diagnosed in women aged 50-64 years.

The most common breast cancer classification system, described further in Section 3.3.3, is the TNM staging system from the American Joint Committee on Cancer (AJCC).[11] It assigns patients to one of five breast cancer stages (0, I, II, III, or IV) based on the size of the tumour (T) and the degree of local infiltration, lymph node involvement (N) and the presence of metastasis (M) beyond the breast and regional lymph nodes. Generally, early-stage breast cancer is defined as a TNM stage classification of I-II or I–IIIa. Therefore, the target population for this rapid HTA includes people with TNM stages I-IIIa. Approximately 80% of all breast cancer cases in women in Ireland are diagnosed at an early stage.[2, 3] The GEP tests

EndoPredict® and MammaPrint® are only indicated for TNM stage I-II. However, Oncotype DX® and Prosigna® are also indicated for TNM stage IIIa (see Section 2.4).

3.3 Natural history of disease

Breast cancer is a disease in which cells in the breast grow out of control, eventually forming a mass of tissue called a tumour. Sections 3.3.1 to 3.3.3 summarise the natural history of breast cancer disease in terms of its aetiology, pathology and prognosis.

3.3.1 Aetiology

The causes of breast cancer are complex and not completely understood, but known risk factors include demographic characteristics (that is, sex and age) and a range of inherited genetic, environmental and lifestyle factors, described below.

Sex and age

Breast cancer can occur in both men and women. However, despite international evidence that incidence in men is increasing,[121] women are at a substantially (approximately 100 times) greater risk of developing breast cancer.[18] Incidence is typically higher in older cohorts, especially in women aged 50 years or older following menopause,[2, 122] but there is also national and international evidence of increasing incidence in younger women.[2, 123] Breast cancer in men typically occurs in those over the age of 60, but can also affect younger men.[124, 125]

Genetics and family history

An inherited genetic predisposition in the BReast CAncer 1 (BRCA1) and BReast CAncer 2 (BRCA2) susceptibility genes is a major risk factor for breast cancer.[116, 126] It is estimated that BRCA1 and BRCA2 gene mutations are present in 80-90% of all hereditary breast cancers, which in turn typically account for 5-10% of breast cancer cases,[127] but are less common in sporadic breast cancers.[128] A 2017 cohort study (N=113,000 women) conducted in the UK found that, compared with women without any affected relatives, the risk of developing breast cancer was 1.75-times higher in women that had a first-degree relative (that is, a parent, sibling or child) with breast cancer, and 2.5-times higher in women that had two or more first-degree relatives with breast cancer.[129]

Reproductive history

The risk of developing breast cancer is related to reproductive history in terms of the age at first menstrual period (that is, menarche), age at final menstrual period (that

is, menopause), the number of times a women has given birth (that is, parity) and the age at first full-term pregnancy.[130]

Published evidence indicates that women who experience menarche at a younger age (for example, before 12 years) or menopause at an older age (for example, after 55 years) are at an elevated risk of breast cancer.[128] Large-scale case-control studies and meta-analyses have consistently shown that younger age at menopause reduces breast cancer risk,[122, 131-135] with each one-year increase in age corresponding to a 3-4% rise in risk.[122, 131, 136] The same study found that age at menarche had the greater impact on breast cancer risk.[122, 131] Although it was previously believed that the influence of age at menarche and menopause on breast cancer risk was due to the duration of exposure to cycling ovarian hormones, a 2012 meta-analysis found that there was no relationship between age at menarche and age at menopause.[122, 131]

Early age at pregnancy can have a protective effect against breast cancer,[137, 138] but large case–control studies and meta-analyses indicate that this protection is restricted to HR+ tumours.[122, 139-141] The number of births, spacing between births and breastfeeding are also associated with a reduction in the risk of developing breast cancer.[122, 128, 142, 143] However, nulliparity (that is, never having given birth to a live baby) and late age at first full-term pregnancy contribute to an increased risk of breast cancer.[137, 144, 145] It has been suggested that parity-induced protection is hormonally-driven, while the mechanisms of breastfeeding-induced protection are largely unknown, but are not limited to ER+ cancers.[122]

Environmental and lifestyle risk factors

Environmental risk factors include exposure to ionising radiation (such as from radiation therapy) and exogenous oestrogens.[18] Exogenous (that is, originating from outside the body) oestrogens, such as the use of hormone replacement therapy (HRT) and oral contraceptives, are associated with the risk of developing breast cancer.[116, 146] Lifestyle factors such as excessive alcohol consumption, smoking, physical inactivity and being overweight or obese post-menopause are all associated with an elevated risk of breast cancer.[116, 147]

3.3.2 Pathology

Symptoms

Most early-stage breast cancers are asymptomatic and discovered during screening by mammography.[148] Symptoms of invasive breast cancer can be variable and depend on stage, but in general can include:

- a breast lump or thickening

- alteration in size, shape or appearance of a breast
- dimpling, redness, pitting or other alteration in the skin
- change in nipple appearance or alteration in the skin surrounding the nipple (areola)
- abnormal nipple discharge (for example, blood-stained or clear fluid).[28]

The symptoms of advanced breast cancer, including locally advanced (stage III) and metastatic (stage IV) breast cancer, can differ from those of early-stage breast cancer and can also differ according to the site of metastasis and disease course.[149] General symptoms include fatigue, difficulty sleeping and or depression. However, symptoms can also include bone and skin complications (such as bone pain, hypercalcaemia and infection), symptoms of the central nervous system (such as severe headaches, confusion, seizures and speech impairment), complications of the gastrointestinal tract (such as nausea, vomiting and diarrhoea), pulmonary complications (such as dyspnoea, haemoptysis and cough) and lymphoedema (that is, lumps or swelling in the lymph node areas).[150] This can lead to mortality if left untreated.

Types of breast cancer

Breast cancer can be classified by the type of tissue or area of the breast in which the cancer originates (such as the ducts, lobules or the tissue between them). Two broad classifications, based on which cell origin is involved, can be applied:

- Carcinoma – arising from the epithelial component of the breast, consisting of the layer of cells that lines the lobules and terminal ducts responsible for producing milk during lactation (that is, breastfeeding).
- Sarcoma – arising from the stromal components of the breast, comprising the connective tissue of the breast (that is, the fibrous tissue in which the epithelial elements are located). Sarcoma is much rarer than carcinoma, comprising less than 1% of primary breast cancers, and is not the focus of this rapid HTA.[151]

Within the heterogeneous group of breast carcinomas there are various subtypes, according to pathological features and invasiveness relative to the primary tumour site. These subtypes have different prognoses and treatment implications. Common breast cancers can be categorised according to whether they are *in situ* (that is, non-invasive), invasive, or metastatic. Invasive cancers are those in which the tumour has spread outside of the breast lobules (glands that are involved in milk production) or ducts (tubes which are involved in transport of milk from lobules to the nipple) to the surrounding tissue, and, potentially, to the local lymph nodes (that is, small organs comprising groups of immune cells, lymphotytes, that filter lymph

fluid through the body's lymphovascular or lymphatic channels). Metastatic cancers are those which have spread to other parts of the body, for example, the bones, liver, lungs or brain.

Pathophysiology

Breast cancer begins with genetic changes in a single cell or a small group of cells in the epithelium (cells which form the lobules and ducts) or stroma (supporting tissue which regulates the epithelium) of the breast, which, in the absence of immune suppression of the abnormal growth, allow cells in the breast to reproduce uncontrollably, eventually forming a tumour. The cancer can then spread via the lymphatic system or the bloodstream, but typically spreads first to the lymph nodes in the axilla (that is, the underarm area) before spreading elsewhere in the body.[44] Spread via the bloodstream can lead to distant metastases in the bone or viscera that are incurable.

The presence or absence of axillary lymph node metastases is a key indicator of disease and prognosis, and is considered as part of decisions regarding the use of adjuvant therapy.[44, 152] These are caused by detachment of a single or small number of cells from the main tumour which circulates via the lymphatic system before growing in the axillary lymph nodes. Axillary metastases, for which there is an inferior prognosis, occur in approximately 41% of cases.[44] When metastases are present, axiliary dissection (that is, surgical removal) is indicated in order to prevent further spread of the disease.

3.3.3 Staging and prognosis

Breast cancer prognosis and treatment response are influenced by a variety of factors, including tumour size, tumour grade (that is, the degree to which cells appear abnormal compared to normal cells under microscopic examination), histologic subtype, lymphovascular invasion (that is, spreading of cancer cells via lymphatic channels) of tumour cells and axillary lymph node status, and HR status.[31] The most common breast cancer classification system is the TNM staging system from the AJCC.[11] This system assigns patients to one of five breast cancer stages (that is, 0, I, II, III, or IV) based on the size of the tumour (T) and the degree of local infiltration, lymph node involvement (N) and the presence of metastasis (M) beyond the breast and regional lymph nodes. The TNM stages are:

- stage 0 (non-invasive) – abnormal cells are present but have not spread to the surrounding tissue
- stage I (invasive) – cancer is present, but contained in the area where the first abnormal cells began to develop

- stage II (invasive) – the cancer is growing, but still contained in the breast or else growth has only extended to the nearby lymph nodes
- stage III (invasive) – the cancer has extended beyond the immediate region of the tumour and may have invaded nearby lymph nodes and muscles, but has not spread to distant organs (can be referred to as 'locally advanced')
- stage IV (metastatic) – the cancer has spread to other distant areas of the body such as the liver, lungs or bones.

Early-stage cancer is variably defined, but is generally categorised as stage I–II or I–IIIa. In these stages, cancer is present, with or without regional lymph node involvement, but without distant metastases. Lymph node involvement can be classified by the presence (lymph node-positive (LN+)) or absence (lymph node-negative (LN-)) of regional lymph nodes. LN+ disease is further classified by the number of positive lymph nodes:

- N1 – one to three positive lymph nodes with or without micrometastases
- N2 – four to nine positive lymph nodes
- N3 – more than nine positive lymph nodes.[31]

Breast cancer can be further classified based on the presence or absence of the three main receptors (that is, HER2, ER and PR) and an additional receptor for the Ki67 protein (that is, a marker of cell proliferation that correlates with how quickly the cancerous cells are growing). There are two main molecular subtypes among breast cancers that are HR+, HER2-:

- luminal A – cancer growth is typically slow and has the best prognosis of the molecular subtypes
- luminal B – cancer growth is slightly faster than that of luminal A and prognosis is inferior.[29, 31]

Early diagnosis of breast cancer is associated with a better prognosis and a higher survival rate. In general, good prognosis is associated with small tumour size, being LN-, younger age, HR+ status and a lack of expression of HER2. Selection and timely administration of appropriate therapeutic options (see Section 2.3.3) is also critical to survival rates.

Relative to patients with HER2+ and triple negative tumours (which are not within the scope of this rapid HTA), patients with early-stage HR+ HER2- breast cancer luminal A or B tumours have a good prognosis, and often have tumours that progress slowly. Following surgery, these women receive adjuvant endocrine therapy (considered standard practice) and may also receive adjuvant chemotherapy (that is, 'chemoendocrine therapy') to reduce the risk of recurrence. In women diagnosed with stage I-III invasive breast cancer, risk of recurrence is highest during the

second year post-diagnosis and first recurrence is associated with more than three positive lymph nodes, age under 40 years, and tumour size over 5cm.[153] In women with early-stage ER+ breast cancer who receive adjuvant endocrine therapy for five years, distant recurrence is estimated to occur in 13-34% of stage I and 19-41% in stage II breast cancer cases, depending on the lymph node status and grade of the original tumour.[154]

It has been demonstrated that women with HR+ breast cancer have better survival outcomes than those with HR- breast cancer,[155, 156] due in part to additional treatment options available for this subgroup.[120] Without the use of chemotherapy, it is estimated that 15% of women with HR+ breast cancer will develop a recurrence within 10 years if treated with adjuvant endocrine therapy alone.[7, 157] However, if all women with HR+ breast cancer were to receive chemotherapy most of these women could be considered to be over-treated; this is due to the relatively low risk of recurrence and the partial effectiveness of chemotherapy for this cohort.[107]

3.4 Morbidity and mortality

This section describes data on the morbidity and mortality associated with early stage breast cancer in Ireland. These data were provided by the NCRI. All new cancer cases, including *in situ* disease, in the population usually resident in Ireland have been registered with the NCRI since 1994.

Only data for HR+, HER2- early-stage (I-IIIa) invasive breast cancers are presented in this section, in keeping with the population of interest for this rapid HTA. Also data are restricted to females only, given the rarity of breast cancer in males. In general, these data are presented according to the stage at diagnosis based on two groupings, stage I-II and stage I-IIIa, as two of the GEP tests assessed in this rapid HTA (EndoPredict® and MammaPrint®) are indicated for TNM stage I-II only and the other two (Oncotype DX® and Prosigna®) are indicated for TNM stage I-IIIa.

The time periods for the requested data were informed by correspondence with the NCRI regarding the level of completeness of these data. The number of deaths is based on death certificates where the official cause of death was recorded as breast cancer (C50).

3.4.1 Incidence

Absolute number of new cases of breast cancer

Between 2014 and 2019, the five-year rolling average of new cases ranged between 1,612 and 1,689 for stage I-II cases and between 1,761 and 1,852 for stage I-IIIa cases. Based on the five-year average from 2015 to 2019, there were 1,641 new

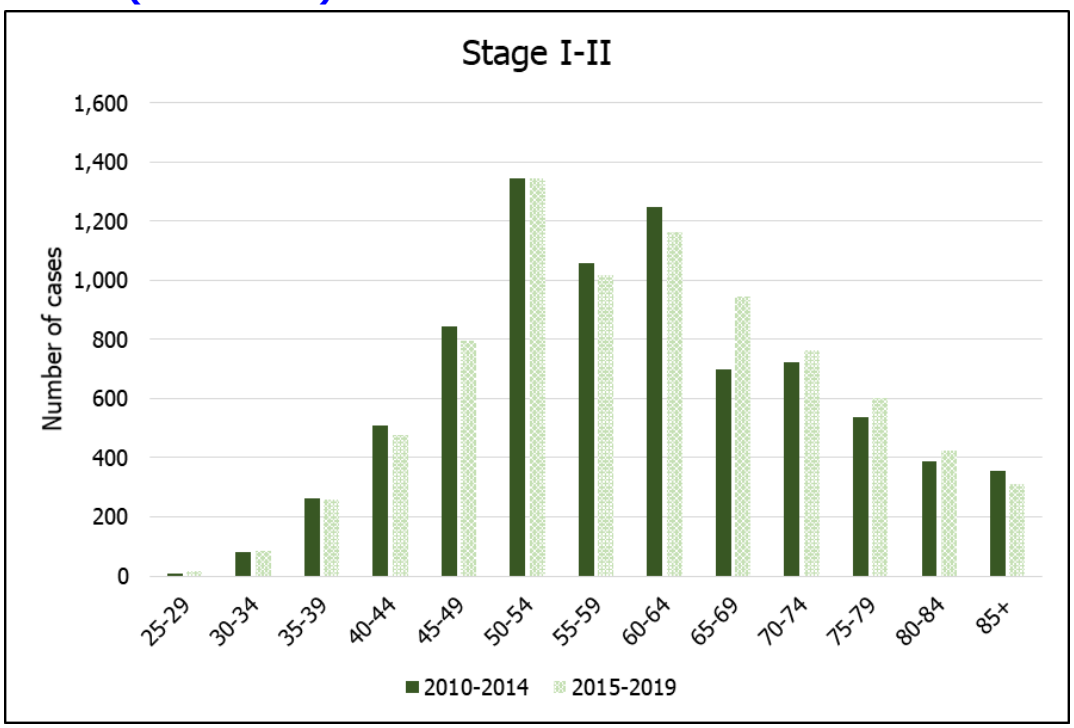
cases of HR+, HER2- stage I-II breast cancer and 1,806 cases of HR+, HER2- stage I-IIIa diagnosed in Ireland.

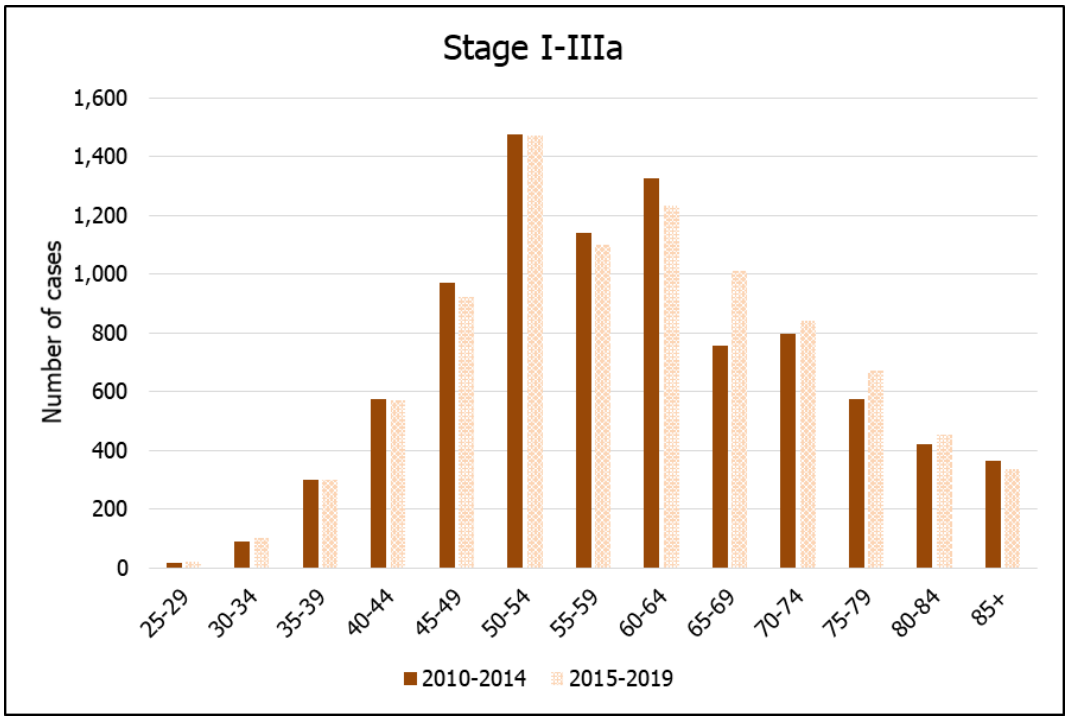
As noted in sections 2.2 and 3.3.3, the presence or absence of the spread of breast cancer to the LNs can impact on prognosis. Overall, 11,277 (69%) of the stage I-II cases from 2010 to 2019 were LN-, while 4,420 (27%) were LN+ (see Figure 3.1 for trends over time). All reported stage IIIa cases in this time period were LN+.

Age at diagnosis

For both stages I-II and I-IIIa, cases of HR+, HER2- breast cancer were most commonly diagnosed in women aged 50 to 64 years. Overall, between 2010 and 2019, 44% of women diagnosed at stage I-II and 43% of women diagnosed at stage I-IIIa were aged 50 to 64 years at diagnosis. The age breakdown of cases by five-year period is presented in Figure 3.1.

Figure 3.1 Incidence of HR+, HER2- early-stage breast cancer, by age (2010-2019)*





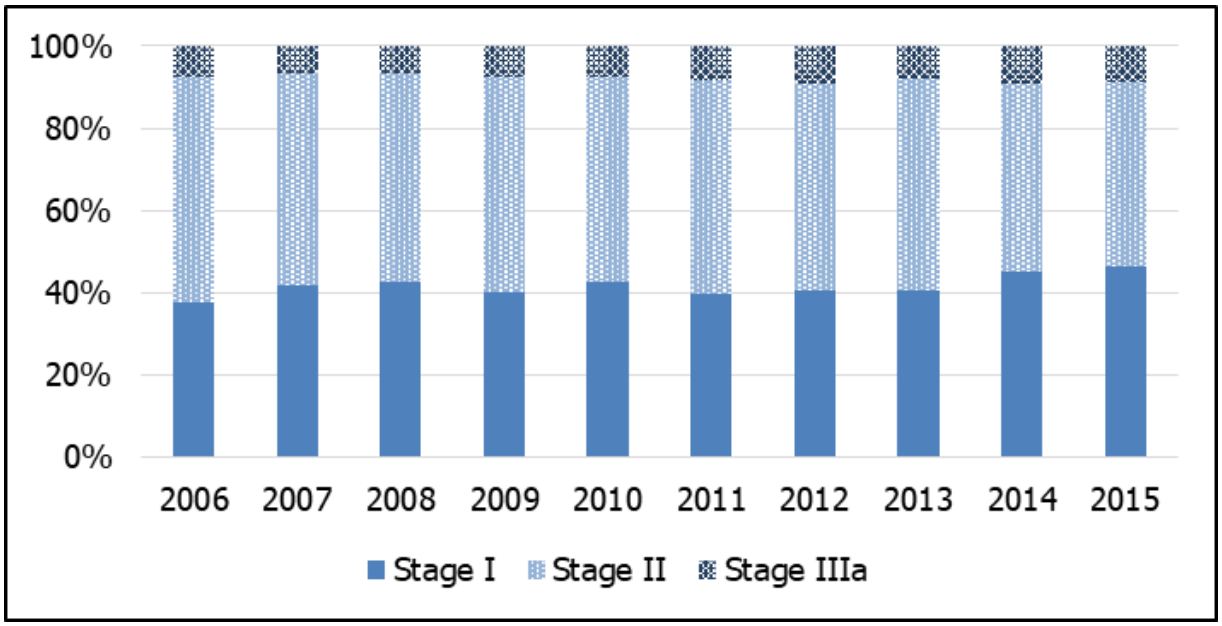
Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; LN – lymph node.
Source: National Cancer Registry Ireland.
* Data for the 15-19 years and 20-24 years age groups have been excluded from the figure given the small number of recorded cases.

3.4.2 Stage and treatment

Stage at diagnosis

As noted in section 3.3.3 breast cancer prognosis and treatment response are influenced by a variety of factors, including the stage at diagnosis. Of all HR+, HER2- early-stage invasive breast cancers diagnosed between 2006 and 2015, 42% (n=6,852) were diagnosed at stage I, 50% (n=8,225) were diagnosed at stage II and 8% (n=1,316) were diagnosed at stage IIIa. There was only slight variation (+/- up to 5%) in the annual proportion of cases diagnosed at each stage between 2006 and 2015 (Figure 3.2).

Figure 3.2 Stage at diagnosis of HR+, HER2- early-stage invasive breast cancer (2006-2015)

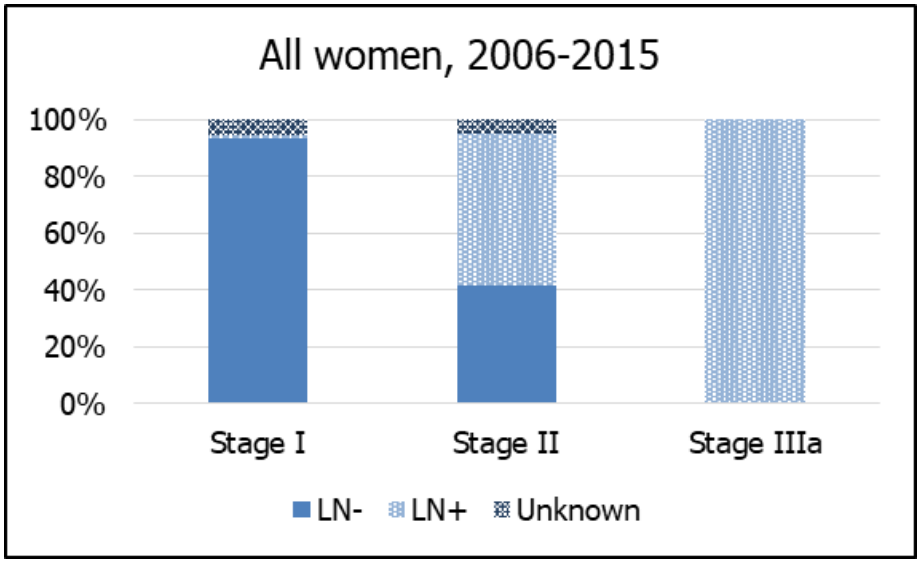


Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor.
Source: National Cancer Registry Ireland.

Stage at diagnosis by lymph node status

The stage by lymph node status at diagnosis of all HR+, HER2- early-stage invasive breast cancers diagnosed between 2006 and 2015 is presented in Figure 3.3. Between 2006 and 2015, at diagnosis, only 1% of stage I cases were LN+ (these were assumed to represent micrometastases), 54% of stage II cases were LN+ and all stage IIIa cases were LN+.

Figure 3.3 Stage at diagnosis of HR+, HER2- early-stage invasive breast cancer, by lymph node status (2006 - 2015)*



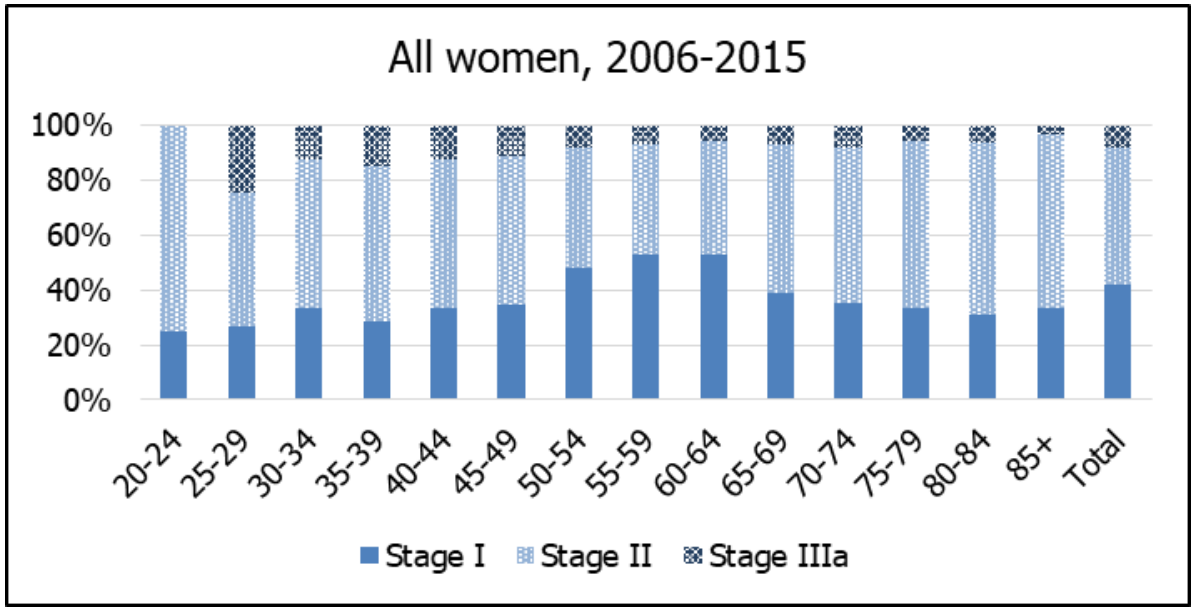
Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; LN – lymph node.
Source: National Cancer Registry Ireland.
* Data from 2016 onwards are not presented due to incompleteness.

Stage at diagnosis by age

The stage by age at diagnosis of all HR+, HER2- early-stage invasive breast cancers diagnosed between 2006 and 2015 is presented in Figure 3.4.

As noted in section 3.4.1, most HR+ HER2- stage I-IIIa breast cancer cases between 2006 and 2015 were diagnosed in those aged 50 to 64 years. In this age group, 55% (n=3,754), 38% (n=3,093) and 39% (n=507) were diagnosed at stage I, II and IIIa, respectively. By comparison, the percentage diagnosed at stage I was consistently lower in those younger than 50 years or older than 64 years.

Figure 3.4 Stage at diagnosis of HR+, HER2- early-stage invasive breast cancer, by age group (2006-2015)*



Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; LN – lymph node.
Source: National Cancer Registry Ireland.
* Data from 2016 onwards are not presented due to incompleteness.

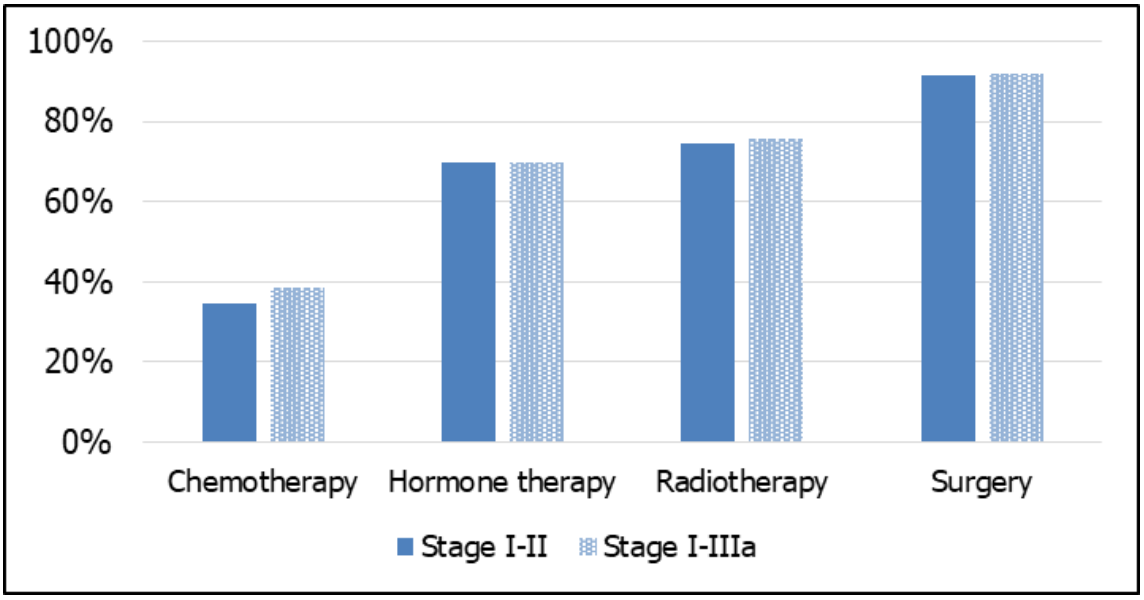
Treatment received following diagnosis

The NCRI capture data on treatments received in the first year following diagnosis. The proportion of women with HR+ HER2- early-stage breast cancer that received chemotherapy, hormone therapy, radiotherapy and or surgery within one year of diagnosis is presented in Figure 3.5.

Of women with stage I-II breast cancer diagnosed between 2010 and 2015 (that is, the most recent year with complete data on therapies, receptor status, stage, and survival), 91% (n=8,955) underwent surgery, 74% (n=7,286) received radiotherapy, 70% (n=6,855) received hormone therapy, and 35% (n=3,393) received chemotherapy within the first year of diagnosis. Of those who received chemotherapy, 99% (n=3,359) also underwent surgery.

Of women with stage I-IIIa breast cancer diagnosed between 2010 and 2015, 92% (n=9,836) underwent surgery, 76% (n=8,106) received radiotherapy, 70% (n=7,497) received hormone therapy, and 38% (n=4,124) received chemotherapy within one year of diagnosis. Of those who received chemotherapy, 99% (n=4,086) also underwent surgery.

Figure 3.5 Proportion of patients receiving treatment, by treatment type, within one year of diagnosis of HR+, HER2- early-stage invasive breast cancer (2010-2015)*



Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; LN – lymph node.
Source: National Cancer Registry Ireland.
* Data from 2016 onwards are not presented due to incompleteness.

3.4.3 Mortality and survival

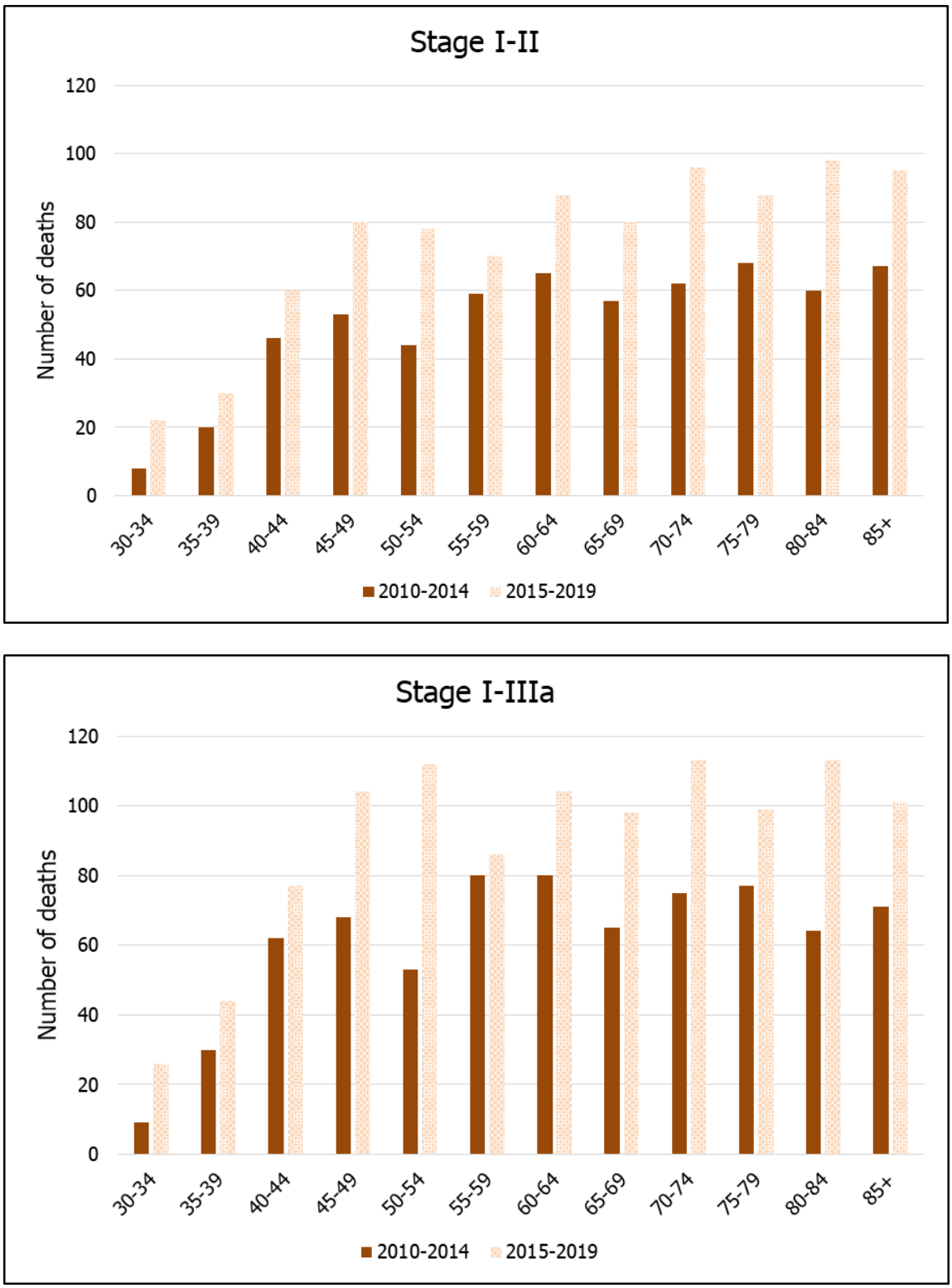
Absolute number of deaths

Between 2014 and 2019, the five-year rolling average for stage I-II cases ranged between 73 (calculated based on the period from 2006 to 2010) and 178 (calculated based on the period from 2015 to 2019) deaths per year. The five-year rolling average for stage I-IIIa cases ranged between 87 (calculated based on the period from 2006 to 2010) and 216 (calculated based on the period from 2015 to 2019) deaths per year.

Age at death

HR+ HER2- breast cancer deaths by age is presented in Figure 3.6. The total number of deaths generally increased with increasing age. However, the age groups with the highest total deaths differed between stage I-II and stage I-IIIa breast cancers and over time.

Figure 3.6 HR+, HER2- early-stage breast cancer mortality, by age group (2010-2019)*



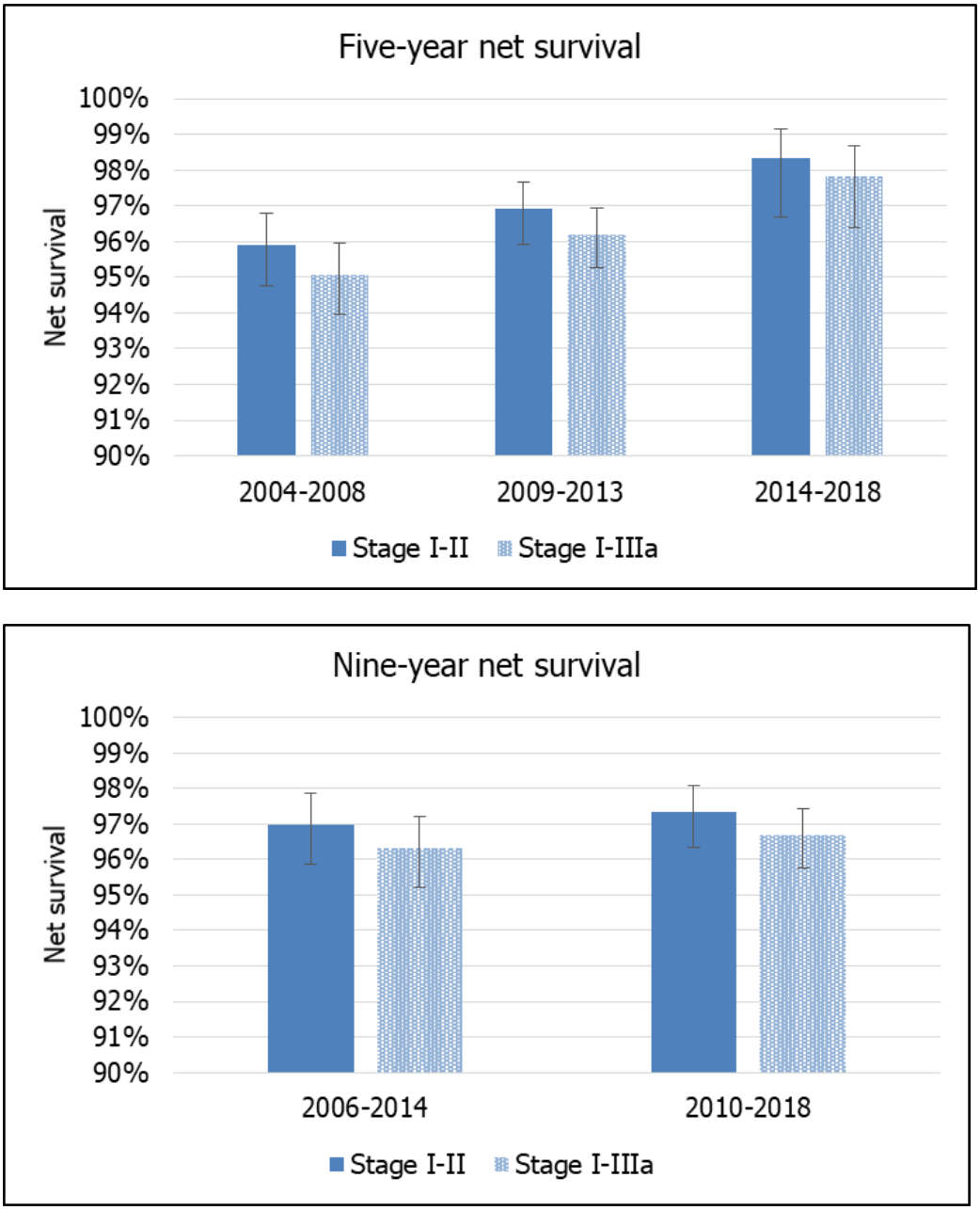
Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; LN – lymph node.
Source: National Cancer Registry Ireland.
* Data for the 15-19 years, 20-24 years and 25-29 years age groups have been excluded from the figure given the small number of recorded deaths.

Survival

The cumulative survival statistics for HR+, HER2- early-stage invasive breast cancer are presented in Figure 3.7. Five-year net survival is presented for the three most recent five-year diagnosis cohorts (that is, 2004-2008, 2009-2016 and 2014-2018). Nine-year net survival is also presented for two diagnosis periods (that is, 2006-2014 and 2010-2018).

Estimates of five-year net survival ranged between 95% (2004-2008) and 98% (2014-2018). Figure 3.7 shows a marginal improvement in survival over time, with a slightly higher net survival (~1% difference) in stage I-II than stage I-IIIa cases. This pattern was mirrored in the nine-year net survival estimates, which ranged between 96% (2006-2014) and 97% (2010-2018) across the time periods.

Figure 3.7 Survival estimates for HR+, HER2- early-stage breast cancer mortality, by time period (2004-2018)



Key: HER2 – human epidermal growth factor receptor 2; HR – hormone receptor; LN – lymph node.
Source: National Cancer Registry Ireland.
Notes:

- Net survival estimates were derived by the NCRI using the method proposed by Perme et al..[158] This method comprises an 'improved' version of relative survival that facilitates better comparability between alternative populations and age groups by taking competing mortality into account, and represents the cumulative probability of a patient surviving a given time in the hypothetical situation in which the disease of interest is the only possible cause of death (that is, survival having controlled for other possible causes of death by comparison of observed survival with the expected survival of people of the same age and gender in the general population).
- Error bars represent 95% confidence intervals.
- Figures are not age-standardised.

3.5 Discussion

The burden of breast cancer is significant; it is estimated to be the leading cause of cancer-related mortality in women worldwide while also being a rarer cause of mortality in men.[18, 159, 160] In Ireland, breast cancer is the most commonly diagnosed invasive cancer in women, affecting one in eight women over their lifetime.[115, 116] This chapter provides an overview of the natural history of breast cancer, informed by international literature, and describes data provided by the NCRI on the morbidity and mortality associated with HR+, HER2- early-stage breast cancer in Ireland.

It is challenging to estimate, based on NCRI data, the population that are eligible for Oncotype DX® or another GEP test as a support to decision-making. There were on average 1,806 new cases of HR+ HER2- stage I-IIIa breast cancer diagnosed annually in Ireland between 2015 and 2019, the majority of whom would be assumed to have undergone surgery within the first year of diagnosis (92% of those diagnosed between 2010 and 2015 underwent surgery). This, therefore, indicates a substantial number of people are eligible for the use of GEP tests as a decision input.

It is important to note that Oncotype DX® was reimbursed under the public system in Ireland in 2011,[9, 10, 15] and chemotherapy uptake trends within the target population may have been influenced by the availability of the test during this period. However, there is a lack of publicly available data to understand national trends in the usage of Oncotype DX® as a support to decision-making, and sufficient chemotherapy usage trend data for the years prior to and post-introduction of Oncotype DX® were not available to enable such an exploratory analysis. Nonetheless, as reported in section 2.3.4, studies conducted in Irish cohorts have demonstrated a reduction in chemotherapy use in women with early stage breast cancer in the years following this reimbursement decision.[3, 16, 17]

In interpreting the epidemiological morbidity and mortality data presented within this chapter, there is important contextual information that should be borne in mind. Firstly, the NCRI published a report in 2022 which included an overview of breast cancer trends between 1994 and 2019.[161] The NCRI report shows that while the absolute number of breast cancer deaths has risen over this time period, survival rates have improved and the mortality rate has reduced. This is likely due to a combination of factors relating to the introduction of population-based screening, better awareness of symptoms, improvements in care, availability of new treatments, and changes in the underlying risk exposure of the population. Notably, some of the overall improvements in breast cancer mortality may be the result of the availability of HER2-directed treatment (such as trastuzumab) for early and locally advanced breast cancer. This treatment could have improved outcomes in the

HER2+ population,[162, 163] which would otherwise have had inferior prognosis to that within the HER- population.[164]

There are a number of limitations relevant to the burden of disease data presented. Firstly, the incidence and mortality data presented in this chapter should be considered underestimates due to under-recording of receptor status in earlier years (such as 2010-2012), while later years (such as 2017-2019) may not yet have all stage or receptor status data available. Secondly, data on lymph node status were also not available on a subset of patients each year. Therefore, whether incidence and mortality of such patients fell into the LN- or LN+ categories was unclear. This level of incompleteness was generally small with regards to incidence (consistently less than 5% missing), but more variable with regards to mortality (7% on average). Thirdly, the hormonal therapy data for the period from 2006-2015 are likely to be incomplete due to the prescription and or administration of some therapies outside of hospital settings along with cases for which date of initiation of hormonal therapy was not known (which may have led to such treatments not being registered).[165]

Due to these three limitations and the likelihood of confounding arising from changes in the completeness of these data, incidence and mortality rates of HR+, HER2- early stage breast cancer have not been presented in this chapter. Finally, published evidence has indicated that age at menopause can impact the risk of breast cancer.[128] Menopausal status is not routinely collected for breast cancer cases registered with the NCRI. Therefore this information is not presented. However, data were presented for the populations younger and older than 50 years, an age range commonly used as a proxy for menopausal status.[166]

4 Clinical effectiveness

Key points

- A systematic review was undertaken to assess the prognostic ability (that is, the ability to accurately predict breast cancer outcomes), predictive ability (that is, the ability to identify people who are most likely to benefit from chemotherapy) and decision impact of four gene expression profiling (GEP) tests (Oncotype DX®, MammaPrint®, Prosigna®, EndoPredict®).
- This review updates a review by the government agency Ontario Health which included 53 relevant studies. The updated search identified an additional 34 studies, resulting in a total of 87 relevant studies included in the current review. These considered the prognostic ability (n=49), predictive ability (n=24), and decision impact (n=14) of GEP tests (note that some studies reported data in more than one category).
- Most prognostic and predictive studies were not effectively designed to address whether an individual GEP test, or GEP tests in comparison to each other, can offer additional prognostic and or predictive information beyond that provided by routinely assessed clinicopathologic factors.

GEP test prognostic abilities

- Evidence from 49 studies suggests that each of the four GEP tests likely has modest prognostic value for providing an estimate of a patient's likely future risk of cancer recurrence and or survival, with greater consistency of evidence among LN- populations than LN+.
- There was considerable variation across study designs, analytic approaches, risk cut-off scores used within GEP tests, choice of outcomes, and study populations examined. Therefore, meaningful quantification of each test's ability to predict cancer recurrence and or survival was not possible.
- Direct comparisons of GEP tests were sparse (LN- populations: n=6; LN+ populations: n=4), making it difficult to differentiate between the prognostic abilities of the tests.
- Each GEP test may add prognostic value beyond that of other prognostic information available to clinicians and patients (that is, clinical and pathological information), although the extent to which such value is added is unclear.

GEP test predictive abilities

- Three tests (Oncotype DX®, MammaPrint® and EndoPredict®) are indicated for predictive use. Of these, RCT evidence for predictive ability is available for two

tests: Oncotype DX® and MammaPrint®. The associated trials are MINDACT for MammaPrint® and TAILORx and RxPONDER for Oncotype DX®.

- Among LN- patients:
    - evidence for the predictive ability of MammaPrint® from the MINDACT trial indicated that MammaPrint® does not offer predictive value beyond that of a modified Adjuvant! Online algorithm; this algorithm incorporated data on estrogen receptor status, human epidermal growth factor receptor 2 (HER2) status, nodal status, tumour grade, and tumour size.
    - evidence for the predictive ability of Oncotype DX® from the TAILORx trial indicated that LN- women with an Oncotype DX® recurrence score (RS) 11-25 could be safely spared chemotherapy, although this finding was uncertain due to major limitations. Specifically, the unbalanced participant flow and participant selection in TAILORx likely biased results and limited generalisability to the Irish setting. Additionally, unlike MINDACT, TAILORx had no comparator for Onctoype DX®, meaning that its relative predictive ability was not assessed.
- Among LN+ patients:
    - findings from the MINDACT trial supported the predictive utility of MammaPrint® among high clinical risk LN+ patients aged 50 years and over; the trial results suggested that patients in this group with a low genomic risk score may be safely spared chemotherapy. However, this finding was uncertain and its generalisability to an Irish context is unclear due to the clinical risk assessment tool used and the use of frozen rather than FFPE tissue samples.
    - the RxPONDER trial supported the predictive ability of Oncotype DX®, suggesting that postmenopausal LN+ women with an RS 0-25 can be safely spared chemotherapy. However, these findings are derived from the first five years of data of a planned 15-year follow up. Further, similar to TAILORx, RxPONDER was limited by the lack of a comparator for Onctoype DX®, meaning that its relative predictive ability was not assessed.
- No trials assessed the predictive abilities of EndoPredict®, the only other test indicated for predictive use.
- There were no direct comparisons of the GEP tests; therefore, differentiating between the predictive abilities of the tests was not feasible.

GEP test decision impact

- Across all GEP tests, 24 studies evaluating the impact of GEP test results on treatment recommendations found that between approximately 20% and 50% of treatment decisions were observed to have changed as a result of test administration. This suggests that the use of GEP tests impacts treatment recommendations. It is important to note that these studies did not assess whether these changes in treatment recommendations led to improved patient outcomes.

Concordance between tests

- Large differences in the categorisation of patients across tests have been observed at an individual patient level. This discordance in risk group assignment, and the fact that there is minimal overlap in the genes assessed across tests, suggests that there may be more than one way of genetically predicting risk. However, despite differences in the individual level categorisation, the overall proportions of patients identified as low, intermediate, or high risk have been found to be comparable across tests.

4.1 Introduction

This chapter reports on a systematic review to address the following research question:

- Among patients with HR+, HER2-, and LN- or LN+ (1-3) early-stage (stages I to IIIa) invasive breast cancer, how do the gene expression profiling (GEP) tests EndoPredict®, MammaPrint®, Prosigna® and Oncotype DX® compare to each other in terms of their:
    - prognostic accuracy
    - predictive accuracy
    - impact on clinical decision-making?

This research question significantly overlaps with a systematic review performed as part of a comprehensive HTA published in 2020 by the government agency Ontario Health and the Canadian Agency for Drugs and Technologies in Health.[29] Therefore, the objective of this chapter is to update the Ontario Health HTA systematic review and review new studies in the context of the prior Ontario Health review findings.

4.2 Methods

The full details on the methods used for this chapter are described in the accompanying protocol. A summary of the methodology is provided here.

4.2.1 Quality assessment of prior review

As the current review updates the search from the Ontario Health review, two researchers independently assessed the quality of this prior review using the AMSTAR 2 critical appraisal tool.[167] This assessment identified that two items from the Critical Domains specified in AMSTAR 2 were not included in the Ontario systematic review:

- A protocol was not registered before commencement of the review
- Justification for excluding individual studies at the full text screening stage was not provided.

Additionally, four other items, which were not deemed 'critical,' were not included in the Ontario systematic review:

- Study selection was not performed in duplicate
- Data extraction was not performed in duplicate
- Funding sources for the studies included in the review were not reported
- Potential sources of conflict of interest, including any funding they received for conducting the review, were not reported.

The completed AMSTAR 2 tool for the Ontario Health systematic review is presented in Appendix Table A1.

Given that the study selection and data extraction components of the Ontario review were not performed in duplicate, as is recommended, there is an increased potential for error to have emerged within these steps. Cross-checking was therefore undertaken within the present review to confirm the appropriateness of the study selection and the data extraction accuracy. Briefly, as the review's study search was run in November 2018, the selection of the studies included in the Ontario review was cross-checked with those studies included in a NICE review of GEP tests which examined papers published as of February 2017.[44]

In order to confirm the inclusion of relevant studies published between February 2017 and November 2018, studies from this time period that were identified by the literature search conducted for the present review were cross-checked with those within the Ontario Health review. This cross-checking did not reveal any studies missed during the Ontario Health review selection process. Next, data were extracted for six randomly selected studies in the Ontario review and cross-checked with the original papers. Similarly, cross-checking of this random sample of data extractions did not reveal any material issues; only the source of the sample size reported for one study was unclear.

4.2.2 Search strategy

The draft search strategy for this present review was peer reviewed by a Health Library Ireland Librarian using the Peer Review of Electronic Search Strategies (PRESS) checklist.[168] The finalised search strategy was run on Medline (EBSCO) and translated to Embase (OVID) and The Cochrane Library on 22 June 2022. As the previous review included articles to November 2018 the search was limited by publication date (2018 – current), and the 'Randomized Controlled Trial' and 'Cohort Study' study design filters developed by the Health Library Ireland Evidence Team were applied. On 23 June 2022, clinicaltrials.gov (https://clinicaltrials.gov/ct2/home) and the International Clinical Trials Registry Platform portal (https://www.who.int/clinical-trials-registry-platform) were also searched for relevant studies. No limits were applied to these searches. Furthermore, a grey literature search was conducted in the International Network of Agencies for Health Technology Assessment database (https://database.inahta.org/), Epistemonikos (https://www.epistemonikos.org), TRIP database (https://www.tripdatabase.com/) and National Institute for Health and Care Excellence (NICE; https://www.nice.org.uk). A google search (https://www.google.com/) was also carried out and the first 100 results were screened. Forward citation searching on seven identified relevant studies was carried out in Google Scholar

(https://scholar.google.com/) on 28 June 2022. All results were deduplicated in Endnote and imported to Covidence for screening. Finally, potentially relevant studies were submitted by GEP test manufacturers after the updated search was completed. These studies were screened and relevant studies included.

4.2.3 Inclusion and exclusion criteria

Prognostic and predictive studies were included for this review update if they:

- included HR+, HER2-, and LN- or LN+ (1-3) early-stage (stages I to IIIa) invasive breast cancer patients
- reported data on freedom from distant recurrence, disease-free survival, or overall survival relating to the prognostic or predictive performance of Oncotype DX®, EndoPredict®, MammaPrint®, or Prosigna®
- comprised 1) randomised controlled trials (RCTs), 2) reanalyses/retrospective analyses of RCTs, or 3) retrospective analyses of data from a prospectively assembled database/registry across multiple institutions.

Prognostic and predictive studies were excluded if they comprised:

- retrospective analyses of data from a prospectively assembled database/registry in a single institution
- analyses of data from a retrospective review of medical records
- an assessment of the effectiveness of GEP tests in the context of extending adjuvant endocrine therapy and not chemotherapy
- patients with only a specific subtype of breast cancer
- examination of non-genetic tests only (for example, immunohistochemical 4 + Clinical score (IHC4+C))
- non-human studies
- conference abstracts and preprints.

Decision impact studies were included if they:

- assessed the change in the number of chemotherapy and no chemotherapy recommendations before and after use of a GEP test across multiple institutions
- were conducted in Europe (that is, the EU27 and Norway, Switzerland, Ukraine, and the UK); this criterion was chosen due to expected geographical differences in chemotherapy uptake rates, such as higher rates in the US compared to Europe.[169]

Decision impact studies were excluded if they comprised:

- retrospective analyses of data from a prospectively assembled database or registry in a single institution
- analyses of data from a retrospective review of medical records

- analyses of patients with only a specific subtype of breast cancer
- examination of non-genetic tests only (for example, IHC4+C)
- conference abstracts and preprints.

All studies were screened by two reviewers and disagreements were resolved by a third reviewer.

4.2.4 Data extraction

For the extraction of data from studies identified within the updated search, a data extraction form was constructed in Microsoft Excel® (Microsoft Corporation, Redmond, WA, US) and piloted using examples of each study design, after which the form was revised and finalised. Data were extracted by one reviewer and checked by a second reviewer. Disagreements were resolved by discussion. Data on study characteristics (for example, country, enrolment dates, design), authors (conflicts of interest), patient characteristics (for example, age, sex, menopausal status), and outcomes of interest (that is, prognostic performance, predictive performance, and or decision impact) were extracted.

Outcomes extracted to assess prognostic and predictive accuracy were:

- freedom from distant recurrence
    - incorporating freedom from distant recurrence (that is, where the cancer has recurred in distant parts of the body post surgery), and freedom from second primary cancer or death.
- disease-free survival
    - incorporating freedom from disease recurrence (for example, ipsilateral breast tumour recurrence, local recurrence, regional recurrence, distant recurrence, contralateral second primary invasive cancer, second primary non-breast invasive cancer [excluding non-melanoma skin cancers]), and freedom from second primary cancer or death from any cause.
- overall survival (that is, freedom from death due to any cause).

Where available, relevant data that did not perfectly match these outcome definitions were extracted and reported along with the most comparable outcomes. For example, breast cancer-specific survival was reported along with disease-free survival.[170, 171] Instances of this are noted throughout the results tables.

Data extracted to assess test decision impact were:

- the proportion of patients whose treatment recommendation changed from chemotherapy to no chemotherapy based on GEP test results

- the proportion of patients whose treatment recommendation changed from no chemotherapy to chemotherapy based on GEP test results
- the proportion of patients whose treatment recommendation changed based on GEP test results.

4.2.5 Risk of bias assessment of included studies

The authors of the Ontario review assessed the risk of bias for studies included in the review using a variety of tools. For the present update, the risk of bias for all of the newly identified studies was assessed using the corresponding tools: Cochrane Risk of Bias[172] tool (RoB2) for randomised controlled trials; the Prediction Model Risk of Bias Assessment Tool[173] for prognostic studies; and the Risk of Bias Assessment Tool for Nonrandomized Studies[174] for nonrandomised predictive ability studies. Risk of bias was assessed by one reviewer and checked by a second reviewer. Disagreements were resolved by discussion. For RoB2 the risk of bias was reported at the study level as risk of bias did not differ across outcomes within studies.

4.2.6 GRADE certainty of evidence

The certainty of a body of evidence may be evaluated using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework.[175] The GRADE assessment conducted within the Ontario review did not differentiate GEP tests (that is, GEP tests were considered as a whole rather than individually).

For the present review, GRADE was conducted only on studies which comparatively evaluated GEP tests, and included studies identified from the Ontario review and those identified in the updated search.

The GRADE assessment uses five factors to determine confidence in the body of evidence. These factors are as follows:

- risk of bias
- inconsistency
- indirectness
- imprecision
- publication bias.

The factors are used to assess the evidence associated with a specific outcome; this body of evidence can thereby be downgraded or upgraded. Evidence was graded as

high, moderate, low or very low, indicating the confidence in the effect. Agreement between two researchers was required for the GRADE outcome to be finalised.

4.2.7 Data presentation and synthesis

Findings of the studies identified within the current review were presented narratively and placed in the context of the previous Ontario review findings. For consistency of reporting and to aid comparability of results identified within the Ontario review and the present updated review, numerical results were transformed where necessary as follows.

In the context of prognostic studies, where studies reported outcomes in terms of the proportion of patients with distant recurrence, disease recurrence, or death, the numerical result was converted for presentation in the current report such that it was expressed in terms of freedom from these outcomes. For example, Pece et al.[176] reported ten-year distant recurrence rates of 10.9% and 24.8% for Low/Intermediate and High-risk groups, respectively; these were reported in the present report in terms of freedom from distant recurrence rates equating to 89.1% and 75.2%, respectively. Outcomes for which this conversion has been performed are noted in the results tables. Hazard ratios were reported such that better outcomes (that is, greater disease-free survival, freedom from distant recurrence, or overall survival rates) in a lower-risk group compared to a higher-risk group were represented by a hazard ratio greater than one.

Similarly, in predictive studies, hazard ratios are reported such that better outcomes in the chemotherapy group compared to the no chemotherapy group are represented by a hazard ratio greater than one. When individual studies represented this with a hazard ratio of less than one, the inverse of the hazard ratio and confidence intervals are presented here to ensure consistency of reporting and aid comparability of results across studies. For example, Kalinsky et al.[70] reported freedom from distant recurrence rates among premenopausal women of 96.1% and 92.8% for those who did and did not undergo chemotherapy, respectively, and the associated hazard ratio reported in the manuscript was 0.58, 95% confidence interval (0.39-0.87). This hazard ratio and confidence interval were inverted (that is, 1/0.58 [1/0.87-1/0.39]) and reported as 1.72 (0.73-1.23) in the current report. Outcomes for which this has been done are noted in the results tables.

Data were summarised and presented in tabular form and as narrative syntheses. Results were stratified by outcome, test, and lymph node status. Differences based on menopausal status were also discussed. When menopausal status was not assessed, but data were presented for patients aged above and below 50 years (an age cut-off which has been used as a proxy for menopausal status in previous

studies[70, 177]), these were discussed. Due to the heterogeneity of included studies, meta-analysis was not undertaken.

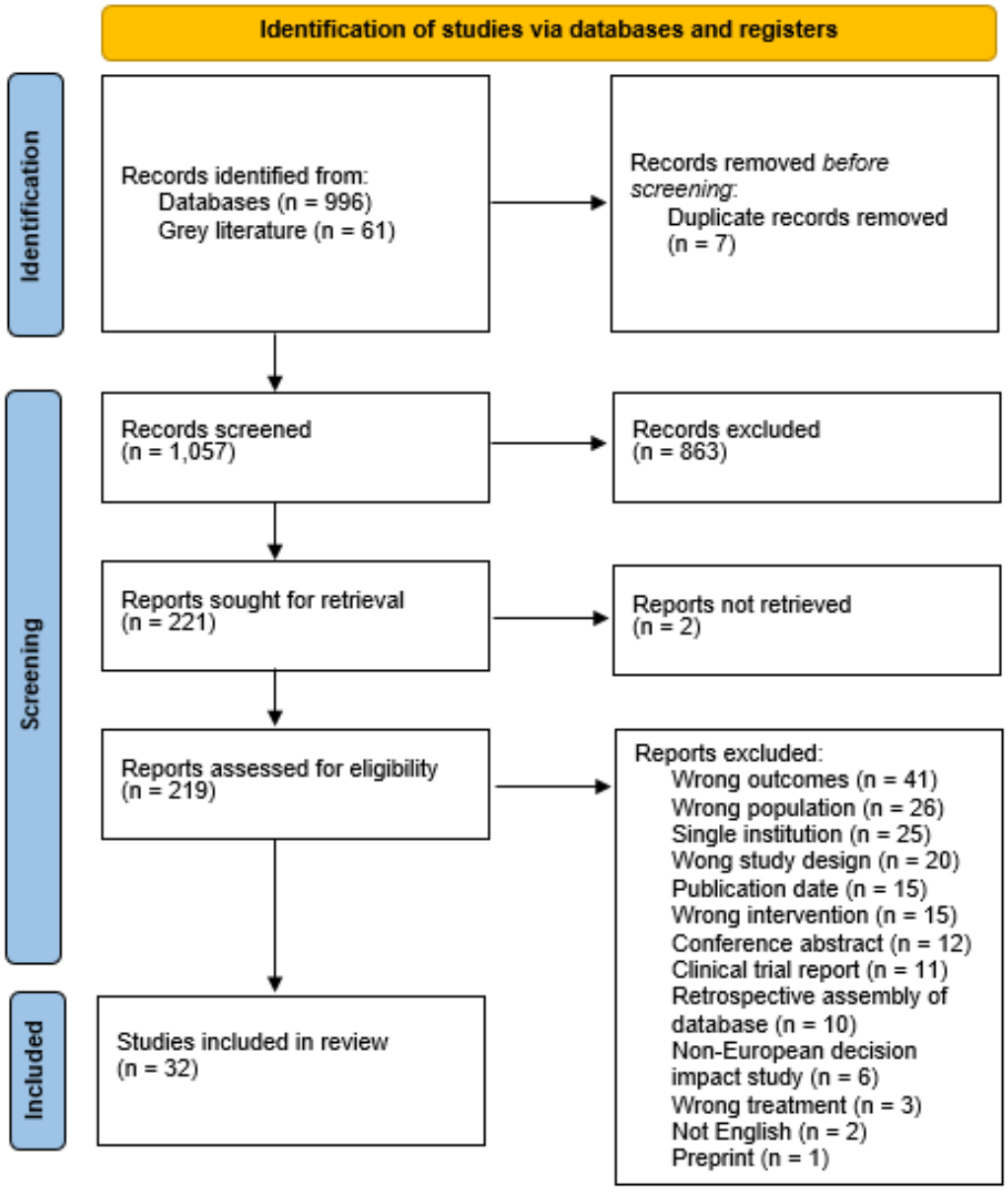
As decision impact studies included in the current review were limited to those conducted in Europe (see 'inclusion and exclusion criteria', above), only decision impact studies included in the Ontario review that were performed in a European setting were discussed here.

4.3 Results

4.3.1 Quantity of new evidence

The database search yielded 996 citations, with a further 61 citations identified from the grey literature search and forward and backward citation searching of included studies. Of these, 836 were excluded at the title and abstract screening stage and 221 full-text articles were screened, of which 189 were excluded. In total, 32 studies published since the Ontario Health HTA were included in this review update. Figure 4.1 presents the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram for the search results. As described in the following sections, the findings of these 32 studies, and an additional two studies published prior to the Ontario Health HTA identified from the manufacturer submissions,[178, 179] were tabulated and narratively appraised in the context of the findings of the Ontario review. The Ontario review identified 53 relevant studies as of November 2018, resulting in a total of 87 studies.

Figure 4.1 PRISMA flow-diagram for the search results in the updated search



4.3.2 Overview of results

The 'Study characteristics' Section 4.3.3 first presents the characteristics of the studies identified within the Ontario Health review, followed by the characteristics of the studies identified within the present review update. For the subsequent main results sections, the overall body of evidence is presented separately in terms of two major patient subgroups; LN- (Section 4.3.4) and LN+ (Section 4.3.5) patients. These two sections are subdivided further into the following three overall categories of outcome:

- **prognostic accuracy of tests**. This refers to the degree to which GEP tests can accurately predict the risk of an outcome and can discriminate people with different recurrence and survival rates (for example, whether survival rates differ between patients with low and high genomic risk). Generally, in the context of this review, observational studies provide the highest certainty of evidence regarding prognostic factors. Secondary analyses of RCTs can also provide evidence regarding prognosis,[180] although eligibility criteria for RCTs usually result in the exclusion of patients relevant for the assessment of prognostic factors. Additionally, eligible patients may decline to participate in RCTs for reasons related to their prognosis. The most commonly reported statistics displayed for prognostic accuracy were survival rates per test-assigned risk group, as estimated using the Kaplan-Meier method, and hazard ratios (HRs) between groups, as estimated using Cox proportional hazard models. When available, adjusted HRs (aHRs) were reported (typically adjusted for patient, treatment, and or tumour characteristics), although most estimates identified in the studies were unadjusted.
- **predictive accuracy of tests.** This refers to the degree to which GEP tests can identify people who will benefit most from chemotherapy. The strongest evidence in this section was derived from RCT data in which patients were randomised to chemotherapy or no chemotherapy. Ideally, a study would test for the interaction between treatment and GEP test risk group to assess whether the GEP test can predict a differential treatment effect between risk groups. Observational studies that reported outcomes among patients within certain risk categories who did and did not undergo chemotherapy were also included. Patients in these studies were not randomised to treatment groups nor were tests for interactions performed. Therefore, such observational studies were at a higher risk of confounding.
- **decision impact of tests.** These studies report treatment recommendations prior to using a GEP test and the proportion of these recommendations that changed after the use of a GEP test. Follow up of patients beyond the post-test treatment recommendation was typically not reported.

Throughout the results sections, the evidence directly comparing GEP tests is presented, followed by the evidence for each individual GEP test. Within the subsections detailing the evidence for each individual GEP test, the evidence identified from the Ontario review is presented first, followed by the new evidence identified in the current systematic review update.

Assessment of the risk of bias and of the certainty of the body of evidence

Data extraction and risk of bias assessments for the evidence derived from the Ontario review were conducted as part of that review and not repeated within the present update. As described in section 4.2.5, risk of bias assessment was performed for all studies identified in the present update (Appendix table A3, A4 and A5).

As described in section 4.2.6, GRADE was performed to assess the certainty of comparative studies that examined combinations of GEP tests. This involved consideration of the total body of such evidence, that is, that derived from both the Ontario review and the present update. The following GEP test comparison pairs were identified as having evidence available for assessment using GRADE:

- Oncotype DX® versus Prosigna® prognostic ability
- Oncotype DX® versus EndoPredict® prognostic ability
- EndoPredict® versus Prosigna® prognostic ability
- Oncotype DX® versus MammaPrint® prognostic ability.

GRADE specifies that RCT evidence starts at high quality for interventional studies and low quality for observational evidence. For prognostic studies, high quality prospective, longitudinal cohort studies provide high confidence. As prognostic ability was the main outcome under evaluation, we considered all prospective cohort studies to be of high certainty evidence at the outset and retrospective analyses of RCTs were of low certainty evidence.

4.3.3 Study characteristics

The total number of studies included in the present review and their source (Ontario Health review versus present updated search) are presented in Table 4.1. The following describes the study characteristics separately for each of these sources.

Table 4.1 Total number of studies included in this review and their source

	Prognostic evidence			Predictive evidence			Decision impact evidence		
	Total	**Ontario**	**Update**	**Total**	**Ontario**	**Update**	**Total**	**Ontario**	**Update**
LN- data	30	21	9	15	6	9	13	8	5
LN+ data	30	21	9	9	4	5	5	2	3
LN- and LN+ mixed	9	1	8	5	2	3	12	12	0
Total	49	30	19	24	10	14	24	19	5

Note: Decision impact studies from the Ontario Health review were included in the present review only if they were based on European data.

Ontario Health review

The Ontario Health review included a total of 53 studies, 30 of which provided evidence on prognostic effects (LN- n=21; LN+ n=21; LN-mixed n=1), 10 of which related to predictive effects (LN- n=6; LN+ n=4; LN-mixed n=2), and 19 of which considered the decision impact of using GEP tests in European countries (LN- n=8; LN+ n=2; LN-mixed n=12); note that some studies reported data in more than one category so the number of studies across categories do not add up to 53. The population which was most represented within the studies was the US (n=15) with other studies being mainly from European countries. Oncotype DX® was the most frequently studied GEP test across prognostic (LN- n=16; LN+ n=19), predictive (LN- n=13; LN+ n=7), and decision impact studies (LN- n=9; LN+ n=5).

Study designs included in Ontario Health review

The following briefly describes the study designs identified (that is, RCTs designed to evaluate GEP tests, followed by retrospective analyses of data from RCTs designed for other purposes) and notes the availability of comparative data and the key limitations of studies identified within the Ontario review.

The Ontario review identified relevant evidence from two RCTs which evaluated the non-inferiority of endocrine therapy alone relative to chemo-endocrine therapy within specific GEP test-assigned risk groups. These included the **TAILORx trial** for Oncotype DX®, as reported by Sparano et al.,[66] and the **MINDACT trial** for MammaPrint®, as reported by Cardoso et al.[100]

MINDACT was an open-label trial designed to evaluate whether endocrine therapy alone is non-inferior to chemoendicrine therapy among patients with discordant genomic and clinical risk. The study initially enrolled patients with LN- disease (79.0%) but was later revised to include women with LN (1-3) disease (20.9%). A total of 6,693 patients were included and categorised based on their genomic and clinical risk (assessed using a modified version of Adjuvant! Online). The proportions in the categories was as follows: low clinical risk and low genomic risk (41%); low clinical and high genomic risk (8.8%); high clinical and low genomic risk (23.2%); and high clinical risk and high genomic risk (27.0%). Patients with discordant results were randomly assigned to the adjuvant chemotherapy group or the no adjuvant chemotherapy group and patients with concordant risk were recommended to receive chemotherapy (high/high-risk patients) or no chemotherapy (low/low-risk patients). Limitations of the MINDACT trial include the open-label nature of the study, changes in patient eligibility requirements, and dropout rates (22%, n=481 of 2,187 patients who underwent randomisation). Additionally, the generalisability of the trial to an Irish setting may be limited as molecular diagnostic testing was performed on frozen samples of the resected tumour whereas current practice is to use a FFPE tissue sample.

The **TAILORx trial** was a randomised trial comparing chemotherapy plus endocrine therapy to endocrine therapy alone designed to show non-inferiority of endocrine therapy alone for invasive disease-free survival among HR+, HER2−, and LN− women with an Oncotype DX® intermediate risk RS score (RS 11-25). Limitations of the trial, from the perspective of this assessment, include: 1) an unbalanced flow of participants randomised to each group (whereby 32% of women assigned to the chemotherapy plus endocrine therapy arm did not complete the study protocol compare to 17% of patients assigned to the endocrine therapy alone arm); 2) the randomisation of patients to chemotherapy who may not have received chemotherapy under an Irish treatment pathway, and 3) the lack of a comparator for Oncotype DX®.

Another key data source was the **TransATAC** cohort derived from the ATAC RCT ('Arimidex, Tamoxifen, Alone or in Combination'). The ATAC trial, initiated in 1996 and conducted across 21 countries, was performed to compare the efficacy and safety of adjuvant treatment approaches in postmenopausal women with early-stage operable breast cancer.[181] The 'TransATAC' project was initiated in 2002 to

retrospectively establish a tissue bank from archival histopathology blocks from HR+ patients in the monotherapy arms of the ATAC trial.[181] By 2010, ten-year follow-up data from the trial were reported, enabling the TransATAC study to support the assessment of the potential for tumour markers to predict long-term outcomes.[182] Several assessments resulted, for example, individually considering the risk scores of Oncotype DX®, Prosigna® and EndoPredict® followed by studies comparing the prognostic performance of risk scores. The latter included the closest fully comparative study identified in this review; this study, conducted by Sestak et al.[183] in the TransATAC cohort, evaluated the prognostic accuracy of EndoPredict®, Oncotype DX®, and Prosigna®.

In the Ontario Health review there were 26 retrospective analyses of data collected as part of trials designed for different purposes, although many of these analyses were conducted using data from the same trials. Data from only 12 trials were used, including the ABCSG-8 trial which was used in eight studies and the ABCSG-6 and the ATAC (TransATAC cohort) trials which were used in three and four studies, respectively. Prospective cohort studies and retrospective analyses of cancer registries such as the Surveillance, Epidemiology, and End Results [SEER] database were also included.

Five studies evaluated multiple GEP tests, although none evaluated the performance of all four tests within the same population.[16, 18, 19-20, 68] The TransATAC study by Sestak et al.[183] which evaluated the prognostic accuracy of EndoPredict®, Oncotype DX®, and Prosigna®, was the closest fully comparative study.

The authors of the Ontario review appraised all of their included studies as generally low to moderate quality, with the most common areas for risk of bias being patient selection and analysis. Methods used for patient recruitment and selection were unclear in some studies (for example, whether consecutive patients were enrolled). Some studies did not perform multivariable analyses to consider potential confounding factors or, where such analyses were performed, it was unclear how potential confounders were chosen for the analyses.

Additional studies identified in the present review update

Characteristics of the additional studies identified are presented in Appendix Table A2. Briefly, 34 new studies were identified, 19 prognostic (LN- n=9; LN+ n=9; LN-mixed n=8), 14 predictive (LN- n=9; LN+ n=5; LN-mixed n=3), and five decision impact (LN- n=5; LN+ n=3; LN-mixed n=0) studies. Similar to the Ontario review, the majority of studies were from the US (14/34 exclusively in the US and 5/34 among more than one country including the US) and Oncotype DX® was the most frequently studied test in prognostic, predictive, and decision impact studies among

both LN- and LN+ patients. The total number of studies considered in the current review and their source are presented in Table 4.1.

Three RCTs relevant to the present review were identified; Kalinsky et al.[70] (the **RxPONDER** trial for Oncotype DX®), Sparano et al.[67] (the **TAILORx** trial for Oncotype DX®), and Piccart et al.[177] (the **MINDACT** trial for MammaPrint®).

TAILORx and **MINDACT** are described above. The **RxPONDER** trial was a randomised comparison of chemoendocrine therapy versus endocrine therapy alone among HR+, HER2−, and LN+ patients with an Oncotype DX® recurrence score (RS) 0-25. It aimed to assess whether the risk of disease recurrence increased with higher RS values and whether the benefit of chemotherapy relative to endocrine therapy also increased with a higher RS. It also examined whether chemotherapy benefit differed by menopausal status.

There were eight retrospective analyses of seven trials, (two were conducted using data from the ABCSG-8 trial). Retrospective analyses of cohort studies and cancer registries such as the National Cancer Database (n=6) and SEER database (n=5), both from the US, were also included.

There is substantial patient overlap across studies when they use data from the same trials and potentially also across RCT cohorts and cancer registries. Similar to studies in the Ontario review, the most common areas for risk of bias across studies were around patient selection and analysis. Risk of bias assessments for each study included in the present review update are presented in Appendix Tables A3-A5.

4.3.4 Lymph node-negative population

4.3.4.1 Prognostic studies

In total, considering the LN- population, 30 studies were identified which considered prognostic ability of GEP tests; these included 21 from the Ontario review and nine from the updated search. Data from the nine studies identified in the updated search are presented in Table 4.2.[67, 102, 184-190] Overall, the number of studies presenting prognostic ability differed for the individual tests: Oncotype DX®: n=16; MammaPrint®: n=6; EndoPredict®: n=3; Prosigna®: n=9.

Prognostic outcomes were freedom from distant recurrence, disease-free survival, and less commonly overall survival. Studies evaluated the prognostic ability of GEP tests by measuring the rates of these outcomes in a population and comparing whether rates differed across GEP test-assigned risk groups.

The most frequent sources of risk of bias across the evidence for all tests were unclear or non-consecutive participant enrolment in the study and inadequate controlling for potential confounders in analyses or study design.

Direct comparisons of prognostic accuracy across tests were sparse (n=4). This section will present the results of the studies first for those which involved head-to-head comparisons and then separately the evidence available for each individual test.

Head-to-head comparisons

Four studies identified in the Ontario Health review compared different GEP tests; all of these studies comprised secondary analyses of the TransATAC cohort (see section 4.3.3 for description of the TransATAC cohort). No studies reported any comparative data for MammaPrint®.

Sestak et al.[183] compared EndoPredict® (EPclin <3.3 and ≥3.3), Oncotype DX® (RS 0-17, 18-31, and 32-100), and Prosigna® (ROR ≤26, 27–68, and ≥69). The amount of prognostic information provided by each GEP alone was assessed individually using HRs and area under the ROC curve statistics for the modelled outcome (distant recurrence). All of the GEPs were found to predict, with statistical significance, distant recurrence during years 0 to 10 (and also years 5 to 10), though Prosigna® and EndoPredict® had the highest strength of prediction (as judged by the relatively higher hazard ratio point estimates and area under the ROC curve statistics than other tests). All GEPs also provided independent prognostic information beyond the clinicopathologic score when this was included as an additional variable in the model. When the three tests were combined with the clinicopathologic score, the Prosigna® ROR score provided the highest prognostic value, although differences across tests were not tested for statistical significance.

Similarly, Sestak et al.[191] compared Oncotype DX® (cut-off scores not reported) and Prosigna® (cut-off scores not reported). The study found that both gave similar prognostic information in years 0 to 5, as the difference in distant recurrence rate between the low- and high-risk groups was approximately 7% for both scores in years 0 to 5. However, Prosigna® had greater prognostic discrimination between the high risk and low risk groups than Oncotype DX® in years 5 to 10 (Difference in distance recurrence rate between risk groups: Risk of Recurrence (ROR) score=15.1% vs RS score=5.4%).

Buus et al.[192] compared EndoPredict's® EP (<5 and ≥5) and EPclin scores (EPclin <3.3 and ≥3) and Oncotype DX® (RS 0-17, 18-31, and 32-100), in endocrine-treated LN- patients. This study used likelihood ratio χ^2 and Kaplan-Meier survival analyses to compare prognostic information. The distant recurrence between tertiles

was compared for each test The difference in ten-year distant recurrence rates between the high-/non-low- vs low-risk groups was higher for EP (HR=5.15, 95% CI=2.44-10.85, $p<0.001$) than for RS (HR=3.72, 95% CI=2.17-6.39, $p<0.001$) and EPclin (HR=3.90, 95% CI=2.33-6.52, $p<0.001$) When this was examined over time, they found that the three test scores gave similar prognostic information in years 0 to 5 (Likelihood of distant recurrence LRχ^2: EP=15.5, LRχ^2: EPclin=17.0 LRχ^2: RS=18.7) but EndoPredict's® EPclin score gave the most prognostic information in years 5 to 10, followed by its EP score, and then Oncotype DX® (Likelihood of distant recurrence LRχ^2: EP=15.5, LRχ^2: EPclin=22.7 LRχ^2: RS=4.8).

Dowsett et al.[193] assessed the relative amount of information provided by Prosigna® or Oncotype DX® compared with each other in a patient population (n=649) with ER+ and LN- disease. Likelihood ratio values (ΔLRχ^2) were used to quantitatively measure the relative amount of information provided by one score compared with another. Both tests added significant prognostic information for distant recurrence beyond clinical parameters in all LN- patients, although more information was added by Prosigna's® ROR score (Difference in the likelihood of distant recurrence; ΔLRχ^2: ROR + clinical vs clinical alone=23.4) than by Oncotype DX® RS (ΔLRχ^2: RS + clinical vs clinical alone=10.2). Moreover, Prosigna® demonstrated better differentiation of intermediate- and higher-risk groups, correctly scoring more patients as high risk and fewer as intermediate risk compared to Oncotype DX®.

No new studies comparing the prognostic ability of individual tests were identified in the updated search.

The certainty of evidence from the four comparative studies was evaluated using the GRADE approach. For each of the comparisons (Oncotype DX® and Prosigna®, Oncotype DX® and EndoPredict®, and EndoPredict® and Prosigna®), the certainty of the evidence was considered to be very low due to imprecision and high risk of bias (Appendix Tables A6-A9).

Oncotype DX®

In the Ontario Health review,[29] nine studies examined the prognostic accuracy (freedom from distant recurrence n=9; disease-free survival n=2; overall survival n=4) of Oncotype DX® in LN- patients. The most common study design was retrospective analysis of RCTs (n=7), with retrospective analyses of databases also included (n=2). Sample sizes ranged between 301 and 40,134. The most common sources of risk of bias were unclear or non-consecutive participant enrolment in the study (n=6) and inadequate controlling for potential confounders in analyses or study design (n=9).

Eight studies from Ontario Health review, using data from four sources (the TransATAC,[182, 183, 191, 192] NSABP B-14,[194, 195] and NSABP B-20[69, 194] trials and the Stemmer Clalit Health Services database[196]) found that the traditional Oncotype DX® recurrence score cut-off scores (that is, RS 0-17, 18-30, and 31-100) were prognostic for freedom from distant recurrence and offered more limited evidence for overall survival (n=2). Two studies[197, 198] examined recurrence score cut-offs comparable with the 2019 TAILORx trial recommended cut-offs (that is, RS 0-10, 11-25, and 26-100). One, a large (n=40,134) retrospective analysis of the SEER database,[197] found that these cut-off scores were significantly prognostic (disease free survival (95% CI); RS 0-11=99.6% (99.4-99.8), RS 12–25=99.3% (99.2-99.4), RS 26-100=96.4% (95.6–97.0), between groups $p<0.01$) for disease-free survival. The other, a smaller analysis (n=465) of the WSG Plan B trial,[198] found that these cut-off scores were significantly prognostic for overall survival (overall survival; RS 0-11=99.2% (98.0–100), RS 12–25=98.3% (97.0–99.5), RS 26-100=96.7% (94.4–99.0), RS 0-11 vs. 26-100: HR=NR; $p<0.05$; RS 12–25 vs. 26-100: HR=NR; $p<0.05$). A retrospective analysis of the Clalit Health Services database,[196] found no significant difference between RS 0-10 and RS 11–25 for freedom from distant recurrence.

The updated search identified a further seven studies, using data from TAILORx,[67, 184] the Young Women's Breast Cancer Study,[185] The Breast Cancer Bank of Tissue,[186] the National Cancer Database (US),[187] and SEER (US),[102, 188] that examine the prognostic accuracy (freedom from distant recurrence n=5; disease-free survival n=4; overall survival n=2) of Oncotype DX® in LN- patients. The most common (n=5) study design was a retrospective analysis of a prospectively assembled database,[102, 185-188] with two retrospective analyses of RCTs[67, 184] also included. Sample sizes ranged between 300 and 119,328. Similar to the studies in the Ontario Health review, the most common sources of risk of bias were unclear or non-consecutive participant enrolment in the study (n=6) and inadequate controlling for potential confounders in analyses or study design (n=4).

Six studies examined the TAILORx trial recommended cut-off scores (that is, RS 0-10, 11-25, and 26-100), finding that they were prognostic for freedom from distant recurrence and disease-free survival.[67, 102, 184, 185] Only two studies reported data for overall survival,[187, 188] with both finding that Oncotype DX® could be prognostic for overall survival, however, results varied across age groups. Data from the TAILORx trial[67] presented in Table 4.2 also suggest that binary clinical-risk stratification based on tumour size and histologic grade may add further prognostic information to Oncotype DX®. Of note, Lynch et al.[186] was the only Irish study included in this review. This study used data from The Breast Cancer Bank of Tissue, an exploratory, translational, non-interventional multicentre biobank sponsored by Cancer Trials

Ireland that aims to identify potential biomarkers. Eligibility required prior registration with the TAILORx trial, participation in trial arms, and having sufficient tumour material available for molecular analysis. In a sample of 404 patients, they found that the low risk group (that is, women aged 50 years or younger with an RS of 0-15 or women aged over 50 years with an RS of 0-25) had a significantly higher ten year freedom from distant recurrence and disease-free survival than their high risk counterparts (both $p<0.001$).

The largest study was conducted by Ibraheem et al.[188] in 119,328 patients from the National Cancer Database. Using TAILORx cut-offs, they reported no difference in overall survival over five years between the Low and Intermediate risk groups, but significantly better overall survival in the Low-risk group compared to the High-risk group. They also observed a significant difference in the performance of Oncotype DX® across racial and or ethnic groups ($p<0.001$) with better performance observed among non-Hispanic white patients compared with non-Hispanic black patients and Hispanic patients, although the number of patients in minority populations, especially Asian American and Hispanic patients, was limited (8% non-Hispanic black, 4% Hispanic, and 4% Asian/Pacific Islander).

MammaPrint®

Six studies considered the prognostic accuracy of Mammaprint®, five from the Ontario Health review and one from the updated search.

In the Ontario Health review, five studies examined the prognostic accuracy (freedom from distant recurrence n=2; disease-free survival n=3; overall survival n=1) of MammaPrint® in LN- patients. Two studies were prospective observational studies from the RASTER study,[199, 200] two were retrospective analyses of RCTs,[201, 202] and one was a retrospective analysis of the Netherlands Cancer Institute database.[203] Sample sizes ranged between 216 and 652. The most common sources of risk of bias were unclear or non-consecutive participant enrolment in the study and inadequate controlling for potential confounders in analyses or study design.

Three studies stratified patients into Low and High risk categories,[200, 202, 203] one into Low and High risk categories which were further stratified by Low and High clinical risk,[199] and one into Ultra-low, Low, and High risk categories.[201] Evidence for the Low and High risk categories suggested that they were significantly prognostic for five-year freedom from distant recurrence (low risk=97.0%, (97% CI 94.7–99.4), high risk=91.7% (97% CI 87.9–95.7), between groups $p=0.03$),[200] ten-year overall survival (low risk=96.7%±2.3, high risk=49.6%±6.1),[203] and possibly ten-year disease-free survival although statistical significance was not assessed (low risk=93% (88–96), high risk=85% (75–91)).[202] Evidence from the RASTER study

suggested that MammaPrint® added further prognostic information to clinicopathological risk estimations from Adjuvant! Online although statistical significance was not assessed.[199] The retrospective analyses of an RCT (n=652) that stratified into three risk groups (Ultra-low, Low, and High) reported that these were significantly prognostic for disease-free survival over 20 years.[201]

This was supported by the one study identified in the updated search.[189] This small (n=80) retrospective analysis of an RCT found that the MammaPrint® Ultra-low, Low, and High risk categories were significantly prognostic (Table 4.2) for freedom of distant recurrence and disease-free survival over ten, 15 and 20 years.

EndoPredict®

Three studies considered the prognostic accuracy of EndoPredict®, all three were from the Ontario Health review with none identified in the updated search.

In the Ontario Health review, three studies examined the prognostic accuracy (freedom from distant recurrence n=3; disease-free survival n=0; overall survival n=0) of EndoPredict® in LN- patients. The three studies were retrospective analyses of RCTs and sample sizes ranged between 378 and 1,166. One study[192] had a high risk of bias for inadequate controlling for potential confounders in analyses or study design and in two studies[183, 204] it was unclear. One study[183] also had a high risk of bias relating to participant recruitment. All three studies found that EPclin binary categories (<3.3 and ≥3.3) were significantly prognostic for freedom from distant recurrence over ten years.

No new studies were identified in the updated search.

Prosigna®

Nine studies considered the prognostic accuracy of Prosigna®, eight from the Ontario Health review and one from the updated search.

In the Ontario Health review, eight studies examined the prognostic accuracy (freedom from distant recurrence n=7;[183, 193, 205-209] disease-free survival n=1;[210] overall survival n=0) of Prosigna® in LN- patients. Six studies were retrospective analyses of RCTs, one was a retrospective analysis of an observational study,[210] and one was a retrospective analysis of a database.[207] Sample sizes ranged between 591 and 1,455. One study was at low risk of bias in all domains,[206] three were at a high risk of bias for participant selection,[183, 207, 208] and six were at a high or unclear risk of bias for confounding.[193, 205, 207-210]

Two studies by Sestak et al. (retrospective analyses of both ABCSG-8 and ATAC RCTs and exclusively of the TransATAC cohort)[183, 208] assessed three risk categories

based on ROR scores ≤26, 27-68, and >69. Both studies found that these categories were significantly prognostic for freedom from distant recurrence over ten years. Four studies, two retrospective analyses of the ABCSG-8 trial,[205, 206] a retrospective analysis the Danish Breast Cancer Cooperative Group database[207] and a retrospective analysis of prospective observational study (Oslo1)[210] assessed risk categories Low (ROR ≤40), Intermediate (ROR 41-60), and High (ROR >60). These studies broadly supported the prognostic ability of these risk categories; however, one study did not report statistical significance,[206] one found no significant difference between the Intermediate and High group in predicting distant recurrence,[207] and one found no significant difference between the Low and Intermediate group in predicting breast cancer-specific survival among people who received tamoxifen only.[210] Two studies that did not report the risk group cut-off scores found that the ROR score added significant prognostic ability beyond clinical treatment score over ten years[193] and that a higher ROR score was associated with poorer 12-year recurrence-free survival, but there was no difference between Low or Intermediate ROR.[209]

One study, a retrospective analysis of The Women's Healthy Eating and Living trial,[190] was identified in the updated search (Table 4.2). It found that Prosigna® categories of Low, Medium and High (ROR categories not specified) was significantly prognostic for breast cancer-specific survival (p=0.007) and disease free survival (p=0.05) over ten years.

Table 4.2 Prognostic ability of GEP tests among LN- patients: studies identified from the updated search

Author, year	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI) or ±SE	Disease-free survival, % (95% CI) or ±SE	Overall survival, % (95% CI) or ±SE
Oncotype DX®						
Sparano et al, 2019[67] TAILORx		**Age ≤50y**				
	64 348	No CT / RS 0-10 / C-high No CT / RS 0-10 / C-low	9	100 ± 0.0 98.2 ± 0.9 HR=NA	90.7 ± 4.5 86.7 ± 2.3 HR=0.68 (0.24–1.92)	NR
	265 835	No CT / RS 11-25 / C-high No CT / RS 11-25 / C-low		87.7±2.4 95.3±1.0 HR=3.06 (1.78–5.25)	80.2±3.0 82.6±1.8 HR=1.27 (0.89–1.83)	
	252 791	CET / RS 11-25 / C-high CET / RS 11-25 / C-low		93.9±1.8 96.1±1.0 HR=2.20 (1.10–4.40)	86.5±3.0 88.7±1.4 HR=1.19 (0.76–1.88)	
	228 175	CET / RS 26-100 / C-high CET / RS 26-100 / C-low		84.8 ± 3.3 93.8 ± 2.5 HR=2.87 (1.23–6.65)	76.0 ± 4.2 85.2 ± 4.2 HR=2.27 (1.22–4.19)	
		Age >50y				
	281 879	No CT / RS 0-10 / C-high No CT / RS 0-10 / C-low	9	92.6 ± 3.4 97.4 ± 0.8 HR=2.20 (0.95–5.08)	72.8 ± 4.5 86.7 ± 1.5 HR=2.09 (1.47–2.96)	NR
	577 1,605	No CT / RS 11-25 / C-high No CT / RS 11-25 / C-low		90.7±1.9 96.5±0.6 HR=2.61 (1.65–4.11)	76.8±2.6; 86.4±1.1; HR=1.56 (1.21–2.00)	
	603 1,568	CET / RS 11-25 / C-high CET / RS 11-25 / C-low		91.7±1.5 96.0±0.7 HR=2.49 (1.60–3.87)	77.4±2.3 84.3±1.3 HR=1.61 (1.27–2.04)	
	542 414	CET / RS 26-100 / C-high CET / RS 26-100 / C-low		80.2 ± 3.9 93.0 ± 2.4 HR=3.35 (1.82–6.14)	67.9 ± 4.4 80.7 ± 3.8 HR=1.85 (1.28–2.66)	
Sparano et al, 2020[184] TAILORx	1,389	CET / RS 26-100	5	93.0 ± 0.8	88.1 ± 1.0	NR
	1,389	CET / RS 26-100	9	86.8 ± 1.7	76.2 ± 2.3	

Author, year	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI) or ±SE	Disease-free survival, % (95% CI) or ±SE	Overall survival, % (95% CI) or ±SE
Poorvu et al, 2019[185] **Young Women's Breast Cancer Study**	 127 125 48 33 195 72	**Traditional cut-offs** RS 0-17 RS 18-30 RS 31-100 **TAILORx cut-offs** RS 0-10 RS 11-25 RS: 26-100	 6 6	 97.5 (90.1-99.4) 93.1 (86.0-96.7) 86.4 (72.0-93.7) *p*=0.006 HR=NA 94.4 (66.6-99.2) 96.9 (92.7 to 98.7) 85.1 (72.9-92.1) *p*<0.001 HR=NA	NR	NR
Lynch et al, 2021[186] **The Breast Cancer Bank of Tissue**	235 169	Low risk: women aged 50 years or younger with an RS of 0-15 or women aged over 50 years with an RS of 0-25) High risk: Women aged ≤50 years with an RS of 16-100 and women aged >50 years with an RS of 26-100	10	97.2% 89.2% HR=1.6 (0.9-2.8)	88.8% 75.7% HR=3.2 (1.2-8.3)	NR
Iles et al, 2022[187] **National Cancer Database**	NR	RS 0-10 (low) RS 11-25 (Int) RS 26-100 (high)	5	NR	NR	High vs Low/Int **<40y** aHR=5.28 (2.61-10.66) **40-69y** aHR=1.89 (1.72-2.08); **≥70y** aHR=1.41 (1.26-1.59)
Ibraheem et al, 2020[188] **National Cancer Database**	27,795 73,951 17,582	RS 0-10 (low) RS 11-25 (Int) RS 26-100 (high)	5	NR	NR	Int v Low HR=0.95 (0.88-1.03) High v Low HR=2.26 (2.06-2.47) per 10-unit RS increase HR=1.32 (1.29-1.35)

Author, year	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI) or ±SE	Disease-free survival, % (95% CI) or ±SE	Overall survival, % (95% CI) or ±SE
Kantor et al, 2021[102] **SEER database**	11,300 30,685 5,309 29,353 12,628 112 2,025 1,544 953 1,184	Stage IA RS 0-10 RS 11-25 RS 26-100 RS 0-17 RS 18-25 Not Stage IA RS 0-10 RS 11-25 RS 26-100 RS 0-17 RS 18-25	5	 Range: 98.4-99.6 Range: 98.7-100 Range: 97.2-97.4 Range: 96.5-97.0	 Range: 99.5-100 Range: 98.3-99.5 Range: 96.4-100 Range: 100 Range: 96.5-96.7 Range: 86.7-93.7	NR
MammaPrint®						
Opdam et al, 2022[189] **IKA**	16 33 31 16 33 31	MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0	10	100 90 (79-100) 66 (51-86) *p*=0.0078	RFI 100; 90 (79-100) 66 (51-86) *p*=0.0078 BCSS 100 93 (84-100) 72 (58-91) *p*=0.0038	NR
	16 33 31 16 33 31	MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0	15	82 (61-100) 90 (79-100) 61 (45-83) *p*=0.0078	RFI 82 (61-100) 90 (79-100) 61 (45-83) *p*=0.0078 BCSS 92 (77-100) 93 (84-100) 66 (50-88) *p*=0.0038	NR

Author, year	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI) or ±SE	Disease-free survival, % (95% CI) or ±SE	Overall survival, % (95% CI) or ±SE
	16 33 31 16 33 31	20 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0		82 (61-100) 90 (79-100) 61 (45-83) *p*=0.0078	RFI 82 (61-100) 90 (79-100) 61 (45-83) p=0.0078 BCSS 92 (77-100) 93 (84-100) 60 (43-85) *p*=0.0038	NR
Prosigna®						
Pu et al, 2020[190]* WHEL	NR	Low Medium High Low Medium High	10	NR	BCSS 97% 95% 90% *p*=0.007 DFS 90%; 87%; 68% *p*=0.05	NR

Key: BCSS=breast cancer specific survival; C=clinical risk; CET=chemotherapy plus endocrine therapy; CI=confidence interval; CT=chemotherapy; DFS=disease-free survival GEP=gene expression profiling; HR=hazard ratio; Int=intermediate; LN-=lymph node negative; NA=not applicable; NR=not reported; RFI=-recurrence free interval; RS=recurrence score; y=year. TAILORx=Trial Assigning Individualized Options for Treatment; SEER=Surveillance, Epidemiology, and End Results programme; IKA=Integraal Kankercentrum Amsterdam (trial); WHEL=Women's Healthy Eating and Living Study

*Probability of distant recurrence, disease recurrence, or death were reported for this study and converted to freedom from these outcomes for presentation in this table

**HRs are inverted

4.3.4.2 Predictive studies

In total, 15 studies were identified that evaluated the predictive ability of GEP tests (Oncotype DX®: n=13; MammaPrint®: n=2) in LN- patients. Six of these studies were from the Ontario Health review[66, 69, 100, 105, 196, 211] and nine from the updated search. Data from eight of the studies identified in the current review are presented in Table 4.3[171, 177, 185, 212-216] (data were not reported for the ninth study[102]). Two RCTs, TAILORx for Oncotype DX® and MINDACT for MammaPrint®, which prospectively evaluated the predictive ability of GEP tests, were identified; all other studies were nonrandomised retrospective analyses of RCTs or databases.

Head-to-head comparisons

No studies were identified that directly compared predictive ability across tests.

Oncotype DX®

Thirteen studies considered the predictive accuracy of Oncotype DX®, five from the Ontario Health review and eight from the updated search.

In the Ontario review, one RCT[66] (the TAILORx trial) and four nonrandomised studies (two retrospective analyses of the NSABP B-20 RCT,[69, 105] one of the Clalit Health Services database,[196] and one of the National Cancer Database[211]), provided evidence for the predictive ability of Oncotype DX® in LN- patients. The TAILORx trial was a randomised comparison of chemotherapy plus endocrine therapy versus endocrine therapy alone and was designed to show non-inferiority of endocrine therapy alone (described in Section 4.3.3) for invasive disease-free survival in women who have an intermediate risk RS score (RS 11-25) as determined using Oncotype DX® (clinical risk was not considered within this study). The nine-year rate of freedom from distant recurrence did not differ between women with RS 11-25 who did and did not receive chemotherapy (95.0%±0.5 vs 94.5%±0.5). Similar results for the predictive ability of Oncotype DX® were found within the nonrandomised studies.[69, 105, 196, 211]

Exploratory analyses in the TAILORx trial indicated that chemotherapy was associated with some benefit for people aged 50 years and under with an RS 16-25.[66] Moreover, a follow-up analysis of TAILORx data by Sparano et al.[67], as reported within the 'prognostic studies' section above, was identified within the present review as also reporting data relevant from a predictive perspective. This study aimed to assess whether clinical risk added prognostic value to Oncotype DX®-assessed genomic risk. In doing so, it indicated that withholding chemotherapy from a subset of women with RS 11-25 (specifically, those aged 50 years and under with high clinical risk) may lead to worse cancer outcomes (see Table 4.2). In further stratifying women by clinical risk based on tumour size and histologic grade, women

aged 50 years and under with an RS of 11-25 and high clinical risk were shown to have better freedom from distant recurrence and disease-free survival at nine years when they received chemotherapy plus endocrine therapy compared to those who received endocrine therapy alone (distant recurrence: 93.9%±1.8 vs 87.7%±2.4; disease-free survival: 86.5%±3.0 vs 80.2%±3.0), although these between-group differences were not tested for statistical significance.

Limitations in the TAILORx trial include no presentation of a per protocol analysis, an unbalanced flow of participants randomised to each group, and the high rate of nonadherence to the assigned treatment (whereby 18% of women assigned to the chemotherapy plus endocrine therapy arm did not receive chemotherapy and only 5% of patients assigned to the endocrine therapy alone arm did receive chemotherapy). Despite this, the TAILORx study results were interpreted by the TAILORx study authors as providing evidence that the Oncotype DX® assay may identify a large proportion of women with early breast cancer who can be spared adjuvant chemotherapy.

In the updated search, eight additional retrospective analyses of databases (SEER,[102, 171, 212, 213] National Cancer Database,[214-216] and Young Women's Breast Cancer Study[185]) were identified (see Table 4.3). Findings across these studies were mixed. For example, among patients with RS 11-25, two found chemotherapy benefit,[212, 216] two found no chemotherapy benefit,[185, 215] and one found no chemotherapy benefit among patients with RS 11-15 but significant chemotherapy benefit among patients with RS 16-20 and 21-25.[213] Further, one study found no chemotherapy benefit among patients with T1-2 N0 disease (tumour size no greater than 5cm, with no nodal involvement) and RS 18-25,[102] another found chemotherapy benefit among patients with RS 26-30 and RS 31-100,[171] and another study found chemotherapy among all RS groups (0-17; 18-30; 31-100).[214] These nonrandomised studies are at increased risk of confounding compared to RCTs as the full range of factors that contributed to the decision of whether a patient should undergo chemotherapy are unclear and may also influence the outcomes of interest.

MammaPrint®

Two studies considered the predictive accuracy of Mammaprint®, one from the Ontario Health review and one from the updated search; both were based on analysis of data from the MINDACT trial. In the Ontario review. Cardoso et al.[100] presented five-year findings of the MINDACT trial, while, in the only study identified in the updated search, five- and eight-year follow up data from this trial were presented by Piccart et al.[177]

As described in Section 4.3.3 'Study characteristics', MINDACT is an open-label non-inferiority trial, designed to evaluate whether the addition of MammaPrint® genomic testing to clinical risk scoring is beneficial in selecting patients for adjuvant chemotherapy. Clinical risk was assessed using Adjuvant! Online. Patients with discordant results (low clinical and high genomic risk (8.8%), or high clinical and low genomic risk (23.2%)) were randomly assigned to the adjuvant chemotherapy group or the no adjuvant chemotherapy group and patients with concordant risk were recommended to receive chemotherapy (high/high-risk patients (27.0%)) or no chemotherapy (low/low-risk patients 41%)). Limitations in the MINDACT trial include the open-label nature of the study, changes in patient eligibility requirements, and dropout rates (22%, n=481 of 2,187 patients who underwent randomisation).

Cardoso et al. reported that there was no significant difference in distant metastasis-free survival after five years between LN- patients in the high clinical risk and low genomic risk group who did or did not receive chemotherapy (95.7% vs 93.2%; HR=0.69, 95% CI (0.39-1.21), *p*=0.193).[100] However, at eight-year follow-up, Piccart et al. reported that patients in this risk group who received chemotherapy had a significantly higher distant metastasis-free survival rate (91.7% vs 89.2%; HR=0.60, 95% CI (0.38-0.96), *p*=NR).[177] Additionally, Piccart et al. reported that at both five- and eight-year follow up, patients in the high clinical risk and low genomic risk group who received chemotherapy had higher disease-free survival (five-year: 93.0% vs 90.1%; eight-year: 87.5% vs 83.4%) and overall survival (five-year: 98.5% vs 96.4%; eight-year: 95.5 vs 93.9%) rates, although these between group differences were not tested for statistical significance. Additionally, there was some uncertainty around these estimates as evidenced by the confidence intervals (up to ±5% around the point estimate).[177] Cardoso et al. reported that there was no significant difference in distant metastasis-free survival between LN- patients who did or did not receive chemotherapy in the low clinical risk and high genomic risk group (96% vs 95.1%; HR=1.09, 95% CI (0.54-2.19) *p*=0.815) which was supported by the five- and eight-year data reported by Piccart et al. (see Table 4.3). The MINDACT trial was limited by the performance of molecular diagnostic testing on frozen samples of the resected tumor whereas a FFPE tissue sample is used in practice currently in Ireland. This may limit the generalisability of the findings to the Irish setting.

EndoPredict®

No relevant studies identified.

Prosigna®

No relevant studies identified.

Table 4.3 Predictive ability of GEP tests among LN- patients: studies identified from the updated search

Author, year Study/ database name	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Oncotype DX®						
Cheng et al, 2020[212] SEER database	190 10,974 4,093 27,638 4412 2232	RS 0-10 / CT RS 0-10 / No CT RS 11-25 / CT RS 11-25 / No CT RS 26-100 / CT RS 26-100 / No CT	7	NR	NR	Survival rates NR. CT vs No CT: *p*=0.871 CT vs No CT: *p*=0.005 CT vs No CT: *p*=0.006
Cheng et al, 2021[213] SEER database**	413 12,639 1,461 10,407 2,225 4,586	RS 11-15 / CT RS 11-15 / No or unknown CT RS 16-20 / CT RS 16-20 / No or unknown CT RS 21-25 / CT RS 21-25 / No or unknown CT	5	NR	NR	HR=1.63 (0.73–3.01), *p*=0.275 HR=2.22 (1.19–2.78), *p*=0.006 HR=1.69 (1.15–2.25), *p*=0.006
Ma et al, 2021[171] SEER database**	943 943 1,194 1,194	 RS 26-30 / CT RS 26-30 / No CT RS 31-100 / CT RS 31-100 / No CT	5	NR	BCSS HR=1.85 (1.12-3.13), *p*=0.02 HR=1.56 (1.06-2.27), *p*=0.02 Interaction assessing relative CT benefit across groups: *p*=0.99	 HR=1.56 (1.09-2.27), *p*=0.02 HR=1.52 (1.12-2.00), *p*=0.006 Interaction assessing relative CT benefit across groups: *p*=0.66
Iorgulescu et al, 2019[214] National Cancer Database	327 3,264	RS 0-17 / CT RS 0-17 / No CT	5	NR	NR	98.8 (94.9-99.7) 96.3 (95.0-97.2) *p*=0.07

Author, year Study/ database name	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
	3,912 2,392 2,378 1,954 3,960 3,626	RS 18-30 / CT RS 18-30 / No CT RS 31-100 / CT RS 31-100 / No CT No RS / CT No RS / No CT				95.5 (94.0-96.6) 93.0 (90.8-94.7) *p*=0.002 91.8 (90.2-93.1) 79.6 (72.4-85.1) *p*<0.001 93.6 (92.4-94.7) 83.3 (81.3-85.0) *p*<0.001
Wang et al, 2020[215] National Cancer Database**	Male 126 54 252 39 Female 17,475 10,307 44,243 8,141	 RS 18-30 / No CT RS 18-30 / CT RS 11-25 / No CT RS 11-25 / CT RS 18-30 / No CT RS 18-30 / CT RS 11-25 / No CT RS 11-25 / CT	5	NR	NR	 CT vs No CT: aHR=0.87 (0.11-50.00) CT vs No CT: aHR=0.20 (0.02-2.17) CT vs No CT: aHR=1.25 (1.02-1.52) CT vs No CT: aHR=1.15 (0.92-1.43)
Weiser et al, 2021a[216] National Cancer Database	NR	RS 0-10 / CT RS 0-10 / No CT RS 11-25 / CT RS 11-25 / No CT RS 26-100 / CT RS 26-100 / No CT	5	NR	NR	97.8% 97.9% *p*>0.05 98.4% 97.9% *p*<0.05 96.1%, 93.6% *p*<0.05
Poorvu et al, 2020[185]	86 109	RS 11-25 / CT	6	97.3 (89.4-99.3) 96.7 (89.9-98.9)	NR	NR

Author, year Study/ database name	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Young Women's Breast Cancer Study		RS 11-25 / No or unknown CT		*p*=0.247		
MammaPrint®						
Piccart et al, 2021[177] MINDACT**	349 350 272 262 349 350 272 262	C-high / G-low / CET C-high / G-low / ET C-low / G-high / CET C-low / G-high / ET C-high / G-low / CET C-high / G-low / ET C-low / G-high / CET C-low / G-high / ET	5 8 	95.1 (92.1-97.0) 94.3 (91.2-96.3) 94.1 (90.3-96.4) 95.6 (92.2-97.5) 91.7 (88.1-94.3) 89.2 (85.2-92.2) HR=1.67, 95% CI (1.04-2.63) 91.1 (86.7-94.0) 91.8 (87.4-94.7)	93.0 (89.6-95.3) 90.1 (86.4-92.9) 91.3 (87.1-94.2) 91.6 (87.5-94.5) 87.5 (83.3-90.7) 83.4 (78.8-87.1) 85.5 (80.3-89.4) 83.6 (78.1-87.8)	98.5 (96.4-99.4) 96.4 (93.7-97.9) 96.1 (92.8-97.9) 98.4 (95.8-99.4) 95.5 (92.5-97.3) 93.9 (90.6-96.1) 92.6 (88.5-95.3) 92.9 (88.6-95.6)

Key: BCSS=breast cancer specific survival; CET=chemotherapy plus endocrine therapy; CI=confidence interval; CT=chemotherapy; ET=endocrine therapy; GEP=gene expression profiling; LN-=lymph node negative; RS=recurrence score; SEER=Surveillance, Epidemiology, and End Results Program.
*Probability of distant recurrence, disease recurrence, or death were reported for this study and converted to freedom from these outcomes for presentation in this table
**HRs are inverted

4.3.4.3 Decision impact studies

In total, 13 studies assessed changes in chemotherapy recommendations and or decisions before and after the use of GEP tests (Oncotype DX®: n=9; MammaPrint®: n=2, EndoPredict®: n=1; Prosigna®: n=2) in LN- patients. Note, these studies primarily assessed pre-test recommendations prospectively and did not include follow-up of clinical outcomes such as recurrence or survival. Changes in treatment recommendations among LN- patients in studies identified in the current review are presented in Table 4.4. These results are summarised below along with findings from studies reported in the Ontario Health review. There were no studies that directly compared the impact on treatment decisions across tests. As decision impact studies included in the current review were limited to those conducted in Europe, due to expected geographical differences in chemotherapy uptake rate (see Section 4.2.3 'inclusion and exclusion criteria'), only decision impact studies included in the Ontario review that were performed in a European setting are discussed here.

Oncotype DX®

Nine studies considered the impact of Oncotype DX® on treatment recommendations, five from the Ontario Health review and four from the updated search.

In the Ontario Health review, five European studies (based in France, Germany, Italy, Spain, and UK) [217-221] evaluated the change in treatment recommendations following use of Oncotype DX® in LN- populations. Sample sizes ranged from 107 to 527 and all patients were HR+ and HER2-. The proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy ranged between 0% and 11%, 2) from chemotherapy to no chemotherapy ranged between 8% and 60%, or 3) at all ranged between 12% and 60%. In the one UK study,[219] 0/136 (0%) patients changed from no chemotherapy to chemotherapy and 82/136 (60%) changed from chemotherapy to no chemotherapy. The risk of bias assessment (ROBANS) conducted by Ontario Health found that it was unclear across the five studies whether outcome assessments were blinded, it was unclear in three of the studies how patients were enrolled,[219-221] and in three of the studies confounding was a concern.[217, 220, 221]

The updated search identified four new studies; three conducted in Italy[222-224] and one in Spain.[225] Sample sizes ranged from 152 to 1,160. In the largest study,[224] chemotherapy recommendations reduced from 512 to 374 following use of Oncotype DX®, although the exact number of patients whose treatment recommendations changed from no chemotherapy to chemotherapy or vice versa were not reported. Across the other three studies, the proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy ranged between 4% and 10%, 2) from chemotherapy to no chemotherapy, ranged between 10% and 35%, or 3) at

all ranged between 14% and 45%. In the two studies where subgroup data was reported, all of the changes from no chemotherapy to chemotherapy were reported in the Intermediate RS 18-30 and High RS 31-100 groups. Changes from chemotherapy no chemotherapy were mostly in the Low RS 0-17 risk group, with some changes in the Intermediate risk group.[222, 225] All studies were at a high risk of bias for confounding and two were at a high risk of bias regarding participant selection.

MammaPrint®

Two studies considered the impact of Mammaprint® on treatment recommendations; one was from the Ontario Health review and one was from the updated search.

In the Ontario Health review, one study, conducted in the Netherlands, evaluated the change in treatment recommendations following use of MammaPrint® in LN- populations.[226] Among 660 patients (all HR+ and 97% HER2-), 38 (6%) recommendations changed from no chemotherapy to chemotherapy and 156 (24%) changed from chemotherapy to no chemotherapy. The risk of bias assessment (ROBANS) conducted by Ontario Health found that it was unclear whether the outcome assessment was blinded. Also, the study was found to be at a high risk of bias for participant selection, incomplete outcome data, and selective outcome reporting.

One new study of 467 patients (all HR+/HER2-), conducted in Spain, was identified in the updated search.[225] For 65 (14%) treatment recommendations changed from no chemotherapy to chemotherapy (all within the high risk group), and 125 (27%) changed from chemotherapy to no chemotherapy (all within the low risk group). This study was at a high risk of bias for confounding but at a low risk of bias across the other five domains.

EndoPredict®

No studies in the Ontario Health review evaluated the change in treatment recommendations following the use of EndoPredict® in LN- populations. One study was identified in the update search to have considered the impact of EndoPredict® on treatment recommendations in LN- populations; this study was identified from the updated search and was conducted in France.[227] Within this study, among 200 patients, 15 (8%) recommendations changed from no chemotherapy to chemotherapy (all within the high risk group) and 57 (29%) recommendations changed from chemotherapy to no chemotherapy (51 of which were within the low-risk category group). This study was at a high risk of bias (assessed by ROBANS) for confounding as no confounders were considered in analyses. However, low risk of bias was found across the other five domains.

Prosigna®

Two studies considered the impact of Prosigna® on treatment recommendations, both were from the Ontario Health review, with none identified from the updated search.

In the Ontario Health review, two studies, conducted in France and Germany, evaluated the change in treatment recommendations following use of Prosigna® in LN- populations.[228, 229] Sample sizes were 194 and 198 and all patients were HR+ and HER2-. The proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy ranged between 11% and 13%, 2) from chemotherapy to no chemotherapy ranged between 3% and 5%, or 3) at all ranged between 14% and 18%. The risk of bias assessment (ROBANS) conducted by Ontario Health found that for both studies it was unclear whether the outcome assessment was blinded and both studies were at a high risk of bias for confounding and selective outcome reporting.

Table 4.4 Changes in treatment recommendations among LN- patients: studies from the updated study

Author, year	Country	No CT to CT, n (%)	CT to no CT, n (%)	Total treatment change, n (%)
Oncotype DX®				
Cognetti et al, 2021[224]	Italy	NR	CT recommendations were reduced from n=512 to n=374	NR
Dieci et al, 2019[223]	Italy	15/152 (9.9%)	27/152 (17.8%)	42/152 (27.6%)
Pérez Ramírez et al, 2020[225]	Spain	Total population: 44/440 (10.0%) RS 0-17: 0/238 (0.0%) RS 18-30: 38/168 (22.6%) RS 31-100: 6/34 (17.6%)	Total population: 152/440 (34.5%) RS 0-17: 127/238 (53.4%) RS 18-30: 25/168 (14.9%) RS 31-100: 0/34 (0.0%)	Total population: 196/440 (44.5%) RS 0-17: 127/238 (53.4%) RS 18-30: 63/168 (37.5%) RS 31-100: 6/34 (17.6%)
Zambelli et al, 2020[222]	Italy	Total population: 12/267 (4.5%) RS 0-17: 0/166 (0.0%) RS 18-30: 6/85 (7.1%) RS 31-100: 6/16 (37.5%)	Total population: 26/267 (9.7%) RS 0-17: 19/166 (11.4%) RS 18-30: 7/85 (8.2%) RS 31-100: 0/16 (0.0%)	Total population: 38/267 (14.2%) RS 0-17: 19/166 (11.4%) RS 18-30: 13/85 (15.3%) RS 31-100: 6/16 (37.5%)
MammaPrint®				
Pérez Ramírez et al, 2020[225]	Spain	Total population: 65/467 (13.9%) Low: 0/297 (0.0%) High: 65/170 (38.2%)	Total population: 125/467 (26.8%) Low: 125/297 (42.1%) High: 0/170 (0.0%)	Total population: 190/467 (40.7%) Low: 125/297 (42.1%) High: 65/170 (38.2%)
EndoPredict®				
Penault-Llorca et al, 2020[227]	France	Total population: 15/200 (7.5%) Low risk: 0/135 (0.0%) High risk: 15/65 (23.1%)	Total population: 57/200 (28.5%) Low risk: 51/135 (37.8%) High risk: 6/65 (9.2%)	Total population: 72/200 (36.0%) Low risk: 51/135 (37.8%) High risk: 21/65 (32.3%)

Key: CT - chemotherapy; LN- - lymph node negative; NR - not reported; RS - recurrence score

4.3.5 Lymph node-positive and mixed lymph node status populations

4.3.5.1 Prognostic studies

In total, 30 studies evaluated the prognostic ability of GEP tests in LN+ patients (Oncotype DX®: n=19; MammaPrint®: n=3; EndoPredict®: n=6; Prosigna®: n=9). A further nine studies evaluated the prognostic ability of GEP tests in mixed LN status patients (Oncotype DX®: n=5; MammaPrint®: n=3; Prosigna®: n=2). Data from the 15 new studies identified in the current review are presented in Table 4.5. Overall, data were predominantly (n=26) presented on distant recurrence, while data on disease-free survival (n=13) and overall-survival (n=11) were less common. Direct comparisons of prognostic accuracy across GEP tests were sparse (n=6). The most frequent risk of bias across the evidence for all GEP tests were from unclear or non-consecutive participant enrolment in the study and inadequate controlling for potential confounders in analyses or study design.

Head-to-head comparisons

Six studies considered head to head comparisons of the prognostic accuracy across GEP tests, four from the Ontario Health review and two from the updated search.

Four studies in the Ontario Health review directly compared different GEP tests, three of which were retrospective analyses of the TransATAC cohort (described in Section 4.3.3),[183, 192, 193] and one was a retrospective analysis of a trial (GEICAM).[192] Sestak et al.[183] compared EndoPredict® (EPclin <3.3 and ≥3.3), Oncotype DX® (RS 0-17, 18-31, and 32-100), and Prosigna® (ROR ≤26, 27–68, and ≥69), finding that all GEP tests provided significant prognostic information among endocrine-treated women with LN+ disease during years 0 to 10; however, the prognostic ability of all three GEP tests was weaker for the LN+ population compared with the LN− population. EndoPredict® and Prosigna® provided the most prognostic value for late distant recurrence (Years 5-10: EndoPredict®=HR 1.87, 95% CI (1.27–2.76); Prosigna®=HR 1.65, 95% CI (1.08–2.51)). Oncotype DX® did not significantly provide prognostic information for late distant recurrence (5-10 years) on its own or in combination with clinical parameters. Dowsett et al.[193] assessed the relative amount of information provided by Prosigna® or Oncotype DX® compared with each other in patient population (n=268) with ER+ and LN+ cancer, 89% of which were also HER2-. Both tests added significant prognostic information for distant recurrence beyond clinical parameters in all LN+ patients, although more information was added by Prosigna's® ROR score than by Oncotype DX's® RS (likelihood ratio value, $\Delta LR\chi^2$=33.9, P <0.001). These findings did not differ in the mixed LN-status and 100% HER2- patient group. Buus et al.[192] compared EndoPredict's® EP (<5 and ≥5) and EPclin scores (EPclin <3.3 and ≥3) and Oncotype DX® (RS 0-17, 18-31, and 32-100) in endocrine-treated LN+ patients. In years 0 to 5 all tests were significantly prognostic for distant

recurrence, although EndoPredict's® EPclin score gave more prognostic information than both its EP score and Oncotype DX®, which were comparable (Likelihood ratios 0-10 years, LRχ^2: EPclin=48.3; LRχ^2: EP=14.5; LRχ^2: RS=8.0). Additionally, EndoPredict's® EPclin score gave stronger prognostic information in years 5 to 10, followed by its EP score, and then Oncotype DX® which was not significantly prognostic (Likelihood ratios 5-10 years, LRχ2: EPclin=16.1; LRχ2: EP=6.6; LRχ2: RS=1.0). In a retrospective analysis of the GEICAM RCT, a phase III trial in 1,246 women that compared adjuvant chemotherapy to adjuvant chemotherapy followed by five-year hormonal therapy, Martin et al.(230) compared EndoPredict® to Prosigna® (PAM50) in 536 ER+, HER2- patients who received endocrine therapy and adjuvant chemotherapy. Predictors including clinical information showed superior prognostic performance compared to molecular scores alone and all tests added prognostic information to common clinical parameters. EndoPredict's® EPclin scores (EPclin <3.3 and ≥3) better predicted freedom from distant recurrence that Prosigna's ROR cut-off scores <29, 29–65, and >65 (P=0.04, note only p-vales were reported), but not ROR cut-off scores <18, 18–65, >65 (P=0.09).

Two studies identified in the updated search considered different GEP tests. Both studies were at a high risk of bias relating to participant enrolment and one also for confounding.(231) In the first study, a retrospective analysis of 432 LN+ patients receiving endocrine therapy in the SCAN-B cohort, Vallon-Christersson et al.(231) reported that both Prosigna® (three groups, cut-off scores not reported) and Oncotype DX® (three groups, cut-off scores not reported) were significantly prognostic for overall survival over six years. In the second study, using data for mixed-LN status patients from the SEER database, Ibraheem et al.(188) propensity matched 5,042 pairs of patients who received either MammaPrint® or Oncotype DX®. Binary MammaPrint® risk groups and two different Oncotype DX® cut-off scores were used: the first had three risk groups, Low (RS 0-10), Intermediate (RS 11-25), and High (RS 26-100); and the second defined two risk groups: Low (LN- and RS 0-25 or LN+ and RS 0-10) and High (LN- and RS 26-100 or LN+ and RS 11-100). In adjusted analyses, all GEP tests were significantly prognostic for overall survival over five years.

The certainty of evidence from these comparative studies was completed using the GRADE approach. For each of the comparisons (for LN+; Oncotype DX® and Prosigna®, Oncotype DX® and EndoPredict®, EndoPredict® and Prosigna®; for LN mixed status; Oncotype DX® and MammaPrint®), low and very low certainty of evidence was achieved (Appendix Tables A6-A9). The low and very low certainty of evidence was due to imprecision and a high risk of bias.

Oncotype DX®

Nineteen studies considered the prognostic accuracy of Oncotype DX® in LN+ patients, 13 from the Ontario Health review and six from the updated search. Additionally there were five studies that considered the prognostic accuracy of Oncotype DX® in mixed LN status patients, all identified from the updated search.

In the Ontario Health review, thirteen studies examined the prognostic accuracy (freedom from distant recurrence n=9; disease-free survival n=4; overall survival n=6) of Oncotype DX® in LN+ patients. The most common study design was retrospective analysis of RCTs (n=9); data came from TransATAC,[182, 183, 192, 193] WSG Plan B,[198, 232] NSABP B-28,[233] PACS-01,[234] and SWOG-8814.[71] Additionally, two studies performed retrospective analysis of the SEER database,[197, 235] one of the Clalit Health Services database,[196] and one was a prospective observational study.[236] Sample sizes ranged between 77 and 6,483. The most common risks of bias were due to unclear or non-consecutive participant enrolment in the study and inadequate controlling for potential confounders in analyses or study design.

Nine studies examined the prognostic ability of the traditional Oncotype DX® recurrence score cut-off scores (that is, RS 0-17, 18-30, and 31-100). Six of these considered freedom from distant recurrence,[182, 183, 192, 196, 233, 234] three disease-free survival,[197, 234, 235] and four overall survival.[196, 234-236] These studies broadly found that these cut-off scores were prognostic for distant recurrence and disease-free survival. Evidence for overall survival was mixed, although the study follow-up periods (2 to 5 years) were likely insufficient to detect survival differences between groups. Two studies[198, 232] examined recurrence score cut-offs comparable with the 2019 TAILORx trial recommended cut-offs (that is, RS 0-10, 11-25, and 26-100). Both performed reanalyses of the WSG Plan B trial, finding that these cut-off scores were prognostic for freedom from distant recurrence and overall survival at five years. A retrospective analysis of the Clalit Health Services database[196] found no significant difference between RS 0-10 and RS 11–25 for freedom from distant recurrence at five years. Four studies[71, 233, 234, 236] examined the Oncotype DX® RS score as a continuous score, finding it to be significantly prognostic for distant recurrence,[233, 234] disease free survival,[71, 234] and overall survival.[236]

The updated search identified a further six studies that examined the prognostic accuracy (freedom from distant recurrence n=2; disease-free survival n=2; overall survival n=3) of Oncotype DX® in LN+ patients. Two studies were retrospective analyses of the National Cancer Database,[187, 188] and one each from SEER database,[102] the Young Women's Breast Cancer Study,[185] SCAN-B,[231] and the SWOG S8814 RCT.[237] Sample sizes ranged between 163 and 25,029. All six studies were at a high risk of bias for unclear or non-consecutive participant enrolment in the study and four were at a high risk of bias for inadequate controlling for potential confounders in analyses or study design.[102, 185, 231, 237]

Four studies examined the TAILORx trial recommended cut-off scores (that is, RS 0-10, 11-25, and 26-100). One large analysis of the National Cancer Database[188, 211] found no significant difference between the Low and Intermediate risk categories in predicting overall survival after five years, but the High risk category performed significantly better than the Low risk category. Additionally, the continuous RS score significantly predicted overall survival, finding that they were prognostic for freedom from distant recurrence and disease-free survival and offering some evidence for overall survival.[188] Evidence from the other three studies are unclear regarding freedom from distant recurrence or disease-free survival, but two of these studies are small (n=163 and n=165)[185, 187] and one stratified their sample into small subsamples based on AJCC pathologic prognostic stage.[211]

One small (n=163) study,[185] a retrospective analysis of the Young Women's Breast Cancer Study, found that the traditional cut-off scores (that is, RS 0-17, 18-30, and 31-100) were significantly prognostic for freedom from distant recurrence over six years, although the Low and Intermediate risk groups had materially the same freedom from distant recurrence rates (Freedom from distant recurrence at 6 years; RS 0-17=85.9%, RS 18-30=87.3%, RS 31-100=62.8%). One study,[237] a retrospective analysis of the SWOG S8814 RCT (n=165), pooled the Intermediate and High-risk groups and found that RS 0-17 and 18-100 were not significantly prognostic for disease-free survival over ten years.

The updated search identified a further five studies that examine the prognostic accuracy (freedom from distant recurrence n=2; disease-free survival n=2) of Oncotype DX® in mixed LN status patients. One was the comparative study described above by Ibraheem et al.[188] Two studies were retrospective analyses of databases (the National Cancer Database[215] and Georgia Cancer Registry[238]) and two were retrospective analyses of RCTs (TransATAC cohort derived from ATAC trial[176] and WSG-Plan B[239]). Sample sizes ranged between 163 and 111,746. All five studies were at a high risk of bias for unclear or non-consecutive participant enrolment in the study and three were at a high risk of bias for inadequate controlling for potential confounders in analyses or study design.[176, 215, 239] Of the two retrospective analyses of RCT, Pece et al.[176] found that Oncotype DX® cut-off scores RS 0-26 and 27-100 were significantly prognostic for freedom from distant recurrence over ten years and Nitz et al.[239] found that cut-off scores RS 0-25 and 26-100 were prognostic for disease-free survival over five years (statistical significance not reported). Collin et al.[238] reported that the traditional cut-off scores and cut-off scores RS 0-25 and 26-100 were significantly prognostic for disease-free survival over ten years, although some ethnic differences were observed. The only data in this report on male breast cancer came from Wang et al.[215] Using data from the National Cancer Database (male n=848; female n=110,898), they found that

both the traditional and TAILORx cut-off scores were significantly prognostic for freedom from distant recurrence over five years among both men and women.

MammaPrint®

Three studies considered the prognostic accuracy of Mammaprint® in LN+ patients, one from the Ontario Health review and two from the updated search. Additionally there were two studies that considered the prognostic accuracy of MammaPrint® in mixed LN status patients both identified from the updated search.

In the Ontario Health review, one study examined the prognostic accuracy of MammaPrint® in LN+ patients. In a retrospective analysis of 295 patients from the Clalit Health Services database, the Low and High risk categories were prognostic for overall survival over ten years.[203] This study had a high risk of bias for participant enrolment and confounding. Two small studies were identified in the updated search. One was a retrospective analysis of an RCT (IKA) with a high risk of bias relating to participant enrolment. In 55 LN+ patients, MammaPrint® (cut-off scores: MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0) was significantly prognostic for freedom from distant recurrence, breast cancer-specific survival, and recurrence free interval over ten, 15 and 20 years.[179, 189] The second was a retrospective analysis of frozen samples from the Netherlands Cancer Institute-Antoni van Leeuwenhoek hospital, Amsterdam, and the European Institute of Oncology, Milan; this study had a high risk of bias for participant enrolment and confounding.[179] In 241 LN+ patients, breast cancer specific survival (96.6% vs 76%) and distant recurrence (91% vs 76%) were significantly more favourable over ten years in the low-risk group compared to the high-risk group (both $p<0.001$). Additionally, three studies examined the prognostic accuracy of MammaPrint® in mixed LN status patients. One was the comparative study described above by Ibraheem et al.[188] Another was a retrospective analysis of the ABCSG-8 RCT that examined the prognostic accuracy of binary MammaPrint® risk categories in a mixed LN-status population, finding that it was significantly prognostic for freedom from distant recurrence at five years but not ten years.[240] Finally, an analysis of data from the MINDACT trial showed patients with ultralow-risk tumours have the best prognosis, distinctive from low risk.[241] Compared to patients with low-risk tumours, patients with ultralow-risk tumours had a lower risk of distant recurrence (aHR=0.65, 95% CI=0.45–0.94) and patients with high-risk tumours had a higher risk of distant recurrence (aHR=2.17, 95% CI=1.68–2.80).

EndoPredict®

Six studies, all among LN+ patients and from the Ontario Health review, considered the prognostic accuracy of EndoPredict®.

In the Ontario Health review, six studies examined the prognostic accuracy (freedom from distant recurrence n=5; disease-free survival n=1) of EndoPredict® in LN+ patients. All studies were retrospective analyses of RCTs: TransATAC cohort derived from ATAC trial,[183, 192] ABCSG-6 and ABCSG-8,[81, 204] ABCSG-8 alone,[242] and GEICAM.[230] Sample sizes ranged between 183 and 1,702. All studies were at a high or unclear risk of confounding, and two were at a high risk of bias regarding participant selection.[183, 242] All five studies found that EPclin binary categories (<3.3 and ≥3.3) were significantly prognostic for ten-year distant recurrence and one study[242] found that EPclin binary categories were significantly prognostic for local recurrence (that is, disease-free survival).

The update search did not identify any new studies which examined the prognostic accuracy of EndoPredict® in LN+ patients.

Prosigna®

Nine studies considered the prognostic accuracy of Prosigna® in LN+ patients, seven from the Ontario Health review and two from the updated search. Additionally there were two studies that considered the prognostic accuracy of Prosigna® in mixed LN status patients one from the Ontario review, and one from the updated search.

In the Ontario Health review, seven studies examined the prognostic accuracy (freedom from distant recurrence n=7) of Prosigna® in LN+ patients. Six studies were retrospective analyses of RCTs (TransATAC cohort derived from ATAC trial,[183, 191, 193] ABCSG-8,[205, 206] and GEICAM[230]) and one was a retrospective analysis of the DBCG database.[207] Six of the seven studies were at unclear risk of confounding,[183, 191, 193, 205, 207, 230] two were at a high risk of bias regarding participant enrolment,[183, 207] and one was at low risk of bias across all domains.[206] Different ROR cut-offs were used across studies, two used ROR categories ≤26, 27–68, and ≥69,[183, 191] two used ≤15, 16–40 and ≥40,[205, 206] one used <29, 29–65, and >65,[230]and one did not report the cut-off scores.[207] All studies found that Prosigna® was significantly prognostic for distant recurrence over 5-10 years. One study also examined the prognostic accuracy of Prosigna® in mixed LN status patients.[243] This retrospective analysis of the DBCG 77B RCT (n=460) found that the Prosigna® ROR cut-offs 8–51, 52-71, and 72-100 were significantly prognostic for disease-free survival and overall survival over ten years.

The updated search identified two new studies. Both studies were at a high risk of bias relating to participant enrolment and confounding. Firstly, in a retrospective analysis of 432 LN+ patients receiving endocrine therapy in the SCAN-B cohort, Vallon-Christersson et al.[231] reported that Prosigna® (three groups, cut-off scores not reported) was significantly prognostic for overall survival over six years. Secondly, in a retrospective analysis of LN+ patients (n=551) in the WHEL RCT, Pu

et al.[190] reported that Prosigna® (three groups, cut-off scores not reported) was significantly prognostic for breast cancer specific survival and disease-free survival over ten years. Additionally, a third study examined the prognostic accuracy of Prosigna® in mixed LN-status patients. In a retrospective analysis of the ABCSG-8 RCT, ROR cut-off scores of <57 and ≥57 were significantly prognostic for local recurrence over five and ten years.[244]

Table 4.5 Prognostic ability of GEP tests among LN+ and LN mixed status patients: studies from the updated study

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Multiple tests						
Vallon-Christersson et al, 2019[231] SCAN-B	130 92 201 129 136 158	Prosigna® Low-risk Medium-risk High risk Oncotype DX® Low-risk Medium-risk High-risk	6	NR	NR	Prosgina® *p*=0.00001 Oncotype DX® *p*=0.005
Ibraheem et al, 2020[188] (LN-mixed) National Cancer Database	2,908 2,134 1,140 3,068 834 3,503 1,539	MammaPrint® Low risk High risk Oncotype DX® RS 0-10, RS 11-25 RS 26-100 Oncotype DX® LN- & RS 0-25 or LN+ & RS 0-10 LN- & RS 26-100 or LN+ & RS 11-100	5	NR	NR	96.6 (95.3-97.6) 90.7 (88.3-92.6) aHR=2.25 (1.56-3.25) 95.3 (92.6-97.0) 94.8 (93.2-96.1) 87.6 (83.2-90.9) aHR (Int v Low)=1.04 (0.66-1.62); aHR (High v Low)=1.81 (1.05-3.09) 95.3 (93.8-96.4) 90.1 (87.3-92.4) aHR=1.63 (1.07-2.48)

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Oncotype DX®						
Woodward et al, 2020[237] SWOG S8814	100 65	RS 0-17 RS 18-100	10	NR	LRR 98.5% 88.9% *p*=0.051	NR
Poorvu et al, 2019[185] Young Women's Breast Cancer Study	54 69 40 14 88 61	Traditional cut-offs RS 0-17 RS 18-30 RS 31-100 TAILORx cut-offs RS 0-10 RS 11-25 RS: 26-100	6	 85.9 (72.6-93.0) 87.3 (76.0-93.5) 62.8 (45.1-76.2) *p*=0.004 92.3 (56.6-98.9) 85.2 (75.3-91.4) 71.3 (57.3-81.5) *p*=0.100	NR	NR
Ibraheem et al, 2020[188] National Cancer Database (US)	5,936 15,920 3,173	RS 0-10 RS 11-25 RS: 26-100	5	NR	NR	Int v Low HR=1.15 (0.97-1.36) High v Low HR=2.94 (2.43-3.56) per 10-unit RS increase HR=1.38 (1.31-1.44)
Iles et al, 2022[187] National Cancer Database (US)	NR	RS 0-10 RS 11-25 RS: 26-100	5	NR	NR	RS 26-100 vs RS 0-25 <40y aHR=1.76 (0.72-4.34) 40-69y aHR=2.82 (2.32-3.43) ≥70y aHR=1.62 (1.31-2.00)
Kantor et al, 2021[102] SEER database	833 2,141 135 2,212 762	Stage IA RS 0-10 RS 11-25 RS 26-100 RS 0-17 RS 18-25	5	 Range: 99.0-99.7 Range: 96.6-97.1	 Range: 99.5-100 Range: 98.3-99.5 Range: 96.4-100	NR

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
	470 1,934 635 1,458 946	Not Stage IA*** RS 0-10 RS 11-25 RS 26-100 RS 0-17 RS 18-25		 Range: 97.7-97.4 Range: 96.5-97.0	Range: 100 Range: 96.5-96.7 Range: 86.7-93.7	
Wang et al, 2020 (LN-mixed)[(215)] ** **National Cancer Database (US)**	Male 496 260 92 294 406 148 Female 65,935 36,174 8,789 25,929 68,882 16,087	Traditional cut-offs RS 0-17 RS 18-30 RS 31-100 TAILORx cut-offs RS 0-10 RS 11-25 RS: 26-100 Traditional cut-offs RS 0-17 RS 18-30 RS 31-100 TAILORx cut-offs RS 0-10 RS 11-25 RS: 26-100	5	 95.7 (92.1-97.7) 87.5 (79.6-92.5) 82.3 (63.9-91.9) *p*=0.02 97.2 (90.3-99.2) 91.0 (86.1-94.3) 83.2 (70.5-90.7) *p*=0.003 97.1 (96.9-97.3) 96.1 (95.8-96.4 91.5 (90.7-92.4) *p*<0.0001 96.6 (96.2-96.9) 97.0 (96.8-97.2) 92.9 (92.3-93.5) *p*<0.0001	NR	NR
Pece et al, 2022[(176)] (LN-mixed)* **TransATAC**	627 149	RS 0-26 RS 27-100	10	89.1 (86.2-91.5) 75.2 (67.1-81.7) HR=2.75 (1.80-4.19), *p*<0.0001	NR	NR

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Nitz et al, 2019[239] (LN-mixed) WSG-Plan B	 649 231 657 254	EC-T RS 0-25 RS 26-100 TC RS 0-25 RS 26-100		NR	 94.9 84.0 94.4 85.8	NR
Collin et al, 2019[238] (LN-mixed) Georgia Cancer Registry	2,580 1,469 369 686 456 190 NR (84%) NR (15%) NR (78%) NR (22%)	 Ethnicity: White; RS: 0-17 Ethnicity: White; RS: 18-30 Ethnicity: White; RS: 31-100 Ethnicity: Black; RS: 0-17 Ethnicity: Black; RS: 18-30 Ethnicity: Black; RS: 31-100 Ethnicity: White; RS: 0-25 Ethnicty: White; RS: 26-100 Ethnicity: Black; RS: 0-25 Ethnicity: Black; RS: 26-100	10	NR	BCSS Reference category aHR=3.93 (2.32-6.66) aHR=4.98 (2.51-9.86) aHR=2.57 (1.27-5.19) aHR=.6.35 (3.45-11.68) aHR=10.82 (5.51-21.25) BCSS Reference category aHR=3.52 (2.23-5.55) aHR=2.08 (1.29-3.33) aHR=6.32 (3.84-10.43)	NR
MammaPrint®						
Opdam et al, 2022[189]	7 26	MPI > 0.355; 0 < MPI ≤ 0.355;	10	 69 (40-100) 78 (61-100)	RFI 69 (40-100); 79 (63-100)	NR

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
IKA	22 7 26 22	MPI ≤ 0 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0		53 (35-79) *p*=0.047	42 (25-70) *p*=0.017 BCSS 83 (58-100) 96 (88-100) 57 (39-83) *p*=0.01	
	7 26 22 7 26 22	MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0	15	69 (40-100) 78 (61-100) 45 (27-75) *p*=0.047	RFI 69 (40-100); 79 (63-100) 42 (25-70) *p*=0.017 BCSS 83 (58-100) 88 (73-100) 42 (24-74) *p*=0.01	NR
	7 26 22 7 26 22	MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0 MPI > 0.355; 0 < MPI ≤ 0.355; MPI ≤ 0	20	69 (40-100) 78 (61-100) 45 (27-75) *p*=0.047	RFI 69 (40-100); 79 (63-100) 42 (25-70) *p*=0.017 BCSS 83 (58-100) 78 (60-100) 42 (24-74) *p*=0.01	NR
Dubsky et al, 2021[240] (LN-mixed) ABCSG-8	512 146 512 146	Low Risk High Risk Low Risk High Risk	5 10	94% 91.6% aHR=13.3 (1.92-92.7), *p*=0.0088 Low Risk 91.3% High Risk 84.8%;	NR	NR

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
				aHR=1.28 (0.69-2.35), p=0.4368		
Lopes Cardozo et al, 2022 (LN-mixed)[**(241)]	1,000 3,295 2,398	Ultralow: MPI > 0.355; Low: 0 < MPI ≤ 0.355; High: MPI ≤ 0	8	97.0 (95.8-98.1) 94.5 (93.6-95.3) 89.2 (87.9-90.5) *p*<0.0001 Ultralow vs Low aHR=1.54 (1.06 to 2.22) High vs Low aHR=2.17 (1.68-2.80)	BCSS 99.6 (99.1-100) 98.2 (97.7-98.7) 93.7 (92.6-94.7) *p*<0.0001	NR
Mook et al (2008)[(179)] Cohort study	99 142 99 142	Low risk High risk Low risk High risk	5 10	98% 80% 91% 76% *p*=0.001 HR=4.13 (1.72–9.96)	BCSS 99% 88% 96% 76% *p*<0.001 HR=5.70 (2.01–16.23)	 NR NR NR NR HR=5.40 (2.11–13.80); *p*<0.001
Prosigna®						
Pu et al, 2020[(190)]* WHEL	551	Low Medium High Low Medium High	10	NR	BCSS 90% 84% 77% *p*=0.003 DFS 81%; 64%; 56% *p*=0.02	NR

Author, year, study	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Fitzal et al, 2021[244] (LN-mixed) ABCSG-8	765 269	<57 ≥57	5	NR	Local recurrence 99.9% (99.3-100) 97.8% (95.4-99.1) SHR=17.18 (2.06-142.88), *p*=0.009	NR
	765 269	<57 ≥57	10		Local recurrence 99.1% (98.0-99.6) 96.2% (93.4-98.1) SHR=3.70 (1.31-10.41), *p*=0.013	

Key: aHR=adjusted hazard ratio; BCSS=breast cancer specific survival; CI=confidence interval; DFS=disease-free survival; EC-T=epirubicin and cyclophosphamide followed by docetaxel; GEP=gene expression profiling; HR=hazard ratio; LN+=lymph node positive; LRR=locoregional recurrence rates; TC=docetaxel and cyclophosphamide; RFI=recurrence free interval; DFS=disease-free survival; SHR=subhazard ratio; SWOG58814=Southwest Oncology group study 58814; SEER=Surveillance, Epidemiology, and End Results Program; TransATAC=The translational study of Arimidex, Tamoxifen, Alone or in Combination cohort; WSG-PlanB=West German Study PlanB Trial; IKA=Integraal Kankercentrum Amsterdam ; ABCSG-8=Austrian Breast and Colorectal Cancer Study Group trial 8; WHEL=Women's Healthy Eating and Living Study; SCAN-B=Sweden Cancerome Analysis Network Breast.

*Probability of distant recurrence, disease recurrence, or death were reported for this study and converted to freedom from these outcomes for presentation in this table

**HRs are inverted

*** Not Stage IA=T2N0 grade 3 or grade 1-2 and RS >11; T3N0 grade 2-3; T2N1mi grade 2-3; T3N1mi grade 1-3 ; T1N1 grade 3; T2N1 grade 2-3; T3N1 grade 1-3; T1-3N2-3 grade 1-3

4.3.5.2 Predictive studies

In total, nine studies evaluated the predictive ability of GEP tests (Oncotype DX®: n=7; MammaPrint®: n=2) in LN+ patients. Five studies were identified for LN mixed patients (OncotypeDX®: n=2; MammaPrint®=1; Prosigna®: n=1; EndoPredict®: n=1). Data from the seven studies identified in the current review are presented in Table 4.6. Two RCTs, RxPONDER for Oncotype DX® (one study) and MINDACT for MammaPrint® (two studies) evaluated the predictive ability of GEP tests and the remaining six studies were nonrandomised retrospective analyses of RCTs or databases.

Oncotype DX®

Seven studies considered the predictive accuracy of Oncotype DX® in LN+ patients, three from the Ontario Health review[71, 196, 211] and four from the updated search.[70, 214, 216, 245] Additionally there were two studies that considered the predictive accuracy of Oncotype DX® in mixed LN status patients, both identified from the updated search.

In the Ontario Health review, three studies evaluating the predictive ability of Oncotype DX® in LN+ patients were identified; two retrospective analyses of databases (National Cancer Database[211] and Clalit Health Services database[196]) and one of the SWOG-8814 trial.[71] As these studies are nonrandomised and do not control who receives chemotherapy and without knowing why those who received chemotherapy did so, these studies are severely limited. Ibraheem at al.[211] found significantly higher 5-year overall survival for patients in three risk groups (that is, RS 11-17, RS 18-25, and RS 26-30) who underwent chemotherapy (between groups $p<0.001$). Stemmer et al.[196] reported higher overall survival for patients in each of three risk groups (that is, RS 0-17, RS 18-30, and RS 31-30) who underwent chemotherapy but did not report the significance of their findings. In the study of the SWOG-8814[71] a significant positive association was found between a continuous RS score and the risk of recurrence in years one to ten among patients undergoing chemotherapy (HR=2.64, 1.33-5.27, $p=0.006$).

One RCT (the RxPONDER trial)[70] and three nonrandomised retrospective analyses of the National Cancer Database[214, 216, 245] were identified in the updated search. The RxPONDER trial was a randomised comparison of chemoendocrine therapy versus endocrine therapy alone among HR+, HER2−, and LN+ women with an RS 0-25. It aimed to assess whether the risk of disease recurrence increased with higher RS values and whether the benefit of chemotherapy relative to endocrine therapy also increased with a higher recurrence score. Although the Oncotype DX® RS score was positively associated with disease-free survival (HR=1.05, 95%CI=1.04 to 1.07; $p<0.001$), Oncotype DX® did not significantly predict any relative benefit of

chemotherapy with respect to invasive disease–free survival (p=0.35). That means that among LN+ women with RS 0-25, Oncotype DX® did not significantly predict any relative benefit of chemotherapy with respect to invasive disease–free survival. In the total population, five-year distant relapse-free survival (94.9% versus 93.9%, p=0.89) and five-year disease-free survival (92.2% versus 91.0%, p=0.10) did not differ between the chemoendocrine and endocrine therapy alone groups (see Table 4.6). In a pre-specified analysis, differences by menopausal status were examined. Among postmenopausal patients, five-year distant relapse-free survival and disease-free survival did not differ between the chemoendocrine and endocrine therapy alone groups (distant relapse-free survival: 94.4% vs 94.4%; aHR=0.95, 0.73-1.23; p=0.70; disease-free survival: 91.3% vs 91.9%; aHR=0.98, 0.79-1.22; p=0.89). However, among premenopausal patients, those who underwent chemoendocrine therapy had significantly better outcomes (distant relapse-free survival: 96.1% vs 92.8%; aHR=1.72, 1.15-2.56; p=0.009; disease-free survival: 93.9% vs 89.0%; aHR=1.67, 1.20-2.33; p=0.002). There were more Grade 3, 4, and 5 adverse events in the chemoendocrine therapy arm (Grade 3 n=385; Grade 4 n=123; Grade 5 n=4) compared to endocrine therapy alone (Grade 3 n=114; Grade 4 n=1; Grade 5 n=2). These findings are derived from the first five years of data of a planned 15-year follow-up. It is possible that findings may change when longer-term results are available, as occurred with MINDACT.[100, 177]

A limitation of the RxPONDER trial is that it did not assess the comparative effectiveness of Oncotype DX® with alternative approaches (for example, clinical risk assessment). As a result, it is unknown whether outcomes differ between patients who use Oncotype DX® or alternative approaches to guide adjuvant chemotherapy use. Additionally, whether all patients would have received adjuvant chemotherapy in the absence of Oncotype DX® testing is unclear, as 12% of participants were classed as having low clinical risk (that is, tumour size <2cm and Grade 1, using a modified version of Adjuvant! Online). Three nonrandomised retrospective analyses of the National Cancer Database found that within intermediate and high risk categories (determined by various cut-off scores) chemotherapy was significantly associated with better overall survival (see Table 4.6).[214, 216, 245]

Two studies reported on LN-mixed status[170, 215] patients and reported on disease-free survival and overall survival. Disease-free survival was specified as breast cancer specific survival, no significant different was seen between patients receiving adjuvant chemotherapy or endocrine therapy only for low clinical risk and high RS score (HR=1.37 (0.27-6.95), p=0.71) or for between high clinical risk and low RS score (HR=0.42 (0.12-2.32), p=0.39).[170] Overall survival was reported for LN-mixed patients both male and female patients by Wang et al.[215] Women, but not men, with RS 18-30 who underwent chemotherapy had significantly higher five-year

overall survival. Five-year overall survival did not differ by treatment among men or women with RS 11-25.[215]

MammaPrint®

Two studies, both based on the MINDACT trial, considered the predictive accuracy of MammaPrint® in LN+ patients; one study was identified from the Ontario Health review and one from the updated search.[100, 177] One study, identified in the manufacturer submission, considered the predictive accuracy of MammaPrint® in mixed LN status patients.[178]

Only the MINDACT trial evaluated the predictive ability of MammaPrint® in LN+ patients. Five-year follow-up is presented by Cordoso et al.[100] in the only study identified in the Ontario Health review, and five- and nine-year follow ups are presented by Piccart et al. in the only study identified in the updated search. A description of the MINDACT trial is presented in Section 4.3.3. Briefly, it is an open-label non-inferiority trial designed to evaluate whether the addition of MammaPrint® genomic testing to standard clinical practice is beneficial in selecting patients for adjuvant chemotherapy. A total of 6,693 patients were categorised based on their genomic and clinical risk (assessed using Adjuvant! Online). Patients with discordant results were randomly assigned to the adjuvant chemotherapy group or the no adjuvant chemotherapy group and patients with concordant risk were recommended to receive chemotherapy (high clinical/high genomic risk patients) or no chemotherapy (low clinical/low genomic risk patients).

The subgroup of LN+ patients in the low clinical risk and high genomic risk group was too small to analyse (n=15). In the high clinical risk and low genomic risk group, there was no significant difference in distant metastasis-free survival after five years between LN+ patients who did or did not receive chemotherapy (aHR=0.88, 0.42-1.82, p=0.724).[100] Piccart et al.[177] reported that patients who received chemotherapy had a statistically nonsignificant but slightly higher distant metastasis-free survival at eight-year follow-up (91.2% vs 89.9%, aHR=1.19, 0.73-1.96) but no difference was seen at five-years (96.0% vs 95.9%). Additionally, disease-free survival was slightly higher among patients who received chemotherapy (five-year: 92.7% vs 91.0%; eight-year: 85.3% vs 82.8%) but differences in overall survival rates were not observed (five-year: 98.4% vs 98.8%; eight-year: 95.5% vs 94.9%). However, these between group differences were not tested for statistical significance.

A predefined exploratory analysis of patients in the high clinical risk and low genomic risk group examined distant metastasis-free survival with and without chemotherapy among women who were aged 50 years and under versus those aged over 50 years.[177] In women aged over 50 years, no difference in distant metastasis-free survival with and without chemotherapy was seen at eight years (chemotherapy

(n=441): 90.2%; no chemotherapy (n=453): 90.0%; aHR 1.22, 0.81-1.82). In women aged 50 years and under, eight-year distant metastasis-free survival was higher with chemotherapy than without (chemotherapy (n=235): 93.6%; no chemotherapy (n=229): 88.6%; aHR=1.85, 1.02-3.33). Limitations in the MINDACT trial include the open label nature of the study, changes in patient eligibility requirements, and dropout rates (22%, n=481 of 2,187 patients who underwent randomisation). Additionally, molecular diagnostic testing was performed on a frozen sample of the resected tumour. This is in contrast with current practice in Ireland whereby a FFPE tissue sample is used. This may limit the generalisability of the findings to the Irish context.

One study of mixed LN status patients, a retrospective analysis of a pooled database from seven previously reported studies, supported the predictive ability of MammaPrint®.[178] However, as patients were not randomised to treatment groups, this study is at a high risk of confounding.

EndoPredict®

No studies were identified in the Ontario Health review or the updated search that evaluated the predictive ability of EndoPredict® in LN+ patients. However, one study that considered the predictive accuracy of EndoPredict® in mixed LN status patients identified from the Ontario Health review.

One study identified in the Ontario Health review indirectly evaluated the predictive ability of EndoPredict® in mixed LN status patients (n=3,746).[246] In this retrospective analysis of five RCTs (GEICAM/9906, GEICAM 2003/02, ABCSG-6, ABCSG-8, and TransATAC cohort derived from ATAC trial), the rate of increase in ten-year distant recurrence and disease recurrence with continuous EPclin score was significantly lower in women who received chemotherapy and endocrine therapy versus endocrine therapy alone (distant recurrence: p=0.022; disease recurrence: p=0.025).

Prosigna®

No studies were identified in the Ontario Health review or the updated search that evaluated the predictive ability of Prosigna® in LN+ patients. However, there was one study that considered the predictive accuracy of Prosigna® in mixed LN status patients identified from the Ontario Health review. However, it is important to note that Prosigna® is not indicated for the prediction of benefit associated with chemotherapy.

One study identified in the Ontario Health review evaluated the predictive ability of Prosigna® in mixed LN status patients (n=460).[243] In this retrospective analysis of

the DBCG 77B trial, the association of continuous Prosigna® ROR with disease-free survival and overall survival did not significantly differ between patients receiving chemotherapy or not (disease-free survival: p=0.32; overall survival: p=0.66). However, the molecular subtypes identified by Prosigna® (that is, luminal A, luminal B, basal-like, and HER2-enriched) could predict chemotherapy benefit, but only among ROR score-defined high-risk patients, not low-risk patients.

Table 4.6 Predictive ability of GEP tests among LN+ and mixed LN status patients: studies from the updated search

Author, year Study or database name	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
Oncotype DX®						
Kalinsky et al, 2021[70] RxPONDER**	2,487 2,497	RS ≤ 25 / CET RS ≤ 25 / ET	5	94.9% 93.9% aHR=1.14 (0.92-1.41), *p*=0.25	92.2% 91.0% aHR=1.16 (0.97-1.39), *p*=0.10	NR
Weiser et al, 2021a[216] National Cancer Database	NR	RS 0-10 / CT RS 0-10 / No CT RS 11-25 / CT RS 11-25 / No CT RS 26-100 / CT RS 26-100 / No CT	5	NR	NR	98.6% 96.9% *p*<0.05 97.9% 96.1% *p*<0.05 93.6% 82.7% *p*<0.05
Weiser et al, 2021b[245] National Cancer Database	12,916	RS 0-11 / CT RS 0-11 / No CT RS 12-17 / CT RS 12-17 / No CT RS 18-100 / CT RS 18-100 / No CT RS 12-25 / CT RS 12-25 / No CT	5	NR	NR	HR=NR, *p*=0.44; HR=3.04 (1.78-5.21), *p*<0.001 HR=2.02 (1.42-2.87), *p*<0.001 aHR=1.91 (1.42-2.57), *p*=NR
Iorgulescu et al, 2019[214] National Cancer Database	213 608 765 423 641 85	RS 0-17 / CT RS 0-17 / No CT RS 18-30 / CT RS 18-30 / No CT RS 31-100 / CT RS 31-100 / No CT	5	NR	NR	93.0 (80.7-97.6) 92.0 (86.9-95.2) *p*=0.27 93.2 (88.4-96.1) 85.7 (77.5-91.1) *p*=0.02 92.4 (88.1-95.2) 66.9 (48.6-79.9) *p*<0.001

Author, year Study or database name	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
	5,697 1,183	No RS / CT No RS / No CT				90.7 (89.5-91.8) 70.6 (66.5-74.3) *p*<0.001
Wang et al, 2020[215] (LN-mixed) National Cancer Database**	Men 164 89 329 67 Women 21,460 14,217 54,902 12,750	 RS 18-30 / No CT RS 18-30 / CT RS 11-25 / No CT RS 11-25 / CT RS 18-30 / No CT RS 18-30 / CT RS 11-25 / No CT RS 11-25 / CT	5	NR	NR	 CT vs No CT: aHR=0.73 (0.17-3.13) CT vs No CT: aHR=0.29 (0.08-1.02) CT vs No CT: aHR=1.20 (1.03-1.41) CT vs No CT: aHR=1.12 (0.95-1.32)
Gao et al, 2020[170] (LN-mixed) SEER Database	 451 311 405 1,791	 C-low / RS-high / CT C-low / RS-high / ET C-high / RS-low / CT C-high / RS-low / ET	2.9	NR	BCSS 99.2% 99.0% HR=1.365 (0.268-6.953), p=0.708 99.5% 99.6% HR=0.418 (0.116-2.32), p=0.391	NR
MammaPrint®						
Piccart et al, 2021[177] MINDACT	326 332 326 332	C-high / G-low / CET C-high / G-low / ET C-high / G-low / CET C-high / G-low / ET	5 8	96.0 (93.1-97.7) 95.9 (93.1-97.6) 91.2 (87.2-94.0) 89.9 (85.8-92.8)	92.7 (89.1-95.1) 91.0 (87.3-93.6) 85.3 (80.6-88.9) 82.8 (78.0-86.6)	98.4 (96.1-99.3) 98.8 (96.7-99.5) 95.5 (92.4-97.4) 94.9 (91.7-96.9)
Knauer et al (2010)[178](LN-mixed) Pooled database**	78 174	Low risk / CET Low risk / ET	5	99% 93% HR=3.85 (0.50-33.33); *p*=0.20	BCSS 99% 97% HR=1.72 (0.20-14.29); *p*=0.62	NR

Author, year Study or database name	Sample size	Risk category	Time period, years	Freedom from distant recurrence, % (95% CI)	Disease-free survival, % (95% CI)	Overall survival, % (95% CI)
	148 141	High risk / CET High risk / ET		88% 76% HR=2.86 (1.41-5.88); *p*<0.01	94% 81% HR=4.76 (1.69-14.29); *p*<0.01	

Key: BCSS=breast cancer specific survival; CET=chemotherapy plus endocrine therapy; CI=confidence interval; ET=endocrine therapy GEP=gene expression profiling; LN+=lymph node positive; C-high=high clinical risk; G-low=low genomic risk. RxPONDER=Clinical Trial Rx for Positive Node, Endocrine Responsive Breast Cancer; SEER=Surveillance, Epidemiology, and End Results Program; MINDACT=Microarray In Node negative Disease may Avoid ChemoTherapy.
*Probability of distant recurrence, disease recurrence, or death were reported for this study and converted to freedom from these outcomes for presentation in this table
**HRs are inverted

4.3.5.3 Decision impact studies

In total, five European studies evaluated the impact of GEP tests on treatment recommendations (Oncotype DX®: n=5; MammaPrint®: n=0, EndoPredict®: n=0; Prosigna®: n=0) in LN+ patients. Note, these studies primarily assessed pre-test recommendations prospectively and did not include follow-up of clinical outcomes such as recurrence or survival. Changes in treatment recommendations among LN+ patients in studies identified in the current review are presented in Table 4.7. These results are summarised below along with findings from the Ontario Health review. Twelve European studies reported in the Ontario Health review examined the impact of GEP tests on treatment recommendations in mixed LN status patient groups (Oncotype DX®: n=6; MammaPrint®: n=2; Prosigna®: n=0; EndoPredict®: n=4). These are also reported below. No studies that directly compared the impact on treatment decisions across tests. As decision impact studies included in the current review were limited to those conducted in Europe, due to expected geographical differences in chemotherapy uptake rate (see Section 4.2.3 'inclusion and exclusion criteria'), only decision impact studies included in the Ontario Health review that were performed in a European setting are discussed here.

Oncotype DX®

Five studies considered the impact of Oncotype DX® on treatment recommendations for LN+ patient, two from the Ontario Health review and three from the updated search. Additionally there were six studies that considered the predictive accuracy of Oncotype DX® in mixed LN status patients, all of which were from the Ontario Health review.

In the Ontario Health review two European studies evaluated the change in treatment recommendations following use of Oncotype DX® in LN+ populations. Across the European studies,[220, 221] sample sizes were 122 to 126 and all patients were HR+ and HER2-. The proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy were 4% and 9%, 2) from chemotherapy to no chemotherapy were 16% and 28%, or 3) at all were 20% and 39%. In their risk of bias assessment (ROBANS), in both studies it was unclear whether the outcome assessment was blinded and both studies were at a high risk of bias for participant selection and confounding. One of the studies[220] was also at a high risk of bias due to incomplete outcome data and one[221] was at a high risk of bias due to selective outcome reporting.

Additionally, in the Ontario Health review, six European studies evaluated the change in treatment recommendations following use of Oncotype DX® in mixed LN status populations. Across the studies,[219, 247-251] sample sizes ranged from 50 to 882 and all patients were HR+ and HER2- except for one study[250] in which 99% of

patients were HER2-. The proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy ranged between 0% and 13%, 2) from chemotherapy to no chemotherapy ranged between 10% and 69%, or 3) at all ranged between 16% and 69%. In their risk of bias assessment (ROBANS), it was unclear in five studies whether the outcome assessment was blinded,[219, 247, 248, 250, 251] four studies were at a high risk of bias for each of participant selection[219, 248-250] and confounding.[247-250]

Three new studies, all from Italy, were identified.[222-224] Sample sizes ranged from 99 to 414 and patients were exclusively LN+. All patients were HR+/HER2-. In the largest study,[224] chemotherapy recommendations reduced from 258 to 110 following use of Oncotype DX®. Although neither the exact number of patients whose treatment recommendations changed from no chemotherapy to chemotherapy or vice versa nor any breakdowns by risk groups were reported. Across the other two studies, the proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy were 0% and 3%, 2) from chemotherapy to no chemotherapy were 18% and 30%, or 3) at all were 18% and 33%. One study[222] reported changes by risk subgroup (RS 0-17, RS 18-30, RS 31-100), with no changes in any risk group from no chemotherapy to chemotherapy. All changes from chemotherapy to no chemotherapy were in the RS 0-17 (14/71 (19.7%)) and RS 18-30 groups (9/48 (18.8%)). All three studies were at a high risk of bias for confounding and two studies[222-224] were at risk of bias for participant selection.

MammaPrint®

No studies examined change in treatment recommendations following use of MammaPrint® in LN+ populations, however two studies both from the Ontario Health review considered the impact in mixed LN status populations.

In the Ontario Health review two European studies[252, 253] evaluated the change in treatment recommendations following use of MammaPrint® in mixed LN status populations. In Cusumano et al.,[252] there were 194 patients of whom 86% were ER+ and 88% were HER2-. In Wuerstlein et al.,[253] of the 430 patients 100% were HR+ and 99% were HER2-. The proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy was 15% and 24%, 2) from chemotherapy to no chemotherapy was 17% and 37%, or 3) at all was 32% and 61%. In their risk of bias assessment (ROBANS), it was unclear in both studies whether the outcome assessment was blinded. Wuerstlein et al.[253] was at a high risk of bias for confounding and selective outcome reporting while Cusumano et al.[252] was at a high risk of bias for participant selection.

No new studies were identified.

EndoPredict®

No studies examined change in treatment recommendations following use of EndoPredict® in LN+ populations, however four studies all from the Ontario Health review considered the impact in mixed LN status populations.

In the Ontario Health review, four studies, all of which were European, with three from the UK and one from Germany, evaluated the change in treatment recommendations following the use of EndoPredict® in mixed LN status populations.[254-257] Sample sizes ranged from 120 to 395, and all patients were HR+ and HER2-. The proportion of people whose recommendations changed 1) from no chemotherapy to chemotherapy ranged between 8% and 38%, 2) from chemotherapy to no chemotherapy ranged between 5% and 35%, or 3) at all ranged between 14% and 43%. In their risk of bias assessment (RoBANS), for two studies[256, 257] it was unclear whether the outcome assessment was blinded, both studies were at a high risk of bias for selective outcome reporting, and one study was also at a high risk of bias for confounding and participant selection.[256] For the other two studies (PROBAST), one[255] was at a high risk of bias for participant selection and had an unclear risk of bias for confounding, while the other[254] was at a high risk of bias for confounding.

No new studies were identified.

Prosigna®

No studies were identified from the Ontario Health review or in the updated search that evaluated the change in treatment recommendations following use of Prosigna® in a LN+ or mixed LN status population.

Table 4.7. Changes in treatment recommendations among LN+ patients: studies from the updated search

Author, year	Country	No CT to CT, n (%)	CT to no CT, n (%)	Total treatment change, n (%)
Oncotype DX®				
Zambelli et al, 2020[222]	Italy	Total population: 0/127 (0.0%) RS 0-17: 0/71 (0.0%) RS 18-30: 0/48 (0.0%) RS 31-100: 0/8 (0.0%)	Total population: 23/127 (18.1%) RS 0-17: 14/71 (19.7%) RS 18-30: 9/48 (18.8%) RS 31-100: 0/8 (0.0%)	Total population: 23/127 (18.1%) RS 0-17: 14/71 (19.7%) RS 18-30: 9/48 (18.8%) RS 31-100: 0/8 (0.0%)
Dieci et al, 2019[223]	Italy	3/99 (3.0%)	30/99 (30.3%)	33/99 (33.3%)
Cognetti et al, 2021[224]	Italy	NR	NR	CT recommendations were reduced from n=258/414 to n=110/414

Key: CT- chemotherapy; LN+-lymph node positive; NR - not reported; RS - recurrence score

4.4 Discussion

This discussion firstly summarises and considers the evidence for the prognostic ability, predictive ability and decision impact of the four GEP tests. This is followed by a consideration of the concordance between the GEP tests, the limitations of the evidence base and implications for future research, before a final conclusion is presented.

4.4.1 Evidence for the prognostic ability of GEP tests

This section considers 1) the evidence on whether the GEP tests do have prognostic value, 2) how the prognostic abilities of GEP tests compare with each other, and 3) the extent to which each test adds prognostic value beyond that of other prognostic information available to clinicians and patients.

The prognostic ability of GEP tests refers to their ability to provide an estimate of a patient's likely future risk of cancer recurrence and or survival after surgery. Typically, large observational studies provide the highest certainty of evidence regarding prognostic factors as RCT eligibility criteria often exclude patients relevant for the assessment of prognostic factors and eligible patients may decline to participate in RCTs for reasons related to their prognosis.[175] In this review, evidence for the prognostic ability of the tests came from retrospective analyses of various RCTs and data registries. Oncotype DX® was the most studied test in both LN- and LN+ populations.

The prognostic ability of each test within LN- and LN+ patients was largely supported, and the tests likely have modest prognostic value. However, due to the high heterogeneity across studies, each test's ability to predict cancer recurrence and or survival could not be meaningfully quantified. Considerable variation across study designs, analytic approaches, cut-off scores used within GEP tests, outcomes, and study populations precludes precise conclusions on the prognostic ability of each test. Further, some studies suggested that the prognostic ability of tests may vary between pre- and post-menopausal women[67, 212] (often defined using 50 years of age as a cut-off) and perform better among white compared with African American patients;[188] however, these differences were not examined for all tests and the evidence was limited.

Few studies directly compared GEP tests. Of those that did, they comprised retrospective analyses of trials designed for other purposes and only prognostic performance was compared, not predictive or decision impact. Among LN- patients, four studies found EndoPredict's® EPClin score and Prosigna® to be more prognostic for distant recurrence than Oncotype DX®.[183, 191-193] Among LN+ patients, three studies found EndoPredict's® EPClin score and Prosigna® to be more prognostic for

distant recurrence than Oncotype DX®[183, 192, 193] and two studies found EndoPredict® to be more prognostic for distant recurrence than Prosigna®.[183, 230] None of these studies compared all four GEP tests, although one study based on the TransATAC cohort included three (EndoPredict®, Oncotype DX®, and Prosigna®).[183] On the basis of ranking the strength of prediction for each individual test, this study concluded that Prosigna® and EndoPredict's® EPClin were more prognostic for overall and late distant recurrence than Oncotype DX® among LN- and LN+ women, although limited independent information was available from any test among LN+ women. The GRADE certainty of the comparative evidence was judged to be Very Low or Low due to imprecision and a high risk of bias.

Few studies examined whether Oncotype DX®, EndoPredict®, and Prosigna® add prognostic information beyond that of other prognostic information available to clinicians and patients, and none examined MammaPrint®. Two retrospective analyses of trial data found that Oncotype DX®,[192, 193] Prosigna®,[193] and EndoPredict's® EPclin score[192] added significant prognostic information beyond clinical parameters. A TransATAC study by Sestak et al.[183] found that each of Prosigna®, EndoPredict's® EPclin score and Oncotype DX® were significantly more prognostic, based on 0-10 year distant recurrence data in LN- patients, than a clinical treatment score comprising nodal status, tumour size, grade, age, and endocrine treatment. Among LN+ patients, the prognostic performance of these three tests beyond the clinical treatment score was significant but weaker than that observed for LN- patients.[183] However, when considering late distant recurrence (5-10 year follow-up), Prosigna®, and EndoPredict® (EPclin), but not Oncotype DX®, were found to provide significant additional prognostic performance in LN- patients, while only EndoPredict® (EPclin) added significant (though limited) prognostic information in LN+ patients.

A TAILORx follow-up study by Sparano et al.[67] reported that clinical risk (assessed by tumour size and grade) adds prognostic information to Oncotype DX®, but did not assess whether Oncotype DX® adds prognostic information to clinical risk.

Across all studies examined in this review, evidence that these tests add prognostic information beyond clinical risk was stronger in LN- patients than LN+ patients. Therefore, it may be concluded from the evidence examined that the GEP tests likely have modest prognostic value. Direct comparisons across tests were sparse and of low certainty. High heterogeneity across studies precludes indirect comparisons of tests. GEP tests may add prognostic value beyond that of other prognostic information available to clinicians and patients, although to what extent is unclear.

4.4.2 Evidence for the predictive ability of GEP tests

This section considers the evidence for the predictive ability of GEP tests, that is, the ability of these tests to identify people who may be more likely to benefit from chemotherapy. The strongest evidence comes from the TAILORx, RxPONDER, and MINDACT RCTs. Evidence from nonrandomised studies exists and is reported in the results section (Section 4.3) but, as these studies do not control who receives chemotherapy and without knowing why those who received chemotherapy did so, these studies are severely limited. Therefore, only the three trials providing the strongest evidence are discussed here.

The optimal RCT would be designed to assess whether a GEP test performs better than other risk prediction methods. MINDACT compared cancer outcomes across patients with discordant clinical and genomic risk (that is, high clinical risk and low genomic risk and vice versa) who were randomised to chemoendocrine therapy or endocrine therapy alone. In doing so, it could assess whether patients benefited more from chemotherapy when it was indicated by clinical risk or by genomic risk. When clinical and genomic risk were in agreement, patients were assigned to the treatment indicated by both risk assessments. In the trial's primary five-year findings, patients in the high clinical risk and low genomic risk group had similar rates of distant metastasis-free survival regardless of whether they received chemotherapy, indicating that genomic risk was a better predictor of five-year chemotherapy outcome than clinical risk.[100] This suggests that, on average, these patients could be safely spared chemotherapy. These findings were replicated by Piccart et al. after eight years among LN+ patients; however, they found that LN- patients in the high clinical risk and low genomic risk group would benefit from chemotherapy.[177] Further, across both LN groups there was a trend, albeit statistically nonsignificant, that outcomes were better among patients who received chemotherapy. For example, after eight years in the total population (that is, LN- and LN+ patients combined) with high clinical risk and low genomic risk, the chemotherapy group had 25% lower deaths (4.3% vs 5.7%), 25% lower distant metastasis recurrence (8.0% vs 10.6%), and 20% lower disease recurrence (13.6% vs 17.1%) relative to the no chemotherapy group.[177] Additionally, the authors reported an underpowered exploratory analysis that indicated that LN+ women aged under 50 years with high clinical risk and low genomic risk may benefit more from chemotherapy, with the authors hypothesising that this is possibly due to ovarian function suppression induced by chemotherapy.[177] Based on this, it is plausible that an important subgroup or subgroups of the high clinical risk and low genomic risk group may benefit from chemotherapy. Finally, the subgroup of LN+ patients in the low clinical risk and high genomic risk group was too small to analyse and outcomes for LN- patients in this risk group did not differ between those who did and did not receive chemotherapy; based on this, there is no evidence that MammaPrint® can predict chemotherapy benefit among patients identified as low clinical risk.

Considering the overall findings of the MINDACT trial to date, there is therefore some evidence that MammaPrint® can help to spare chemotherapy among high clinical risk LN+ patients aged 50 years and over, although this is uncertain. The generalisability of these findings to the Irish setting is unclear as the molecular diagnostic testing in MINDACT was performed on frozen samples of the resected tumour whereas in practice currently in Ireland a FFPE tissue sample is used.

In contrast, both the TAILORx and RxPONDER trials report the relative performance of Oncotype DX® with itself and do not compare it with another risk prediction method. Therefore, the predictive performance of Oncotype DX® relative to any other risk prediction method has not been tested. Primary findings from the TAILORx trial indicated that there was no chemotherapy benefit among LN- women with RS 11-25.[66] The authors therefore concluded that this population can be spared chemotherapy, although they note that exploratory analysis indicated that there may be some chemotherapy benefit for women aged 50 years and under with an RS 16-25. However, a further analysis of the trial aimed to assess whether combining clinical risk (defined using a modified Adjuvant! Online algorithm) with Oncotype DX has further prognostic value beyond Oncotype DX® alone, thereby allowing an indirect comparison of Oncotype DX® with clinical risk for predicting chemotherapy benefit.[67] In this analysis, LN- women aged 50 years and under with high clinical risk and RS 11-25 (a population who would be spared chemotherapy based on the primary trial findings) had lower distance recurrence rates (6.1% vs 12.3%) when they received chemotherapy; however, this difference was not tested for statistical significance as it was not the primary aim of the paper. Based on this, it is plausible that there is an important subgroup or subgroups of the RS 11-25 group who may benefit from chemotherapy.

Two major limitations of the TAILORx trial are noted. Firstly, although all women in the study were candidates for chemotherapy, a notable proportion of those randomised to receive chemotherapy may not have received chemotherapy under an Irish treatment pathway. For example, 29% of women within the TAILORx trial had Grade 1 tumours which typically may not have been treated with chemotherapy in Ireland.[17] Additionally, 74% of women randomised to chemoendocrine therapy were classified as clinically low risk based on a modified Adjuvant! Online algorithm. It is relevant to note that prior to the introduction of Oncotype DX® for LN- women in Ireland in 2011, Adjuvant! Online was used by clinicians to predict the magnitude of benefit from chemotherapy. As some of the population who had chemotherapy in the trial would not be routinely offered chemotherapy in an Irish pathway, the generalisability of the TAILORx findings is unclear. Assigning participants who may not be expected to benefit from chemotherapy to the chemoendocrine therapy arm of the trial may have reduced the magnitude of the chemotherapy effect, thus

biasing the results towards finding non-inferiority of chemoendocrine therapy compared to endocrine therapy alone. Secondly, more women assigned to chemotherapy plus endocrine therapy did not complete the study protocol compared to those assigned to endocrine therapy alone (32% vs 17%). Of the women assigned to chemoendocrine therapy who did not complete the protocol, 55% did not receive chemotherapy. This unbalanced flow of participants may have diminished the difference in effects between the two trial arms, thus further biasing the results towards finding non-inferiority of chemoendocrine therapy compared to endocrine therapy alone.

Therefore, TAILORx suggests that 1) all LN- women aged over 50 years and 2), based on the exploratory analysis, LN- women aged 50 and under with an RS 11-15 may be spared chemotherapy; however, the trial had notable limitations and did not assess whether Oncotype DX® is a better predictor of chemotherapy benefit than clinical risk.

Primary findings from the RxPONDER trial indicated that premenopausal LN+ women with an Oncotype DX® RS score of 0-25, may benefit from chemotherapy, but not postmenopausal women.[70] Therefore, the authors concluded that postmenopausal LN+ women with an RS 0-25 can be spared chemotherapy. However, these findings are derived from the first five years of data of a planned 15-year follow-up. Findings may change when longer-term results are available, as occurred with MINDACT.[100, 177] Although it is likely most participants would have received adjuvant chemotherapy in an Irish pathway in the absence of Oncotype testing, it is unclear whether all of them would as 12% had low clinical risk (that is, tumour size <2cm and Grade 1). Additionally, as with TAILORx, whether Oncotype DX® performs better than other risk prediction methods was not assessed.

In summary, among LN- patients, evidence for the predictive ability of MammaPrint® from the MINDACT trial indicated that it does not offer predictive value beyond a modified Adjuvant! Online algorithm that uses data on ER status, HER2 status, nodal status, tumour grade, and tumour size. Evidence for the predictive ability of Oncotype DX® from the TAILORx trial had major limitations. The trial had no comparator for Onctoype DX®, meaning that whether it offers predictive value beyond a modified Adjuvant! Online algorithm, for example, was not assessed. Additionally, the unbalanced participant flow and participant selection likely biased results and limit generalizability to an Irish setting. Among LN+ patients, evidence for MammaPrint® from the MINDACT trial was uncertain and its generalisability was limited by the performance of molecular diagnostic testing on frozen samples of the resected tumour whereas in practice currently a FFPE tissue sample is used. The evidence for the predictive ability of Oncotype DX® from the RxPONDER trial was

limited by the lack of comparator for Onctoype DX® meaning, as with TAILORx, whether Onctoype DX® offers predictive value beyond a modified Adjuvant! Online algorithm, for example, was not assessed. Additionally, it is unclear whether all participants would have received adjuvant chemotherapy in an Irish pathway in the absence of Oncotype testing as 12% had low clinical risk (that is, tumour size <2cm and Grade 1); however, it is likely that most would have received chemotherapy. The findings therefore support that postmenopausal LN+ women with an Onctoype DX® RS 0-25 can be spared chemotherapy. No randomised controlled trials assessed the predictive abilities of EndoPredict® which is the only other test indicated as predictive.

4.4.3 Evidence for treatment recommendation changes due to GEP tests

This section considers the evidence for changes in treatment recommendations regarding chemotherapy after the use of a GEP test. Across tests there was generally a 20% to 50% change in treatment decisions from pre- to post-test administration; although evidence among LN+ specific populations was sparse. There were changes in both recommending chemotherapy, generally to those with high genomic risk scores, and in withholding chemotherapy, generally from those with low genomic risk scores. Risk of confounding was the most common limitation across these studies as factors contributing to treatment decision-making were not identified and therefore not controlled for in analyses. Evidence also suggests that Oncotype DX® has a substantial impact on treatment recommendations in Ireland.[3] Based on pre-Oncotype DX® treatment recommendations established by consensus through surveying lead breast medical oncologists from each of the eight cancer centres in Ireland, a pre-Oncotype DX® treatment recommendation was retrospectively assigned to 960 patients with HR+ early stage breast cancer whose tumour samples were tested with the 21-gene assay between October 2011 and February 2019. The change in treatment recommendation as a result of 21-gene testing was then calculated. In total, 63% of treatment recommendations changed, 62% of people were spared chemotherapy and 1% were recommended it. Overall, the evidence suggests that gene expression profiling tests can impact treatment recommendations; however, the impact of these changes on patient outcomes is unclear as studies did not collect follow-up data.

4.4.4 Concordance between GEP tests

Concordance between GEP tests refers to the degree to which the tests assign the same patients to the same risk groups. Relevant, high quality evidence on the concordance of tests is provided by the OPTIMA Prelim study, a feasibility phase of Optimal Personalised Treatment of early breast cancer usIng Multi-parameter

Analysis (OPTIMA), a UK trial that aims to test the effectiveness of multiparameter testing in identifying patients who can be spared chemotherapy.[258] OPTIMA Prelim was designed to help select which tests to include in the trial. Participants were 302 women aged 40 years and over with ER+, HER2– early-stage breast cancer with either 1–9 positive lymph nodes or a tumour of ≥ 30 mm if LN- (that is, women who would routinely be offered chemotherapy). The study compared three GEP tests of interest in the current review (Oncotype DX®, MammaPrint®, and Prosigna®), and three other tests (MammaTyper, NexCourse Breast by Aqua [IHC4-AQUA], and IHC4 by conventional immunohistochemistry). Oncotype DX®, MammaPrint®, and Prosigna® showed moderate agreement with each other when dichotomising results between high versus low/intermediate risk (kappa range: 0.40 to 0.53), with concordance higher in the low/intermediate risk groups than in the high risk groups. In comparing Oncotype DX® with Prosigna®, the two GEP tests with three risk categories (low, intermediate, and high), 29 (9.7%) tumours were classified as low risk by one and high risk by the other. Other findings comparing the three GEP tests specifically are not reported; however, across all six tests, disagreement in risk categorisation for 60.6% of tumours was observed. All tests agreed on a low/intermediate risk categorisation for 30.8% of tumours and a high risk categorisation for 8.6% of tumours. Although large differences were observed in the categorisation of individual patients, the proportions of patients identified as low, intermediate, or high risk were comparable across the three tests. This is perhaps unsurprising given the large variation in genes assessed by each test: Oncotype DX® shares just nine genes with Prosigna®, one with EndoPredict®, and one with MammaPrint®, MammaPrint® shares three genes with Prosigna®, and Prosigna® shares two genes with EndoPredict®.[259] Considering this, the observed discordance across tests suggests that there may be more than one way of predicting risk and that no test should be the ultimate discriminator of risk for individual patients.[258]

4.4.5 Limitations of the evidence base

There are a number of limitations throughout the evidence base. Most studies were not designed in a way that could assess whether GEP tests add prognostic or predictive information compared to routinely assessed clinicopathologic factors.

Many prognostic studies comprised retrospective analyses of RCTs and observational databases which carry their own risks of biases. In the retrospective analyses of databases, the selection of patients on the basis of them having received a test may have introduced bias as patients not offered a test may systematically differ from those who were. In the retrospective analyses of RCTs, RCT eligibility criteria often exclude patients relevant for the assessment of prognostic factors and eligible patients may decline to participate in RCTs for reasons related to their prognosis. Further, observational studies that limit their population to those who did not receive

chemotherapy may introduce bias as patients not given chemotherapy may systematically differ from those who were. Conversely, including patients who received chemotherapy may impact estimates of prognostic performance.

Considering decision impact studies, the generalisability of this evidence base to the Irish setting is limited as the use of chemotherapy differs across countries. Further, with no follow-up data, these studies simply conclude that these tests can influence treatment decisions. They do not provide evidence on whether these changes in treatment decisions were ultimately beneficial to patients.

Many studies, including the three key RCTs described within this report (MINDACT, TAILORx and RxPONDER), were funded by industry. This introduces a risk of bias that should be considered when interpreting the evidence base.

The generalisability of the three primary RCTs to an Irish context is also unclear. In TAILORx and RxPONDER it is likely that some of the population who were randomised to receive chemotherapy would not be routinely offered chemotherapy in an Irish pathway. In MINDACT, the method used to predict clinical risk (that is, a modified Adjuvant! Online algorithm) is no longer used in Ireland and molecular diagnostic testing was performed on frozen samples of the resected tumour, whereas current practice in Ireland is to use a FFPE tissue sample; this may introduce variability in gene expression profiles.[260] Finally, TAILORx had notable limitations fully detailed in Section 4.4.2 that likely influenced key findings.

4.4.6 Future research

Currently, the evidence is unclear on whether the use of GEP tests leads to improved patient care, such as avoiding unnecessary treatments or supporting the addition of treatment(s) when required, and better cancer-related outcomes. This is primarily because existing studies are not designed in a way that enabled them to provide this information. Future planned prospective studies should establish:

1. Does each GEP test offer additional prognostic and or predictive information compared to routinely assessed clinicopathologic factors?
2. How do the prognostic and or predictive abilities of the tests compare with each other?
3. Does each test's performance vary based on patient characteristics such as menopausal status?

Additionally, more research among populations that will likely benefit most from GEP tests is required. In an Irish context, nationally collected follow-up data on patients who use Oncotype DX® could evaluate the impact these tests are having on patient care and outcomes.

4.4.7 Conclusions

The four GEP tests described within this review likely have modest prognostic value; however, due to considerable variation across study designs, analytic approaches, risk cut-off scores used within GEP tests, outcomes, and study populations, each test's ability to predict cancer recurrence and or survival could not be meaningfully quantified within the body of evidence available thus far. Moreover, direct comparisons of tests were sparse, making it difficult to differentiate between the prognostic abilities of the tests and the extent to which each test adds prognostic value beyond that of other prognostic information available to clinicians and patients (that is, clinical and pathological information) is unclear.

The evidence on whether GEP tests have the ability to identify 1) patients who could safely be spared the addition of chemotherapy to endocrine therapy or 2) patients who would benefit from the addition of chemotherapy to endocrine therapy is limited. MammaPrint® was the only test for which predictive performance was compared to another risk prediction method (clinical risk as assessed by Adjuvant! Online). A low genomic risk score from MammaPrint® could potentially be used as an indicator to safely spare chemotherapy for high clinical risk LN+ patients aged 50 years and over, although this is uncertain. It is possible that Oncotype DX® can differentiate between patients (both LN- and LN+) who would or would not benefit from chemotherapy; however, this is uncertain, the evidence is limited, and its ability to do so relative to other risk prediction methods has not been assessed. There were no direct comparisons of the tests. Therefore, differentiating between the predictive abilities of the tests was not feasible.

Across all GEP tests, between approximately 20% and 50% of treatment decisions were observed to have changed from pre- to post-test administration, suggesting that GEP tests can impact treatment recommendations. It is important to note that these studies did not assess whether these changes in treatment recommendations led to improved patient outcomes. Although not examined in the systematic review, large differences in the categorisation of patients across tests have been observed at the individual level. However, despite differences in the individual level categorisation, the overall proportions of patients identified as low, intermediate, or high risk have been found to be comparable across tests.

5 Discussion

GEP tests are intended to provide information on disease prognosis and to predict the potential for benefit from adjuvant chemotherapy among patients with early-stage invasive breast cancer. In Ireland, the Oncotype DX® GEP test was first reimbursed by the HSE in 2011 for LN- patients and in 2019 for LN+ patients. Approximately 1,800 cases of HR+, HER2- stage I-IIIa breast cancer were diagnosed annually between 2015 and 2019. As per communication with the EAG, the majority of these patients receive Oncotype DX® testing (exact figure is commercially sensitive). Three other commercially available GEP tests, MammaPrint®, EndoPredict®, and Prosigna®, are not reimbursed by the HSE. The purpose of this rapid HTA is to provide advice to the HSE on these alternatives to Oncotype DX® in terms of their potential use to inform decision-making in relation to the management of early-stage invasive breast cancer.

5.1 Technological considerations

Surgery is considered the first-line treatment for most types of breast cancer. Following surgery, further ('adjuvant') therapy, including chemotherapy, may be needed to control disease in the breast, lymph nodes and surrounding areas to reduce the risk of recurrence and or metastasis. GEP tests may be used alongside clinicopathological factors (such as tumour size, disease stage and age) to inform decisions regarding the use of adjuvant chemotherapy. There is a large variation in the genes assessed by each test: Oncotype DX® (21 genes) shares nine genes with Prosigna® (50 genes), one with EndoPredict® (12 genes), and one with MammaPrint® (70 genes), MammaPrint® shares three genes with Prosigna®, and Prosigna® shares two genes with EndoPredict®.[259] All four GEP tests use formalin-fixed paraffin-embedded (FFPE) samples, which are routinely prepared during diagnostic testing. EndoPredict® and Prosigna® can be performed in local laboratories with the relevant equipment (platforms, assays, kits and reagents), while analysis of Mammaprint® and Oncotype DX® is limited to centralised laboratories in the US and additionally the Netherlands in the case of Mammaprint®. The anticipated test turnaround times range from 3 to 10 days following receipt of the sample at the relevant laboratory, although the total turnaround time can be longer due to preparation and transport time prior to laboratory receipt.

There are a number of organisational issues to consider in relation to the use of GEP tests for informing adjuvant chemotherapy decisions in breast cancer. Firstly, analysing alternative GEP tests in local Irish laboratories could have capacity implications for laboratory resources and workflows. The use of locally analysed tests could necessitate additional resources in terms of labour and equipment in

order to meet demand. However, it is also noted that it may be feasible to use other partner laboratories within the EU to avoid such challenges in local implementation.

Secondly, breast cancer is a heterogeneous disease which can differ greatly between patients and even within individual tumours from the same patient.[261, 262] Certain aspects of variability within and or between FFPE tumour blocks from the same patient can be addressed by careful dissection of the tumour sample; for example, manual microdissection may be required in cases where there are large amounts of non-tumour elements.[262] Accordingly, variations in the sampling method and preparation techniques used can influence the findings of GEP tests.[263] Best practice guidance and education of the relationship between this variability and the findings of GEP tests (including consideration of intrinsic subtypes) are currently lacking in the Irish context and production of such guidance to address this gap is warranted.

5.2 Clinical effectiveness

As outlined in Chapter 4, the clinical effectiveness evidence suggests that the prognostic abilities of the four GEP tests are modest, with greater consistency among LN- populations compared to LN+, aligning with the findings of previous HTAs.[29, 31, 44, 45, 90] In their 2011 advice to the HSE, the Irish Society of Medical Oncology reported that prognostic estimates may be optimised by combining Adjuvant! Online and Oncotype DX® and recommended that both decision tools be used when discussing treatment options with LN- patients.[9] However, the body of evidence examining the added prognostic value of the four GEP tests beyond clinicopathological information already available to clinicians and patients has remained small and limited since the 2011 recommendations. Adjuvant! Online, no longer available, was a free online tool that considered multiple clinical and pathological factors to produce estimates for recurrence and mortality; it was commonly used to provide prognostic and predictive information to clinicians and patients in 2011 when the HSE approved reimbursement of Oncotype DX® for use among LN- patients. Although Adjuvant! Online is not currently available, free web-based tools, such as the NHS Predict tool (link), that use clinical and pathological information to help clinicians and patients make decisions about treatment following surgery, remain available. The current review found GEP tests may add prognostic value to such tools, but to what extent they may add value is unclear due to limitations in the underlying evidence base.

Three of the tests, Oncotype DX®, MammaPrint®, and EndoPredict®, are indicated as predictive (that is, they are intended to predict patients who may benefit from chemotherapy). RCT data representing patients who were randomised to chemotherapy or to no chemotherapy provide the strongest predictive evidence.

This review also considered observational studies that reported outcomes among patients within certain risk categories who did and did not undergo chemotherapy; however, patients in these studies were not randomised to treatment groups and such observational studies are therefore at high risk of confounding. While no RCTs examined the predictive abilities of EndoPredict®, the evidence from three RCTs for the predictive abilities of Oncotype DX® and MammaPrint® was weak. Among LN- patients, evidence for the predictive ability of MammaPrint® from the MINDACT trial indicated that it does not offer predictive value beyond that of a modified Adjuvant! Online algorithm that used data on ER status, HER2 status, nodal status, tumour grade, and tumour size. Evidence for the predictive ability of Oncotype DX® in LN- patients from the TAILORx trial indicated that there was no chemotherapy benefit among LN- women with RS 11-25, although this was uncertain due to major limitations. Specifically, the unbalanced participant flow and participant selection in TAILORx likely biased results and limited generalisability to the Irish setting. Additionally, the trial had no comparator for Onctoype DX®, meaning that the performance of Oncotype DX® relative to another potential prognostic or predictive tool was not assessed. It therefore remains plausible that other tools may be as effective as Oncotype DX® in the total population or for specific subgroups while offering additional benefits in factors such as cost and feasibility of use. For example, one small study (n=46 patients) showed high concordance between Oncotype DX®, the NHS Predict tool, and the Nottingham Prognostic Index among low risk patients.[264] If such findings can be verified and replicated in Ireland, there may be minimal benefit to using Oncotype DX® among low clinical risk patients.

Among LN+ patients, findings from the MINDACT trial suggest that there may be merit in using MammaPrint® among high clinical risk (as assessed by a modified Adjuvant! Online algorithm) LN+ patients aged 50 years and over, as patients in this group with a low genomic risk score may be safely spared chemotherapy. However, this finding was uncertain and its generalisability to an Irish context is unclear as Adjuvant! Online is no longer available and, in practice, the molecular diagnostic testing is performed on an FFPE tissue sample rather than frozen samples of the resected tumour, as was done in MINDACT, which may introduce variability in gene expression profiles.[260] The RxPONDER trial supported the predictive ability of Oncotype DX® among LN+ patients, finding that postmenopausal LN+ women with an RS 0-25 can be safely spared chemotherapy. It is unclear whether all participants would have received adjuvant chemotherapy in an Irish pathway in the absence of Oncotype DX® testing as 12% had low clinical risk (that is, tumour size <2cm and Grade 1, using a modified version of Adjuvant! Online). However, it is likely that most would have received adjuvant chemotherapy. Additionally, as with TAILORx, RxPONDER was limited by the lack of a comparator for Onctoype DX®, meaning that

whether it offers predictive value beyond standard clinical and pathological information was not assessed.

Considering the findings of other assessments before the publication of RxPONDER findings, the Australian Medical Services Advisory Committee found that there was insufficient evidence to support using any GEP test to determine which patients could safely avoid chemotherapy.[90] Similarly, the Haute Autorité de Santé reported that there was insufficient evidence of clinical utility to recommend the routine use of GEP tests.[92] The current report notes many of the same limitations in the clinical effectiveness evidence base as were identified in these prior reports. However, conclusions across HTAs reviewing broadly the same evidence base have varied considerably. Some have supported the use of Oncotype DX® for guiding adjuvant chemotherapy decisions among LN- and LN+ patients.[44, 45, 88, 89] One concluded that the initial findings from the MINDACT trial[100] did not provide sufficient evidence to support withholding adjuvant chemotherapy based on MammaPrint® testing[87] while, based on the same trial, two HTAs concluded that women with high clinical risk early-stage invasive breast cancer may safely forego adjuvant chemotherapy if they have a low MammaPrint® risk score.[88, 89] These varying interpretations of broadly the same evidence base are likely due to issues such as study heterogeneity and flawed study designs that generally do not assess the relative prognostic and predictive value of GEP tests making it difficult to interpret the evidence.

5.3 Need for further data collection, research and guidance

In their 2011 advice to the HSE, the Irish Society of Medical Oncology recommended the establishment of a National Registry to collect clinical and pathological characteristics of patients whose breast cancer specimens were sent for Oncotype DX® testing so as to allow continuous assessment of how the test is used in Ireland and direct comparison to its use in other publicly-funded healthcare systems.[9] While certain databases have been established,[265] for example, the Breast Cancer Ireland and Royal College of Surgeons in Ireland Breast Cancer Patient Biobank, the NUI Galway Breast Cancer Biobank, and the TAILORx Breast Cancer Tissue Bank, a national registry, as recommended, was not established; therefore, currently, patient-level Oncotype DX® test data cannot be linked to treatment or patient outcomes in Ireland at a national level. Collecting data for women with HR+, HER2- breast cancer who are treated in Irish public hospitals and who receive Oncotype DX®, is vital in order to examine the relationships between test use, treatment, and patient outcomes in an Irish context. Such data would also be conducive to performing a meaningful cost-effectiveness analysis from the Irish healthcare perspective.

The need for data collection to provide transparent evidence for the clinical effectiveness of GEP tests has also been identified by NICE in the UK and Haute Authorité de Santé in France.[45, 92] NICE recommended EndoPredict®, Oncotype DX® and Prosigna® for reimbursement in the NHS to guide adjuvant chemotherapy decisions in ER+ HER2- LN- patients, conditional on the collection of data on test use within the NHS to address remaining uncertainty on their clinical impact. Similarly, Haute Authorité de Santé concluded that the evidence of clinical utility was insufficient to recommend routine use of GEP tests, noting that GEP tests are not intended to replace standard clinicopathlogical criteria. However, they recommended conditional funding of EndoPredict®, MammaPrint®, Oncotype DX® and Prosigna® with a view to collecting comparative prospective data in LN- patients who are at intermediate risk of recurrence based on clinicopathological factors and uncertain predicted benefit from chemotherapy.

GEP tests may help spare patients the potential side effects and late complications of chemotherapy, and reduce expenditure on both the chemotherapy itself and the treatment of these adverse effects, without increasing the patient's risk of future distant recurrence or cancer-related death. The evidence reviewed in the current rapid HTA attempts to clarify the risk of omitting chemotherapy; however, there is little information provided in the included evidence on the side effects of chemoendocrine therapy relative to endocrine therapy alone. Publishing data on chemotherapy toxicity and comparing adverse events between patients who underwent chemoendocrine and endocrine therapy alone in trials such as MINDACT and TAILORx would provide further clarity regarding the risks associated with undergoing chemotherapy.

Currently, in Ireland there is no guidance on the use of Oncotype DX®. Clear guidelines on which patients are indicated for the test and how to interpret test results could improve consistency in test use and may particularly benefit trainees or clinicians who do not focus exclusively on breast cancer. Further, tumour heterogeneity and variation in the sampling method and preparation techniques used can influence the findings of GEP tests.[263] Guidelines and education may be beneficial to ensure accuracy and consistency of biopsy selection.

Additional studies investigating the use of GEP tests in other clinical areas have also been conducted or are ongoing; these include studies regarding the use of GEP tests for guiding clinical decision-making regarding the use of neoadjuvant chemotherapy,[266] extended endocrine therapy,[29] and radiation therapy (ongoing trials: DEBRA, 04852887; EXPERT, 02889874; IDEAL, NCT 02400190; PRECISION, NCT2653755; TAILOR RT, NCT 03488693), and the use of these tests for different patient populations, such as patients with ductal carcinoma in situ.[267] The current

rapid HTA only examines the four commercially available GEP tests for guiding adjuvant chemotherapy among patients with HR+, HER2-, and LN- or LN+ (1-3) early-stage (stages I to IIIa) invasive breast cancer. As the indications for GEP tests expand, new evidence will need to be reviewed, and new guidelines will need to be developed.

5.4 Strengths and limitations of the HTA

This rapid HTA has a number of strengths. A robust approach to the assessment process was employed with the publication of a protocol for the HTA, adherence to the HTA Core Model® proposed by the European Network for Health Technology Assessment (EUnetHTA),[268] covering the epidemiology, burden of disease, description of the technology, and clinical effectiveness domains, and the establishment of an Expert Advisory Group (EAG) comprising a range of key stakeholders to support the assessment. All chapters were reviewed and updated in line with recommendations from the EAG. The technical characteristics of each of the GEP tests in Chapter 2 were reviewed by the test manufacturers to ensure that each test was accurately described. The systematic search in Chapter 4 used a comprehensive approach, leveraging Cochrane methodology to systematically review the international evidence and sourcing relevant studies from the GEP test manufacturers to ensure that the most relevant clinical effectiveness evidence was included. Nonetheless, this assessment has some limitations which should be considered in the interpretation of the evidence.

The international HTAs and clinical guidelines detailed throughout this rapid HTA were not identified through a systematic search of the international literature. Therefore, all relevant HTAs and international guidelines may not have been captured. Additionally, a number of the included guidance documents were available in non-English languages only and therefore, interpretation may be subject to translation error. Further, international practice evolves over time, often coinciding with the publication of important trials, and this guidance is therefore likely to change.

In 2020, a comprehensive HTA comparing EndoPredict®, MammaPrint®, Prosigna® and Oncotype DX® was conducted by the government agency Ontario Health and the Canadian Agency for Drugs and Technologies in Health.[29] Due to the recency of that review and overlap with the current rapid HTA, the current systematic review updated the review conducted as part of the Ontario Health HTA. However, this approach carries the risk that relevant sources of data may have been missed or data incorrectly extracted or interpreted. To mitigate this risk, 1) data from a random selection of studies were cross-checked with their primary sources, 2) included studies were cross-checked with related reviews and pre-2018 studies

identified in the current search, and 3) GEP test manufacturers were invited to submit lists of relevant studies.

Finally, there were considerable limitations in the clinical effectiveness evidence base. Studies assessing the prognostic ability of the tests varied substantially by study design, analytic approach, risk cut-off scores used within GEP tests, outcome, and study population. The predictive performance of only two GEP tests, Oncotype DX® and MammaPrint®, was assessed by RCTs. These RCTs had major limitations and findings were unclear. As a result, comparisons of the prognostic and predictive abilities of GEP tests with each other and with standard clinical and pathological information already available to clinicians and patients were limited.

5.5 Conclusions

This rapid HTA examined the four commercially available GEP tests (that is, Oncotype DX®, MammaPrint®, EndoPredict®, and Prosigna®) to guide adjuvant chemotherapy among patients with HR+, HER2-, and LN- or LN+ (1-3 nodes) early-stage (stages I to IIIa) invasive breast cancer. GEP tests are intended to provide information on disease prognosis (that is, distant recurrence and survival) and the predicted benefit of chemotherapy (that is, identify the people who are most likely to benefit from chemotherapy). Oncotype DX® is currently the only GEP test that is reimbursed by the Health Service Executive in Ireland.

Advice relating to GEP tests from previous HTAs and international guidelines varies substantially, despite being grounded broadly in the same evidence base. This is likely to be influenced by issues such as study heterogeneity and study designs that generally do not assess the relative prognostic and predictive value of GEP tests, resulting in a complicated and unclear evidence base. The current review found that the prognostic accuracy evidence is comparable across the four GEP tests. Three tests are indicated for predictive use (Oncotype DX®, MammaPrint®, and EndoPredict®), of which two have had their predictive ability assessed in RCTs (Oncotype DX® and MammaPrint®). Considering predictive ability, although there are limited data to differentiate between Oncotype DX® and MammaPrint®, the available evidence supports the continued use of Oncotype DX® among LN- patients and the evidence most strongly supports the continued use of Oncotype DX® in postmenopausal women, based on available five-year follow-up data among LN+ patients.

Decision-making regarding the use of GEP tests should take into account of differences in factors such as test indications, test costs and feasibility of use, particularly with respect to laboratory resources, in addition to clinical effectiveness. Finally, several steps could be taken to help optimise the management and use of

GEP tests. These may include the collection, as part of a national registry, of clinical and pathological characteristics and tumour gene expression profiles of all patients in Ireland whose breast cancer specimens undergo GEP testing. Furthermore, there may also be a role for the development of guidance to outline the patient subgroups in which gene expression profiling testing should be used, the appropriate tumour sampling methods and preparation techniques, and the interpretation of test results.

References

1. National Cancer Registry Ireland. Cancer Incidence Projections For Ireland 2020-2045. 2019.

2. National Cancer Registry Ireland. Cancer Trends 37 - Breast cancer 1994-2016 2019 [Available from: https://www.ncri.ie/publications/cancer-trends-and-projections/cancer-trends-37-breast-cancer-1994-2016].

3. McSorley LM, Tharmabala M, Al Rahbi F, McSorley K, Chew S, Evoy D, et al. Real-world analysis of clinical and economic impact of 21-gene recurrence score (RS) testing in early-stage breast cancer (ESBC) in Ireland. Breast cancer research and treatment. 2021;188(3):789-98.

4. Health Service Executive. Treatment - Breast cancer in women 2022 [Available from: https://www2.hse.ie/conditions/breast-cancer-women/treatment/].

5. National Clinical Effectiveness Committee. Diagnosis, staging and treatment of patients with breast cancer: National Clinical Guideline No. 7. 2015.

6. Munoz D, Near AM, van Ravesteyn NT, Lee SJ, Schechter CB, Alagoz O, et al. Effects of Screening and Systemic Adjuvant Therapy on ER-Specific US Breast Cancer Mortality. JNCI: Journal of the National Cancer Institute. 2014;106(11).

7. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. The Lancet. 2005;365(9472):1687-717.

8. Ribnikar D, Cardoso F. Tailoring chemotherapy in early-stage breast cancer: based on tumor biology or tumor burden? American Society of Clinical Oncology Educational Book. 2016;36:e31-e8.

9. Irish Society of Medical Oncology. Oncotype-DX® gene expression profile and chemotherapy decision-making in patients with early stage breast cancer. 2011.

10. Health Service Executive National Cancer Control Programme. Breast Chemotherapy Regimens [Available from: https://www.hse.ie/eng/services/list/5/cancer/profinfo/chemoprotocols/breast/].

11. Giuliano AE, Edge SB, Hortobagyi GN. Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. Annals of Surgical Oncology. 2018;25(7):1783-5.

12. Department of Health. Diagnosis, Staging and Treatment of Patients with Breast Cancer Dublin2015 [Available from: https://www.gov.ie/en/collection/6fb16f-diagnosis-staging-and-treatment-of-patients-with-breast-cancer/.

13. Department of Health. National Cancer Strategy 2017 - 2026 [Available from: https://www.gov.ie/en/publication/a89819-national-cancer-strategy-2017-2026/].

14. Health Service Executive. Breast screening - BreastCheck [Available from: https://www2.hse.ie/conditions/breast-screening/].

15. National Centre for Pharmacoeconomics. Oncotype DX® 2011 [Available from: http://www.ncpe.ie/news/oncotype-dx/].

16. McVeigh TP, Hughes LM, Miller N, Sheehan M, Keane M, Sweeney KJ, et al. The impact of Oncotype DX testing on breast cancer management and chemotherapy prescribing patterns in a tertiary referral centre. European Journal of Cancer. 2014;50(16):2763-70.

17. Smyth L, Watson G, Walsh EM, Kelly CM, Keane M, Kennedy MJ, et al. Economic impact of 21-gene recurrence score testing on early-stage breast cancer in Ireland. Breast cancer research and treatment. 2015;153(3):573-82.

18. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2019;30(8):1194-220.

19. Gradishar WJ, Anderson BO, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast Cancer, Version 3.2020, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network : JNCCN. 2020;18(4):452-78.

20. Schünemann HJ, Lerda D, Dimitrova N, Alonso-Coello P, Gräwingholt A, Quinn C, et al. Methods for Development of the European Commission Initiative on Breast Cancer Guidelines: Recommendations in the Era of Guideline Transparency. Annals of internal medicine. 2019;171(4):273-80.

21. National Screening Service. [Available from: https://www.screeningservice.ie/index.php].

22. Fitzpatrick PE, Greehy G, Mooney MT, Flanagan F, Larke A, Connors A, et al. Evolution of the National Breast Screening Programme in Ireland: Two-year interval analysis (2004–2013) of BreastCheck. Journal of Medical Screening. 2017;25(4):191-6.

23. Health Information and Quality Authority. HTA of High Risk Breast Cancer Surveillance 2013 [Available from: https://www.hiqa.ie/reports-and-publications/health-technology-assessment/hta-high-risk-breast-cancer-surveillance.

24. National Screening Service. BreastCheck Programme Report 2018 and 2019 [Available from: https://www.screeningservice.ie/publications/.

25. World Health Organization. WHO Classification of Tumours of the Breast WHO Classification of Tumours, 4th Edition, Volume 4 2012 [Available from: https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Breast-2012].

26. Schettini F, Brasó-Maristany F, Kuderer NM, Prat A. A perspective on the development and lack of interchangeability of the breast cancer intrinsic subtypes. npj Breast Cancer. 2022;8(1):85.

27. Cardoso F, Bese N, Distelhorst SR, Bevilacqua JL, Ginsburg O, Grunberg SM, et al. Supportive care during treatment for breast cancer: resource allocations in low- and middle-income countries. A Breast Health Global Initiative 2013 consensus statement. Breast (Edinburgh, Scotland). 2013;22(5):593-605.

28. World Health Organization. Breast cancer 2021 [Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer].

29. Ontario Health (Quality). Gene Expression Profiling Tests for Early-Stage Invasive Breast Cancer: A Health Technology Assessment. Ontario health technology assessment series. 2020;20(10):1-234.

30. Lohrisch C, Paltiel C, Gelmon K, Speers C, Taylor S, Barnett J, et al. Impact on survival of time from definitive surgery to initiation of adjuvant chemotherapy for early-stage breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2006;24(30):4888-94.

31. Institute of Health Economics Alberta Canada. Clinical effectiveness and cost-effectiveness of Oncotype DX and Prosigna genetic testing in early-stage breast cancer 2019 [Available from: https://open.alberta.ca/dataset/clinical-effectiveness-and-cost-effectiveness-of-oncotype-dx-and-prosigna-genetic-testing/resource/3f99d16d-f108-48b6-b968-0485e28276ea].

32. Peto R, Davies C, Godwin J, Gray R, Pan HC, Clarke M, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. The Lancet. 2012;379(9814):432-44.

33. Partridge AH, Burstein HJ, Winer EP. Side effects of chemotherapy and combined chemohormonal therapy in women with early-stage breast cancer. Journal of the National Cancer Institute Monographs. 2001(30):135-42.

34. Kayl AE, Meyers CA. Side-effects of chemotherapy and quality of life in ovarian and breast cancer patients. Current opinion in obstetrics & gynecology. 2006;18(1):24-8.

35. Kuchuk I, Bouganim N, Beusterien K, Grinspan J, Vandermeer L, Gertler S, et al. Preference weights for chemotherapy side effects from the perspective of women with breast cancer. Breast cancer research and treatment. 2013;142(1):101-7.

36. Odle TG. Adverse effects of breast cancer treatment. Radiologic technology. 2014;85(3):297M-319M.

37. Plana JC, Galderisi M, Barac A, Ewer MS, Ky B, Scherrer-Crosbie M, et al. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. European Heart Journal–Cardiovascular Imaging. 2014;15(10):1063-93.

38. Bao T, Basal C, Seluzicki C, Li SQ, Seidman AD, Mao JJ. Long-term chemotherapy-induced peripheral neuropathy among breast cancer survivors: prevalence, risk factors, and fall risk. Breast cancer research and treatment. 2016;159(2):327-33.

39. Department of Health. Minister for Health announces abolition of public in-patient charges for children under 16 [Available from: https://www.gov.ie/en/press-release/b7867-minister-for-health-announces-abolition-of-public-in-patient-charges-for-children-under-16/].

40. Irish Cancer Society. Inpatient charges [Available from: https://www.cancer.ie/about-us/about-the-irish-cancer-society/what-we-do/cancer-advocacy/inpatient-charges].

41. Pisu M, Azuero A, McNees P, Burkhardt J, Benz R, Meneses K. The out of pocket cost of breast cancer survivors: a review. Journal of cancer survivorship : research and practice. 2010;4(3):202-9.

42. Iragorri N, de Oliveira C, Fitzgerald N, Essue B. The Out-of-Pocket Cost Burden of Cancer Care-A Systematic Literature Review. Current Oncology. 2021;28(2):1216-48.

43. Blumen H, Fitch K, Polkus V. Comparison of Treatment Costs for Breast Cancer, by Tumor Stage and Type of Service. American health & drug benefits. 2016;9(1):23-32.

44. Harnan S, Tappenden P, Cooper K, Stevens J, Bessey A, Rafia R, et al. Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer: a systematic review and economic analysis. Health technology assessment (Winchester, England). 2019;23(30):1-328.

45. National Institute for Health and Care Excellence. Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer: NICE; 2018 [Available from: https://www.nice.org.uk/guidance/dg34].

46. Blok EJ, Bastiaannet E, van den Hout WB, Liefers GJ, Smit V, Kroep JR, et al. Systematic review of the clinical and economic value of gene expression profiles for invasive early breast cancer available in Europe. Cancer treatment reviews. 2018;62:74-90.

47. Hess V. [Adjuvant!Online--an Internet-based decision tool for adjuvant chemotherapy in early breast cancer]. Therapeutische Umschau Revue therapeutique. 2008;65(4):201-5.

48. Fong Y, Evans J, Brook D, Kenkre J, Jarvis P, Gower-Thomas K. The Nottingham Prognostic Index: five- and ten-year data for all-cause survival within a screened population. Annals of the Royal College of Surgeons of England. 2015;97(2):137-9.

49. Engelhardt EG, van den Broek AJ, Linn SC, Wishart GC, Rutgers EJT, van de Velde AO, et al. Accuracy of the online prognostication tools PREDICT and Adjuvant! for early-stage breast cancer patients younger than 50 years. European journal of cancer (Oxford, England : 1990). 2017;78:37-44.

50. Bartlett JMS, Sgroi DC, Treuner K, Zhang Y, Ahmed I, Piper T, et al. Breast Cancer Index and prediction of benefit from extended endocrine therapy in breast cancer patients treated in the Adjuvant Tamoxifen-To Offer More? (aTTom) trial. Annals of Oncology. 2019;30(11):1776-83.

51. Sparano JA. TAILORx: trial assigning individualized options for treatment (Rx). Clinical breast cancer. 2006;7(4):347-50.

52. Expert Advisory Group. Personal communication 2022.

53. Gasparini G, Pozza F, Harris AL. Evaluating the potential usefulness of new prognostic and predictive indicators in node-negative breast cancer patients. Journal of the National Cancer Institute. 1993;85(15):1206-19.

54. Hayes DF, Trock B, Harris AL. Assessing the clinical impact of prognostic factors: when is "statistically significant" clinically useful? Breast cancer research and treatment. 1998;52(1-3):305-19.

55. Foukakis T, Bergh J. Prognostic and predictive factors in early, non-metastatic breast cancer [Available from: https://www.uptodate.com/contents/prognostic-and-predictive-factors-in-early-non-metastatic-breast-cancer].

56. Febbo PG, Ladanyi M, Aldape KD, De Marzo AM, Hammond ME, Hayes DF, et al. NCCN Task Force report: Evaluating the clinical utility of tumor markers in oncology. Journal of the National Comprehensive Cancer Network : JNCCN. 2011;9 Suppl 5:S1-32.

57. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genetics in medicine : official journal of the American College of Medical Genetics. 2009;11(1):3-14.

58. Sargent DJ, Conley BA, Allegra C, Collette L. Clinical trial designs for predictive marker validation in cancer treatment trials. Journal of Clinical Oncology. 2005;23(9):2020-7.

59. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. Journal of the National Cancer Institute. 2009;101(21):1446-52.

60. Clancy S, Brown W. Translation: DNA to mRNA to Protein. Nature Education. 2008;1(1):101.

61. Marchionni L, Wilson RF, Marinopoulos SS, Wolff AC, Parmigiani G, Bass EB, et al. Impact of gene expression profiling tests on breast cancer outcomes. Evidence report/technology assessment. 2007(160):1-105.

62. Govindarajan R, Duraiyan J, Kaliyappan K, Palanisamy M. Microarray and its applications. Journal of pharmacy & bioallied sciences. 2012;4(Suppl 2):S310.

63. Lěvěque N, Renois F, Andréoletti L. The microarray technology: facts and controversies. Clinical microbiology and infection. 2013;19(1):10-4.

64. Shen M, Zhou Y, Ye J, Abdullah Al-Maskri AA, Kang Y, Zeng S, et al. Recent advances and perspectives of nucleic acid detection for coronavirus. Journal of pharmaceutical analysis. 2020;10(2):97-101.

65. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. New England Journal of Medicine. 2015;373(21):2005-14.

66. 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. The New England journal of medicine. 2018;379(2):111-21.

67. Sparano JA, Gray RJ, Ravdin PM, Makower DF, Pritchard KI, Albain KS, et al. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. New England journal of medicine. 2019;380(25):2395-405.

68. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. Journal of Clinical Oncology. 2008;26(5):721-8.

69. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor–positive breast cancer. Journal of Clinical Oncology. 2006;24(23):3726-34.

70. Kalinsky K, Barlow WE, Gralow JR, Meric-Bernstam F, Albain KS, Hayes DF, et al. 21-gene assay to inform chemotherapy benefit in node-positive breast cancer. New England Journal of Medicine. 2021;385(25):2336-47.

71. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh I-T, et al. Prognostic and predictive value of the 21-gene recurrence score assay in a randomized trial of chemotherapy for postmenopausal, node-positive, estrogen receptor-positive breast cancer. The Lancet Oncology. 2010;11(1):55.

72. Buus R, Szijgyarto Z, Schuster EF, Xiao H, Haynes BP, Sestak I, et al. Development and validation for research assessment of Oncotype DX® Breast Recurrence Score, EndoPredict® and Prosigna®. NPJ breast cancer. 2021;7(1):1-8.

73. Exact Sciences. TAILORx [Available from: https://precisiononcology.exactsciences.com/healthcare-providers/treatment-determination/breast-cancer/oncotype-dx-breast-recurrence-score/TAILORx].

74. Tang G, Cuzick J, Costantino JP, Dowsett M, Forbes JF, Crager M, et al. Risk of recurrence and chemotherapy benefit for patients with node-negative, estrogen

receptor–positive breast cancer: recurrence score alone and integrated with pathologic and clinical factors. Journal of Clinical Oncology. 2011;29(33):4365.

75. Exact Sciences. How to Order the Oncotype DX® Breast Cancer Assay in UK and Ireland [Available from: https://www.oncotypeiq.com/en-IE/breast-cancer/healthcare-professionals/oncotype-dx-breast-recurrence-score/how-to-order-a-test].

76. Rizvi SAJ, Eldeeb H, Soulsby R, Taylor A, Karina M. Oncotype Dx in Node-negative Breast Cancer: Our Experience. Clinical Oncology. 2019;31(7):e113.

77. Delahaye L, Drukker CA, Dreezen C, Witteveen A, Chan B, Snel M, et al. A breast cancer gene signature for indolent disease. Breast cancer research and treatment. 2017;164(2):461-6.

78. Agendia. Comparing MammaPrint and BluePrint results between core needle biopsy and surgical resection breast cancer specimens 2021 [Available from: https://agendia.com/wp-content/uploads/2021/07/2021_St.-Gallen-poster_Core-vs-SR.pdf].

79. Myriad Oncology. A provider guide to EndoPredict [Available from: https://myriad-library.s3.amazonaws.com/endopredict/EndoPredict+Provider+Clinical+Guide.pdf].

80. Warf MB, Rajamani S, Krappmann K, Doedt J, Cassiano J, Brown K, et al. Analytical validation of a 12-gene molecular test for the prediction of distant recurrence in breast cancer. Future science OA. 2017;3(3):FSO221.

81. Dubsky P, Brase J, Jakesz R, Rudas M, Singer C, Greil R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2− breast cancer patients. British journal of cancer. 2013;109(12):2959-64.

82. Australian Clinical Labs. EndoPredict [Available from: https://www.clinicallabs.com.au/endopredict/].

83. Buus R, Sestak I, Kronenwett R, Ferree S, Schnabel CA, Baehner FL, et al. Molecular Drivers of Oncotype DX, Prosigna, EndoPredict, and the Breast Cancer Index: A TransATAC Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2021;39(2):126-35.

84. Prosigna. Package Insert Prosigna® Breast Cancer Prognostic Gene Signature Assay 2019 [Available from: https://37z6701s3hobk4wnh3t79rq1-wpengine.netdna-ssl.com/wp-content/uploads/2022/03/LBL-C0223-07-1.pdf].

85. Kulkarni MM. Digital multiplexed gene expression analysis using the NanoString nCounter system. Current protocols in molecular biology. 2011;Chapter 25:Unit25B.10.

86. Veracyte. Package Insert Prosigna® Breast Cancer Prognostic Gene Signature Assay 2022 [Available from: https://www.prosigna.com/en/resources/].

87. European Network for Health Technology Assessment. Final assessment report on MammaPrint® – Added value of using the gene expression signature test MammaPrint® for adjuvant chemotherapy decision-making in early breast cancer 2018 [Available from: https://www.eunethta.eu/final-assessment-report-on-mammaprint-added-value-of-using-the-gene-expression-signature-test-mammaprint-for-adjuvant-chemotherapy-decision-making-in-early-breast-cancer/].

88. Washington State Health Care Authority. Gene expression profile testing of cancer tissue Final evidence report. 2018.

89. Oregon Health Authority. Health Evidence Review Commission (HERC) Coverage Guidance: Genome Expression Profiling for Breast Cancer. 2018.

90. Medical Services Advisory Committee. 1408 - A prognostic RT-qPCR test run locally for ER+ve /HER2-ve breast cancer that determines the risk of early and late metastasis in node negative and positive cancer under endocrine treatment 2022 [Available from: http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1408-public].

91. Agenas. Added value of using the multigene expression signature test for adjuvant chemotherapy decision-making in early breast cancer 2020 [Available from: https://www.agenas.gov.it/aree-tematiche/organizzazione-dei-servizi-sanitari/rete-nazionale-tumori-rari?view=category&id=95].

92. Haute Authorité de Santé. Clinical utility of genomic signatures in early stage breast cancer. 2019.

93. Medical Services Advisory Committee. 1342.5 - Oncotype DX® breast cancer assay to quantify the risk of disease recurrence and predict adjuvant chemotherapy benefit (Resubmission) 2019 [Available from: http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1342.5-public#:~:text=Oncotype%20DX%C2%AE%20is%20a,HER2%2D)%2C%20invasive%20breast%20cancer].

94. Medical Technology Product Council. Gene expression analysis before decisions on adjuvant treatment of breast cancer 2022 [Available from:

https://www.janusinfo.se/nationelltinforandeavmedicinteknik/rekommendationer.4.74a8b4c1170b06b45db47225.html].

95. IQWiG. Biomarker-based tests to support the decision for or against adjuvant systemic chemotherapy in primary breast cancer - state of knowledge 2020 [Available from: https://www.iqwig.de/en/projects/d19-01.html].

96. Meleth S, Reeder-Hayes K, Ashok M, Clark R, Funkhouser W, Wines R, et al. Technology Assessment of Molecular Pathology Testing for the Estimation of Prognosis for Common Cancers [Internet]. 2014.

97. University of Calgary HTAU, . Use of Oncotype DX for guiding adjuvant chemotherapy decisions in early stage invasive breast cancer patients in Alberta : a health technology assessment. 2013.

98. Narain T, L. A. Gene Expression Tests for Women with Early Stage Breast Cancer: A Review of Clinical Utility and Cost-Effectiveness. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health: 2017

99. San Miguel L, Vlayen J, De Laet C. Gene expression profiling and immunohistochemistry tests for personalised management of adjuvant chemotherapy decisions in early breast cancer – a Rapid Assessment. Health Technology Assessment (HTA). Brussels: Belgian Health Care Knowledge Centre (KCE), 2015.

100. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. New England Journal of Medicine. 2016;375(8):717-29.

101. Andre F, Ismaila N, Henry NL, Somerfield MR, Bast RC, Barlow W, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: ASCO Clinical Practice Guideline Update-Integration of Results From TAILORx. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2019;37(22):1956-64.

102. Kantor O, King TA, Shak S, Russell CA, Giuliano AE, Hortobagyi GN, et al. Expanding Criteria for Prognostic Stage IA in Hormone Receptor–Positive Breast Cancer. JNCI: Journal of the National Cancer Institute. 2021;113(12):1744-50.

103. Stein RC, Dunn JA, Bartlett JM, Campbell AF, Marshall A, Hall P, et al. OPTIMA prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. Health technology assessment (Winchester, England). 2016;20(10):xxiii-xxix, 1-201.

104. Nitz U, Gluz O, Graeser M, Christgen M, Kuemmel S, Grischke E-M, et al. De-escalated neoadjuvant pertuzumab plus trastuzumab therapy with or without weekly paclitaxel in HER2-positive, hormone receptor-negative, early breast cancer (WSG-ADAPT-HER2+/HR-): survival outcomes from a multicentre, open-label, randomised, phase 2 trial. The Lancet Oncology. 2022;23(5):625-35.

105. Geyer CE, Jr., Tang G, Mamounas EP, Rastogi P, Paik S, Shak S, et al. 21-Gene assay as predictor of chemotherapy benefit in HER2-negative breast cancer. NPJ Breast Cancer. 2018;4:37.

106. Duffy M, Harbeck N, Nap M, Molina R, Nicolini A, Senkus E, et al. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). European Journal of Cancer. 2017;75:284-98.

107. Giorgi Rossi P, Lebeau A, Canelo-Aybar C, Saz-Parkinson Z, Quinn C, Langendam M, et al. Recommendations from the European Commission Initiative on Breast Cancer for multigene testing to guide the use of adjuvant chemotherapy in patients with early breast cancer, hormone receptor positive, HER-2 negative. British Journal of Cancer. 2021;124(9):1503-12.

108. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. Journal of Clinical Oncology. 2016;34(10):1134-50.

109. Krop I, Ismaila N, Andre F, Bast RC, Barlow W, Collyar DE, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Focused Update. Journal of Clinical Oncology. 2017;35(24):2838-47.

110. Burstein HJ, Curigliano G, Thürlimann B, Weber WP, Poortmans P, Regan MM, et al. Customizing local and systemic therapies for women with early breast cancer: the St. Gallen International Consensus Guidelines for treatment of early breast cancer 2021. Annals of Oncology. 2021;32(10):1216-35.

111. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA: a cancer journal for clinicians. 2017;67(4):290-303.

112. Andre F, Ismaila N, Allison KH, Barlow WE, Collyar DE, Damodaran S, et al. Biomarkers for adjuvant endocrine and chemotherapy in early-stage breast cancer: ASCO guideline update. Journal of Clinical Oncology. 2022;40(16):1816-37.

113. EUR-Lex. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices [Available from: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A31998L0079].

114. EUR-Lex. REGULATION (EU) 2022/112 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 January 2022 amending Regulation (EU) 2017/746 as regards transitional provisions for certain in vitro diagnostic medical devices and the deferred application of conditions for in-house devices 2022 [Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R0112&from=EN].

115. National Cancer Registry Ireland. Cancer Factsheet Female breast. 2018.

116. Sun Y-S, Zhao Z, Yang Z-N, Xu F, Lu H-J, Zhu Z-Y, et al. Risk Factors and Preventions of Breast Cancer. Int J Biol Sci. 2017;13(11):1387-97.

117. Zehr KR. Diagnosis and Treatment of Breast Cancer in Men. Radiologic technology. 2019;91(1):51m-61m.

118. Rudlowski C. Male Breast Cancer. Breast care (Basel, Switzerland). 2008;3(3):183-9.

119. Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LA, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. Journal of the National Cancer Institute. 2014;106(5).

120. O'Brien KM, Mooney T, Fitzpatrick P, Sharp L. Screening status, tumour subtype, and breast cancer survival: a national population-based analysis. Breast cancer research and treatment. 2018;172(1):133-42.

121. Konduri S, Singh M, Bobustuc G, Rovin R, Kassam A. Epidemiology of male breast cancer. Breast (Edinburgh, Scotland). 2020;54:8-14.

122. Dall GV, Britt KL. Estrogen Effects on the Mammary Gland in Early and Late Life and Breast Cancer Risk. Frontiers in Oncology. 2017;7.

123. Ilic L, Haidinger G, Simon J, Hackl M, Schernhammer E, Papantoniou K. Trends in female breast cancer incidence, mortality, and survival in Austria, with focus on age, stage, and birth cohorts (1983–2017). Scientific Reports. 2022;12(1):7048.

124. Health Service Executive. Symptoms and diagnosis - Breast cancer in men [Available from: https://www2.hse.ie/conditions/breast-cancer-men/].

125. National Health Service. Overview - Breast cancer in men [Available from: https://www.nhs.uk/conditions/breast-cancer-in-men/].

126. Majeed W, Aslam B, Javed I, Khaliq T, Muhammad F, Ali A, et al. Breast cancer: major risk factors and recent developments in treatment. Asian Pacific Journal of Cancer Prevention. 2014;15(8):3353-8.

127. Apostolou P, Fostira F. Hereditary breast cancer: the era of new susceptibility genes. BioMed research international. 2013;2013:747318.

128. Dumitrescu RG, Cotarla I. Understanding breast cancer risk -- where do we stand in 2005? Journal of cellular and molecular medicine. 2005;9(1):208-21.

129. Brewer HR, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Family history and risk of breast cancer: an analysis accounting for family structure. Breast cancer research and treatment. 2017;165(1):193-200.

130. Anderson RL, Ingman WV, Britt KL. Editorial: How Reproductive History Influences Our Breast Cancer Risk. Frontiers in Oncology. 2017;7.

131. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. The Lancet Oncology. 2012;13(11):1141-51.

132. Gao YT, Shu XO, Dai Q, Potter JD, Brinton LA, Wen W, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. International journal of cancer. 2000;87(2):295-300.

133. Reeves GK, Pirie K, Green J, Bull D, Beral V, for the Million Women Study C. Reproductive factors and specific histological types of breast cancer: prospective study and meta-analysis. British Journal of Cancer. 2009;100(3):538-44.

134. O'Brien KM, Sun J, Sandler DP, DeRoo LA, Weinberg CR. Risk factors for young-onset invasive and in situ breast cancer. Cancer Causes & Control. 2015;26(12):1771-8.

135. Sisti JS, Collins LC, Beck AH, Tamimi RM, Rosner BA, Eliassen AH. Reproductive risk factors in relation to molecular subtypes of breast cancer: Results from the nurses' health studies. International journal of cancer. 2016;138(10):2346-56.

136. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. The Lancet. 1996;347(9017):1713-27.

137. McPherson K, Steel CM, Dixon JM. ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. BMJ 2000;321(7261):624-8.

138. Pathak DR, Osuch JR, He J. Breast carcinoma etiology: current knowledge and new insights into the effects of reproductive and hormonal risk factors in black and white populations. Cancer. 2000;88(5 Suppl):1230-8.

139. Anderson KN, Schwab RB, Martinez ME. Reproductive risk factors and breast cancer subtypes: a review of the literature. Breast cancer research and treatment. 2014;144(1):1-10.

140. Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. Breast Cancer Research. 2006;8(4):R43.

141. Ursin G, Bernstein L, Lord SJ, Karim R, Deapen D, Press MF, et al. Reproductive factors and subtypes of breast cancer defined by hormone receptor and histology. British Journal of Cancer. 2005;93(3):364-71.

142. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. The Lancet. 2002;360(9328):187-95.

143. Huo D, Adebamowo CA, Ogundiran TO, Akang EE, Campbell O, Adenipekun A, et al. Parity and breastfeeding are protective against breast cancer in Nigerian women. British Journal of Cancer. 2008;98(5):992-6.

144. Bernstein L. Epidemiology of endocrine-related risk factors for breast cancer. Journal of mammary gland biology and neoplasia. 2002;7(1):3-15.

145. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiologic reviews. 1993;15(1):36-47.

146. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. The Lancet. 2003;362(9382):419-27.

147. Shah R, Rosso K, Nathanson SD. Pathogenesis, prevention, diagnosis and treatment of breast cancer. World journal of clinical oncology. 2014;5(3):283-98.

148. Alkabban FM, Ferguson T. Breast Cancer: StatPearls Publishing; [updated 07.08.2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482286/.

149. Galipeau N, Klooster B, Krohe M, Tang DH, Revicki DA, Cella D. Understanding key symptoms, side effects, and impacts of HR+/HER2- advanced breast cancer: qualitative study findings. Journal of patient-reported outcomes. 2019;3(1):10.

150. Irvin W, Jr., Muss HB, Mayer DK. Symptom management in metastatic breast cancer. Oncologist. 2011;16(9):1203-14.

151. Feng Y, Spezia M, Huang S, Yuan C, Zeng Z, Zhang L, et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes & diseases. 2018;5(2):77-106.

152. Fisher B, Bauer M, Wickerham DL, Redmond CK, Fisher ER, Cruz AB, et al. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. Cancer. 1983;52(9):1551-7.

153. Geurts YM, Witteveen A, Bretveld R, Poortmans PM, Sonke GS, Strobbe LJA, et al. Patterns and predictors of first and subsequent recurrence in women with early breast cancer. Breast cancer research and treatment. 2017;165(3):709-20.

154. Pan H, Gray R, Braybrooke J, Davies C, Taylor C, McGale P, et al. 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. New England Journal of Medicine. 2017;377(19):1836-46.

155. Onitilo AA, Engel JM, Greenlee RT, Mukesh BN. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. Clinical medicine & research. 2009;7(1-2):4-13.

156. Ihemelandu CU, Leffall LD, Jr., Dewitty RL, Naab TJ, Mezghebe HM, Makambi KH, et al. Molecular breast cancer subtypes in premenopausal and postmenopausal African-American women: age-specific prevalence and survival. The Journal of surgical research. 2007;143(1):109-18.

157. Early Breast Cancer Trialists' Collaborative Group. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. The Lancet. 2011;378(9793):771-84.

158. Perme MP, Henderson R, Stare J. An approach to estimation in relative survival regression. Biostatistics (Oxford, England). 2009;10(1):136-46.

159. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: a cancer journal for clinicians. 2021;71(3):209-49.

160. Xu S, Liu Y, Zhang T, Zheng J, Lin W, Cai J, et al. The Global, Regional, and National Burden and Trends of Breast Cancer From 1990 to 2019: Results From the Global Burden of Disease Study 2019. Frontiers in Oncology. 2021;11:689562.

161. National Cancer Registry Ireland. Cancer Trends 38 - Breast, cervical and colorectal cancer 1994-2019. 2022.

162. Moja L, Tagliabue L, Balduzzi S, Parmelli E, Pistotti V, Guarneri V, et al. Trastuzumab containing regimens for early breast cancer. Cochrane Database of Systematic Reviews. 2012(4).

163. Bradley R, Braybrooke J, Gray R, Hills R, Liu Z, Peto R, et al. Trastuzumab for early-stage, HER2-positive breast cancer: a meta-analysis of 13 864 women in seven randomised trials. The Lancet Oncology. 2021;22(8):1139-50.

164. Tovey SM, Brown S, Doughty JC, Mallon EA, Cooke TG, Edwards J. Poor survival outcomes in HER2-positive breast cancer patients with low-grade, node-negative tumours. British Journal of Cancer. 2009;100(5):680-3.

165. National Cancer Registry Ireland. Personal Communication. 2022.

166. Rajarajan S, Korlimarla A, Alexander A, Anupama CE, Ramesh R, Srinath BS, et al. Pre-Menopausal Women With Breast Cancers Having High AR/ER Ratios in the Context of Higher Circulating Testosterone Tend to Have Poorer Outcomes. Frontiers in endocrinology. 2021;12:679756.

167. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.

168. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. Journal of Clinical Epidemiology. 2016;75:40-6.

169. Benowitz S. Chemotherapy Culture Differences Persist, but Molecular Insights Foster Change. JNCI: Journal of the National Cancer Institute. 2012;104(8):576-8.

170. Gao W, Lin L, Fei X, Chen X, Shen K. Decision-making of Adjuvant Chemotherapy for Breast Cancer Patients with Discordant Risk Classifications

between Clinical-Pathological Factors and 21-gene Recurrence Score. Journal of Cancer. 2020;11(9):2509.

171. Ma SJ, Oladeru OT, Farrugia M, Mikucki M, Iovoli A, Shekher R, et al. Survival outcome of adjuvant chemotherapy and high 21-gene recurrence score in early-stage breast cancer. The Breast Journal. 2021;27(1):27-34.

172. Sterne JAC, Savović 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.

173. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Annals of internal medicine. 2019;170(1):51-8.

174. Kim SY, Park JE, Lee YJ, Seo H-J, Sheen S-S, Hahn S, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. Journal of Clinical Epidemiology. 2013;66(4):408-14.

175. Schünemann HJ, J. B, Guyatt G, Oxman AD, editors. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach.2013.

176. Pece S, Sestak I, Montani F, Tillhon M, Maisonneuve P, Freddi S, et al. Comparison of StemPrintER with Oncotype DX Recurrence Score for predicting risk of breast cancer distant recurrence after endocrine therapy. European Journal of Cancer. 2022;164:52-61.

177. Piccart M, van't Veer LJ, Poncet C, Cardozo JML, Delaloge S, Pierga J-Y, et al. 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. The Lancet Oncology. 2021;22(4):476-88.

178. Knauer M, Mook S, Rutgers EJ, Bender RA, Hauptmann M, Van de Vijver MJ, et al. The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer. Breast cancer research and treatment. 2010;120(3):655-61.

179. Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1–3 positive lymph nodes in an independent validation study. Breast cancer research and treatment. 2008;116(2):295-302.

180. Foroutan F, Guyatt G, Zuk V, Vandvik PO, Alba AC, Mustafa R, et al. GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. Journal of clinical epidemiology. 2020;121:62-70.

181. Dowsett M, Allred C, Knox J, Quinn E, Salter J, Wale C, et al. Relationship between quantitative estrogen and progesterone receptor expression and human epidermal growth factor receptor 2 (HER-2) status with recurrence in the Arimidex, Tamoxifen, Alone or in Combination trial. Journal of Clinical Oncology. 2008;26(7):1059-65.

182. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(11):1829-34.

183. Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C, et al. Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA oncology. 2018;4(4):545-53.

184. Sparano JA, Gray RJ, Makower DF, Albain KS, Saphner TJ, Badve SS, et al. Clinical outcomes in early breast cancer with a high 21-gene recurrence score of 26 to 100 assigned to adjuvant chemotherapy plus endocrine therapy: a secondary analysis of the TAILORx randomized clinical trial. JAMA oncology. 2020;6(3):367-74.

185. Poorvu PD, Gelber SI, Rosenberg SM, Ruddy KJ, Tamimi RM, Collins LC, et al. Prognostic impact of the 21-gene recurrence score assay among young women with node-negative and node-positive ER-positive/HER2-negative breast cancer. Journal of Clinical Oncology. 2020;38(7):725.

186. Lynch SM, Russell NM, Barron S, Wang C-JA, Loughman T, Dynoodt P, et al. Prognostic value of the 6-gene OncoMasTR test in hormone receptor–positive HER2-negative early-stage breast cancer: Comparative analysis with standard clinicopathological factors. European Journal of Cancer. 2021;152:78-89.

187. Iles K, Roberson ML, Spanheimer P, Gallagher K, Ollila DW, Strassle PD, et al. The impact of age and nodal status on variations in oncotype DX testing and adjuvant treatment. NPJ breast cancer. 2022;8(1):1-8.

188. 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(17):4013-22.

189. Opdam M, van der Noort V, Kleijn M, Glas A, Mandjes I, Kleiterp S, et al. Limiting systemic endocrine overtreatment in postmenopausal breast cancer patients with an ultralow classification of the 70-gene signature. Breast cancer research and treatment. 2022:1-14.

190. Pu M, Messer K, Davies SR, Vickery TL, Pittman E, Parker BA, et al. based PAM50 signature and long-term breast cancer survival. Breast cancer research and treatment. 2020;179(1):197-206.

191. Sestak I, Dowsett M, Zabaglo L, Lopez-Knowles E, Ferree S, Cowens JW, et al. Factors predicting late recurrence for estrogen receptor–positive breast cancer. Journal of the National Cancer Institute. 2013;105(19):1504-11.

192. Buus R, Sestak I, Kronenwett R, Denkert C, Dubsky P, Krappmann K, et al. Comparison of EndoPredict and EPclin with oncotype DX recurrence score for prediction of risk of distant recurrence after endocrine therapy. Journal of the National Cancer Institute. 2016;108(11):djw149.

193. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with onco type DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. Journal of Clinical Oncology. 2013;31(22):2783-90.

194. Mamounas EP, Tang G, Fisher B, Paik S, Shak S, Costantino JP, et al. Association between the 21-gene recurrence score assay and risk of locoregional recurrence in node-negative, estrogen receptor–positive breast cancer: results from NSABP B-14 and NSABP B-20. Journal of Clinical Oncology. 2010;28(10):1677.

195. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. New England Journal of Medicine. 2004;351(27):2817-26.

196. Stemmer SM, Steiner M, Rizel S, Soussan-Gutman L, Ben-Baruch N, Bareket-Samish A, et al. Clinical outcomes in patients with node-negative breast cancer treated based on the recurrence score results: evidence from a large prospectively designed registry. NPJ breast cancer. 2017;3(1):1-8.

197. 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(1):1-9.

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

199. Drukker CA, Bueno-de-Mesquita J, Retèl VP, van Harten WH, van Tinteren H, Wesseling J, et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. International journal of cancer. 2013;133(4):929-36.

200. Drukker CA, Nijenhuis MV, Bueno-de-Mesquita J, Retèl VP, van Harten WH, van Tinteren H, et al. Optimized outcome prediction in breast cancer by combining the 70-gene signature with clinical risk prediction algorithms. Breast cancer research and treatment. 2014;145(3):697-705.

201. Esserman LJ, Yau C, Thompson CK, Van't Veer LJ, Borowsky AD, Hoadley KA, et al. Use of molecular tools to identify patients with indolent breast cancers with ultralow risk over 2 decades. JAMA oncology. 2017;3(11):1503-10.

202. van't Veer LJ, Yau C, Yu NY, Benz CC, Nordenskjöld B, Fornander T, et al. Tamoxifen therapy benefit for patients with 70-gene signature high and low risk. Breast cancer research and treatment. 2017;166(2):593-601.

203. Van De Vijver MJ, He YD, Van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. New England Journal of Medicine. 2002;347(25):1999-2009.

204. 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 OnlyPrediction of Recurrence in ER+, HER2− Breast Cancer. Clinical Cancer Research. 2019;25(13):3865-72.

205. Filipits M, Nielsen TO, Rudas M, Greil R, Stöger H, Jakesz R, et al. The PAM50 Risk-of-Recurrence Score Predicts Risk for Late Distant Recurrence after Endocrine Therapy in Postmenopausal Women with Endocrine-Responsive Early Breast CancerPAM50 ROR Score and Late Distant Recurrence in Breast Cancer. Clinical Cancer Research. 2014;20(5):1298-305.

206. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. Annals of oncology. 2014;25(2):339-45.

207. Lænkholm A-V, Jensen M-B, Eriksen JO, Rasmussen BB, Knoop AS, Buckingham W, et al. PAM50 risk of recurrence score predicts 10-year distant recurrence in a comprehensive Danish cohort of postmenopausal women allocated to 5 years of endocrine therapy for hormone receptor–positive early breast cancer. Journal of Clinical Oncology. 2018;36(8):735-40.

208. Sestak I, Cuzick J, Dowsett M, Lopez-Knowles E, Filipits M, Dubsky P, et al. Prediction of late distant recurrence after 5 years of endocrine treatment: a combined analysis of patients from the Austrian breast and colorectal cancer study group 8 and arimidex, tamoxifen alone or in combination randomized trials using the PAM50 risk of recurrence score. Journal of Clinical Oncology. 2015;33(8):916-22.

209. Liu S, Chapman JA, Burnell MJ, Levine MN, Pritchard KI, Whelan TJ, et al. Prognostic and predictive investigation of PAM50 intrinsic subtypes in the NCIC CTG MA.21 phase III chemotherapy trial. Breast cancer research and treatment. 2015;149(2):439-48.

210. Ohnstad HO, Borgen E, Falk RS, Lien TG, Aaserud M, Sveli MAT, et al. Prognostic value of PAM50 and risk of recurrence score in patients with early-stage breast cancer with long-term follow-up. Breast Cancer Research. 2017;19(1):120.

211. 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(2):213-22.

212. Cheng R, Kong X, Wang X, Fang Y, Wang J. Oncotype DX Breast Recurrence Score Distribution and Chemotherapy Benefit Among Women of Different Age Groups With HR-Positive, HER2-Negative, Node-Negative Breast Cancer in the SEER Database. Frontiers in Oncology. 2020;10:1583.

213. Cheng R, Wang Z, Kong X, Wang J, Fang Y, Qi L. Factors associated with chemotherapy benefit in breast cancer patients with midrange Oncotype DX breast recurrence scores. Cancer letters. 2021;503:213-9.

214. Iorgulescu JB, Freedman RA, Lester SC, Mittendorf EA, Brock JE. 21-Gene Recurrence Score Adds Significant Value for Grade 3 Breast Cancers: Results From a National Cohort. JCO precision oncology. 2019;3.

215. Wang F, Reid S, Zheng W, Pal T, Meszoely I, Mayer IA, et al. Sex Disparity Observed for Oncotype DX Breast Recurrence Score in Predicting Mortality Among Patients with Early Stage ER-Positive Breast Cancer. Clinical Cancer Research. 2020;26(1):101-9.

216. Weiser R, Polychronopoulou E, Haque W, Hatch SS, He J, Qiu S, et al. Prognosis and Chemotherapy Use in Breast Cancer Patients with Multiple Lymphatic Micrometastases: An NCDB Analysis. Annals of Surgical Oncology. 2021;28(13):8717-27.

217. Albanell J, Gonzalez A, Ruiz-Borrego M, Alba E, Garcia-Saenz J, Corominas J, et al. Prospective transGEICAM study of the impact of the 21-gene Recurrence Score assay and traditional clinicopathological factors on adjuvant clinical decision making in women with estrogen receptor-positive (ER+) node-negative breast cancer. Annals of Oncology. 2012;23(3):625-31.

218. Albanell J, Svedman C, Gligorov J, Holt SD, Bertelli G, Blohmer J-U, et al. Pooled analysis of prospective European studies assessing the impact of using the 21-gene Recurrence Score assay on clinical decision making in women with oestrogen receptor–positive, human epidermal growth factor receptor 2–negative early-stage breast cancer. European Journal of Cancer. 2016;66:104-13.

219. Bundred N. Erratum to "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 43 (5)(2017) 931–937]. European Journal of Surgical Oncology. 2018;44(1):194.

220. Dieci MV, Guarneri V, Giarratano T, Mion M, Tortora G, De Rossi C, et al. First prospective multicenter Italian study on the impact of the 21-gene recurrence score in adjuvant clinical decisions for patients with ER positive/HER2 negative breast cancer. The oncologist. 2018;23(3):297-305.

221. Eiermann W, Rezai M, Kümmel S, Kühn T, Warm M, Friedrichs K, et al. The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. Annals of oncology. 2013;24(3):618-24.

222. 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 (Edinburgh, Scotland). 2020;52:1-7.

223. Dieci MV, Guarneri V, Zustovich F, Mion M, Morandi P, Bria E, et al. Impact of 21-Gene Breast Cancer Assay on Treatment Decision for Patients with T1-T3, N0-N1, Estrogen Receptor-Positive/Human Epidermal Growth Receptor 2-Negative Breast Cancer: Final Results of the Prospective Multicenter ROXANE Study. Oncologist. 2019;24(11):1424-31.

224. 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(1):47.

225. Pérez Ramírez S, Del Monte-Millán M, López-Tarruella S, Martínez Jáñez N, Márquez-Rodas I, Lobo Samper F, et al. Prospective, multicenter study on the economic and clinical impact of gene-expression assays in early-stage breast cancer from a single region: the PREGECAM registry experience. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2020;22(5):717-24.

226. Kuijer A, Straver M, den Dekker B, Van Bommel A, Elias SG, Smorenburg CH, et al. Impact of 70-gene signature use on adjuvant chemotherapy decisions in patients with estrogen receptor–positive early breast cancer: results of a prospective cohort study. Journal of Clinical Oncology. 2017;35(24):2814-9.

227. Penault-Llorca F, Kwiatkowski F, Arnaud A, Levy C, Leheurteur M, Uwer L, et al. Decision of adjuvant chemotherapy in intermediate risk luminal breast cancer patients: A prospective multicenter trial assessing the clinical and psychological impact of EndoPredict® (EpClin) use (UCBG 2-14). Breast (Edinburgh, Scotland). 2020;49:132-40.

228. Hequet D, Callens C, Gentien D, Albaud B, Mouret-Reynier M-A, Dubot C, et al. Prospective, multicenter French study evaluating the clinical impact of the Breast Cancer Intrinsic Subtype-Prosigna® Test in the management of early-stage breast cancers. PloS one. 2017;12(10):e0185753.

229. Wuerstlein R, Sotlar K, Gluz O, Otremba B, von Schumann R, Witzel I, et al. The West German Study Group Breast Cancer Intrinsic Subtype study: a prospective multicenter decision impact study utilizing the Prosigna assay for adjuvant treatment decision-making in estrogen-receptor-positive, HER2-negative early-stage breast cancer. Current medical research and opinion. 2016;32(7):1217-24.

230. Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L, et al. Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-

negative breast cancer. A GEICAM/9906 sub-study. Breast cancer research and treatment. 2016;156(1):81-9.

231. Vallon-Christersson J, Häkkinen J, Hegardt C, Saal LH, Larsson C, Ehinger A, et al. Cross comparison and prognostic assessment of breast cancer multigene signatures in a large population-based contemporary clinical series. Scientific reports. 2019;9(1):1-16.

232. Gluz O, Nitz UA, Christgen M, Kates RE, Shak S, Clemens M, et al. West German Study Group Phase III PlanB Trial: first prospective outcome data for the 21-gene recurrence score assay and concordance of prognostic markers by central and local pathology assessment. Journal of Clinical Oncology. 2016.

233. Mamounas EP, Liu Q, Paik S, Baehner FL, Tang G, Jeong J-H, et al. 21-gene recurrence score and locoregional recurrence in node-positive/ER-positive breast cancer treated with chemo-endocrine therapy. JNCI: Journal of the National Cancer Institute. 2017;109(4).

234. Penault-Llorca F, Filleron T, Asselain B, Baehner FL, Fumoleau P, Lacroix-Triki M, et al. The 21-gene Recurrence Score® 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(1):1-10.

235. Roberts MC, Miller DP, Shak S, Petkov VI. Breast cancer-specific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX Recurrence Score results in the SEER database. Breast cancer research and treatment. 2017;163(2):303-10.

236. King TA, Lyman JP, Gonen M, Voci A, De Brot M, Boafo C, et al. Prognostic impact of 21-gene recurrence score in patients with stage IV breast cancer: TBCRC 013. Journal of Clinical Oncology. 2016;34(20):2359.

237. Woodward WA, Barlow WE, Jagsi R, Buchholz TA, Shak S, Baehner F, et al. Association between 21-gene assay recurrence score and locoregional recurrence rates in patients with node-positive breast cancer. JAMA oncology. 2020;6(4):505-11.

238. Collin LJ, Yan M, Jiang R, Ward KC, Crawford B, Torres MA, et al. Oncotype DX recurrence score implications for disparities in chemotherapy and breast cancer mortality in Georgia. NPJ breast cancer. 2019;5(1):1-7.

239. Nitz U, Gluz O, Clemens M, Malter W, Reimer T, Nuding B, et al. West German Study PlanB trial: adjuvant four cycles of epirubicin and cyclophosphamide plus docetaxel versus six cycles of docetaxel and cyclophosphamide in HER2-negative early breast cancer. Journal of Clinical Oncology. 2019;37(10):799-808.

240. Dubsky P, Van't Veer L, Gnant M, Rudas M, Bago-Horvath Z, Greil R, et al. A clinical validation study of MammaPrint in hormone receptor-positive breast cancer from the Austrian Breast and Colorectal Cancer Study Group 8 (ABCSG-8) biomarker cohort. ESMO open. 2021;6(1):100006.

241. Lopes Cardozo J, Drukker C, Schmidt M, van't Veer L, Glas A, Witteveen A, et al. Outcome of patients with an ultralow risk 70-gene signature in the MINDACT trial. Wolters Kluwer Health; 2021.

242. Fitzal F, Filipits M, Rudas M, Greil R, Dietze O, Samonigg H, et al. The genomic expression test EndoPredict is a prognostic tool for identifying risk of local recurrence in postmenopausal endocrine receptor-positive, her2neu-negative breast cancer patients randomised within the prospective ABCSG 8 trial. British journal of cancer. 2015;112(8):1405-10.

243. Jensen M-B, Lænkholm A-V, Nielsen TO, Eriksen JO, Wehn P, Hood T, et al. The Prosigna gene expression assay and responsiveness to adjuvant cyclophosphamide-based chemotherapy in premenopausal high-risk patients with breast cancer. Breast Cancer Research. 2018;20(1):1-10.

244. Fitzal F, Filipits M, Fesl C, Rudas M, Greil R, Balic M, et al. PAM-50 predicts local recurrence after breast cancer surgery in postmenopausal patients with ER+/HER2–disease: results from 1204 patients in the randomized ABCSG-8 trial. British Journal of Surgery. 2021;108(3):308-14.

245. Weiser R, Haque W, Polychronopoulou E, Hatch SS, Kuo YF, Gradishar WJ, et al. The 21-gene recurrence score in node-positive, hormone receptor-positive, HER2-negative breast cancer: a cautionary tale from an NCDB analysis. Breast cancer research and treatment. 2021;185(3):667-76.

246. Sestak I, Martín M, Dubsky P, Kronenwett R, Rojo F, Cuzick J, et al. Prediction of chemotherapy benefit by EndoPredict in patients with breast cancer who received adjuvant endocrine therapy plus chemotherapy or endocrine therapy alone. Breast cancer research and treatment. 2019;176(2):377-86.

247. Voelker H-U, Frey L, Strehl A, Weigel M. Practical consequences resulting from the analysis of a 21-multigene array in the interdisciplinary conference of a breast cancer center. International Journal of Breast Cancer. 2018;2018.

248. Pestalozzi BC, Tausch C, Dedes KJ, Rochlitz C, Zimmermann S, von Moos R, et al. Adjuvant treatment recommendations for patients with ER-positive/HER2-negative early breast cancer by Swiss tumor boards using the 21-gene recurrence score (SAKK 26/10). BMC cancer. 2017;17(1):1-10.

249. Del Prado PM, Alvarez-López I, Domínguez-Fernández S, Plazaola A, Ibarrondo O, Galve-Calvo E, et al. Clinical and economic impact of the 21-gene recurrence score assay in adjuvant therapy decision making in patients with early-stage breast cancer: pooled analysis in 4 Basque Country university hospitals. Clinicoeconomics and Outcomes Research: CEOR. 2018;10:189.

250. Curtit E, Vannetzel J-M, Darmon J-C, Roche S, Bourgeois H, Dewas S, et al. Results of PONDx, a prospective multicenter study of the Oncotype DX® breast cancer assay: Real-life utilization and decision impact in French clinical practice. The Breast. 2019;44:39-45.

251. Kuchel A, Robinson T, Comins C, Shere M, Varughese M, Sparrow G, et al. The impact of the 21-gene assay on adjuvant treatment decisions in oestrogen receptor-positive early breast cancer: a prospective study. British journal of cancer. 2016;114(7):731-6.

252. Cusumano P, Generali D, Ciruelos E, Manso L, Ghanem I, Lifrange E, et al. European inter-institutional impact study of MammaPrint. The Breast. 2014;23(4):423-8.

253. Wuerstlein R, Kates R, Gluz O, Grischke E, Schem C, Thill M, et al. Strong impact of MammaPrint and BluePrint on treatment decisions in luminal early breast cancer: results of the WSG-PRIMe study. Breast cancer research and treatment. 2019;175(2):389-99.

254. Mokbel K, Wazir U, Wazir A, Kasem A, Mokbel K. The Impact of EndoPredict Clinical Score on Chemotherapy Recommendations in Women with Invasive ER+/HER2− Breast Cancer Stratified as Having Moderate or Poor Prognosis by Nottingham Prognostic Index. Anticancer Research. 2018;38(8):4747-52.

255. Mokbel K, Wazir U, Chehade HEH, Manson A, Choy C, Moye V, et al. A Comparison of the performance of EndoPredict clinical and NHS PREDICT in 120 patients treated for ER-positive breast cancer. Anticancer Research. 2017;37(12):6863-9.

256. Ettl J, Klein E, Hapfelmeier A, Grosse Lackmann K, Paepke S, Petry C, et al. Decision impact and feasibility of different ASCO-recommended biomarkers in early

breast cancer: prospective comparison of molecular marker EndoPredict and protein marker uPA/PAI-1. PloS one. 2017;12(9):e0183917.

257. Fallowfield L, Matthews L, May S, Jenkins V, Bloomfield D. Enhancing decision-making about adjuvant chemotherapy in early breast cancer following EndoPredict testing. Psycho-Oncology. 2018;27(4):1264-9.

258. Bartlett JM, Bayani J, Marshall A, Dunn JA, Campbell A, Cunningham C, et al. Comparing Breast Cancer Multiparameter Tests in the OPTIMA Prelim Trial: No Test Is More Equal Than the Others. Journal of the National Cancer Institute. 2016;108(9).

259. Vieira AF, Schmitt F. An update on breast cancer multigene prognostic tests—emergent clinical biomarkers. Frontiers in medicine. 2018;5:248.

260. Mittempergher L, De Ronde JJ, Nieuwland M, Kerkhoven RM, Simon I, Th. Rutgers EJ, et al. Gene expression profiles from formalin fixed paraffin embedded breast cancer tissue are largely comparable to fresh frozen matched tissue. PloS one. 2011;6(2):e17163.

261. Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. Frontiers in medicine. 2017;4:227.

262. Baehner FL. The analytical validation of the Oncotype DX Recurrence Score assay. ecancermedicalscience. 2016;10.

263. Pearce DA, Arthur LM, Turnbull AK, Renshaw L, Sabine VS, Thomas JS, et al. Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies. Scientific reports. 2016;6(1):1-8.

264. Hillyar C, Rizki H, Abbassi O, Miles-Dua S, Clayton G, Gandamihardja T, et al. Correlation between Oncotype DX, PREDICT and the Nottingham prognostic index: implications for the Management of Early Breast Cancer. Cureus. 2020;12(4).

265. Precision Oncology. Biobanking: Resources & Infrastructure [Available from: https://www.precisiononcology.ie/research/researchareas/resourcesinfrastructure/p41/.

266. Boland M, Al-Maksoud A, Ryan ÉJ, Balasubramanian I, Geraghty J, Evoy D, et al. Value of a 21-gene expression assay on core biopsy to predict neoadjuvant chemotherapy response in breast cancer: systematic review and meta-analysis. British Journal of Surgery. 2021;108(1):24-31.

267. Solin LJ, Gray R, Baehner FL, Butler SM, Hughes LL, Yoshizawa C, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. Journal of the National Cancer Institute. 2013;105(10):701-10.

268. EUnetHTA. Joint Action 2 Work Package 8. HTA Core Model ® version 3.0 [Available from: https://www.htacoremodel.info/BrowseModel.aspx].

Appendix A

Table A.1 AMSTAR Checklist for Ontario HTA Series Gene Expression Profiling Tests for Early-Stage Invasive Breast Cancer: A Health Technology Assessment (2020)

AMSTAR2 Question		Answer	Reference
1.	Did the research questions and inclusion criteria for the review include the components of PICO?	YES	Clinical evidence (page 22) Participants Interventions - Head to head comparative studies including two or more of the included GEP tests Outcomes measures Economic evidence (page 51) Population Interventions Outcome measures Data Extraction- Methods (e.g., study design, analytic technique, perspective, time horizon, population, intervention[s], comparator[s])
2.	Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	NO	
3.	Did the review authors explain their selection of the study designs for inclusion in the review?	YES	Clinical evidence (page 26) "In general, we are more confident of estimates of prognosis from observational studies than from RCTs because the eligibility criteria for RCTs tend to be very specific and may exclude potentially relevant patients. Eligible patients may also decline to participate in an RCT for reasons related to their prognosis. Appropriate study designs for prognostic studies are reanalyses of RCTs or prospective studies, which was reflected among the prognostic studies included in our analysis."
4.	Did the review authors use a comprehensive literature search strategy?	YES	*Searched two databases:* (Clinical evidence page 21) MEDLINE and Embase. (Economic evidence) Ovid MEDLINE, Embase, Cochrane Database of Systematic Reviews, CRD Health Technology Assessment Database, Cochrane Central Register of Controlled Trials, and NHS Economic Evaluation Database.

			Provided key work and/or search strategy: Keywords used: Oncotype, OncotypeDX, Prosigna, EndoPredict, MammaPrint, gene profiling, genetic profiling, expression profiling (page 162). Multiple search strategies e.g. PRISMA clinical search strategy (page 24); PRISMA economic search strategy (page 52); Literature search- search strategy peer reviewed using the PRESS checklist (page 104). *Justified publication restrictions (e.g. language):* Eligibility criteria (page 21) – inclusion and exclusion criteria for example: English-language full-text publications Studies published between January 1, 2018, and November 28, 2018 Randomized controlled trials (RCTs), studies with prospectively enrolled nonrandomized (cohort) patients, and prospectively collected tumour specimens Retrospective analyses of RCTs or studies with prospectively enrolled nonrandomized (cohort) patients and prospectively collected tumour specimens *Searched reference lists/ bibliographies of included studies:* Results Clinical literature search (page 24) "identified seven studies from the literature search, 13 from reference lists and experts, and two from auto-alerts." *Searched trial/study registries:* Clinical evidence- Methods (page 21) "performed a targeted grey literature search of clinical trial registries." *Included/consulted content experts in the field:* Expert Consultation (page 20) "engaged with experts in the specialty areas of medical oncology, pathology, breast cancer surgery, health services research, and health economics." *Where relevant, searched for grey literature:* Clinical evidence- Methods (page 21) "performed a targeted grey literature search of clinical trial registries." *Economic evidence- Methods (page 50)* "performed a targeted grey literature search of health technology assessment agency websites, clinical trial and systematic review registries, and the Tufts Cost-Effectiveness Analysis Registry." Conducted search within 24 months of completion of the review: Clinical evidence search date November 28, 2018 – published March 2020 (approximately 15 months after search) Economic evidence search date: December 4, 2018 – published March 2020 (approximately 15 months after search)

5.	Did the review authors perform study selection in duplicate?	NO	"A single reviewer conducted an initial screening of titles and abstracts using Covidence and then obtained the full texts of studies that appeared eligible for review according to the inclusion criteria. A single reviewer then examined the full-text articles and selected studies eligible for inclusion." (page 23 & page 106)
6.	Did the review authors perform data extraction in duplicate?	NO	"A single reviewer conducted an initial screening of titles and abstracts using Covidence and then obtained the full texts of studies that appeared eligible for review according to the inclusion criteria. A single reviewer then examined the full-text articles and selected studies eligible for inclusion." (page 23 & page 106)
7.	Did the review authors provide a list of excluded studies and justify the exclusions?	NO	
8.	Did the review authors describe the included studies in adequate detail?	YES/ PARTIAL YES (?)	Appendices 3(?), 5, 9
9.	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	YES	"assessed the risk of bias of each included study using either the Cochrane Risk of Bias tool, Prediction model Risk Of Bias Assessment Tool (PROBAST), or Risk of Bias Assessment tool for Non-randomized Studies (RoBANS), depending on the type of study and outcome of interest, and the quality of the body of evidence according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group criteria." (page 3)
10.	Did the review authors report on the sources of funding for the studies included in the review?	NO	Characteristics of Included Studies, Clinical Evidence Review A nonspecific statement in the text, "Some studies were funded by test manufacturers." (Appendix 5-Table A7)
11.	If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	N/A	"After determining that a meta-analysis to provide an overall statistical summary of the effect estimate was inappropriate for a broad summary of the quantitative evidence on preferences, we chose a descriptive approach using text or tables." (page 106)
12.	If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	N/A	
13.	Did the review authors account for RoB in individual studies when	YES	The risk of bias of the included studies are presented in Tables A8, A9, and A10 (Appendix 5). In general, the studies were of low to moderate quality."

	interpreting/ discussing the results of the review?		
14.	Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	YES	"undertook a narrative summary of the results due to the heterogeneity of patient populations and the reported endpoints of outcomes within studies." (page 23)
15.	If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	N/A	
16.	Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	NO	"This report was developed by a multidisciplinary team from the Quality business unit at Ontario Health and the Canadian Agency for Drugs and Technologies in Health (CADTH). From Ontario Health (Quality), the clinical epidemiologist was Myra Wang, the health economist was Yuan Zhang, the patient and public partnership analyst was Ammara Shafique, and the medical librarian was Melissa Walter" (page 2)

Key: AMSTAR=A MeaSurement Tool to Assess systematic Reviews; HTA=Health Technology Assessment; PICO=Population Intervention Comparison Outcome; RoB=Risk of Bias.

Table A.2 Characteristics of included studies

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
Oncotype DX®											
Kalinsky et al, 2021[70]	RCT (RxPONDER)	US, Canada, Mexico, Colombia, Ireland, France, Spain, South Korea, Saudi Arabia	2011 to 2017	5,018	Low/int: RS 0-25	Low/int: 5,018 (100)	Median: 58 Range: 18-88 Pre: 33% Post: 67%	100% HR+ 100% HER2-	100% LN (1-3)	ET, CET	Yes
Sparano et al, 2019[67]	RCT (TAILORx)	US, Australia, Canada, Ireland, New Zealand, Peru	2006 to 2010	9,427	Low: RS 0-10 Int: RS 11-25 High: RS 26-100	Low: 1,572 (17) Int: 6,496 (69) High: 1,359 (14)	≤50: 31% >50: 69% Pre/Post: NR	100% HR+ 100% HER2-	100% LN-	No CT, CT	Yes
Nitz et al, 2019[239]	Retrospective analysis of RCT (WSG-Plan B)	Germany	2009 to 2011	2,449	Low/int: RS 0-25 High: RS 26-100	Low/int: 1,413 (58) High: 518 (21) NA: 518 (21)	Range: 18-75 Pre: 35% Post: 57% NA: 8%	82% HR+ 100% HER2-	59% LN- 41% LN (1-3)	CT	Yes
Pece et al, 2022[176]	Retrospective analysis of RCT (TRANSATAC)	UK	1996 to 2000	776	Low/int: RS 0-26; High: RS 27-100	Low/int: 627 (81) High: 149 (19)	Median: 64 IQR: 58-71 Post: 100%	100% HR+ 100% HER2-	76.6% LN- 23.4% LN (1-3)	ET	Yes
Sparano et al, 2020[184]	Retrospective analysis of RCT (TAILORx)	US, Australia, Canada, Ireland, New Zealand, Peru	2006 to 2010	9,719	High: RS 26-100	High: 1,389 (14)	Mean: 55.8±9.4 Median: 56 Range: 23-75 Pre: 29% Post: 71%	100% HR+ 100% HER2-	100% LN-	CT	Yes
Woodward et al, 2020[237]	Retrospective analysis of RCT (SWOG S8814)	US	1989 to 1995	316	Low: RS 0-17 Int/ high: RS 18-100	Low: 38% Int/ high: 62%	Mean: 60.4 Range: 44-81 Post: 100%	96.5% HR+, 88.0% HER2-	100% LN+	ET, CET	Yes
Cognetti et al, 2021[224]	Prospective observational study (PONDx)	Italy	2016 to 2017	1,738	Low: RS 0-17 Int: RS 18-30	Low: 987 (57) Int: 588 (34) High: 163 (9)	<50: 20% ≥50: 80% Pre: 36% Peri: 8%	100% HR+ 100% HER2-	68.9% LN- 24.6% LN (1-3)	ET, CET	Yes

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
					High: RS 31-100		Post: 56%				
Dieci et al, 2019[223]	Prospective observational study (ROXANE)	Italy	2017 to 2018	251	Low: RS 0-10 Int: RS 11-25 High: RS: 26-100	Low: 63 (25) Int: 143 (57) High: 45 (18)	Median: 57 Range: 28-79 Pre: 42% Post: 58%	100% HR+ 100% HER2-	62% LN- 38% LN (1-3)	ET, CET	Yes
Zambelli et al, 2020[222]	Prospective observational study (BONDX)	Italy	2017 to 2018	394	Low: RS 0-17 Int: RS 18-30 High: RS 31-100	Low: 60% Int: 34% High: 6%	Median: 62 Range: 34-80 Pre: 27% Post: 73%	100% HR+ 100% HER2-	67.8% LN+ 32.2% LN-	ET, CET	Yes
Collin et al, 2019[238]	Retrospective analysis of the Georgia Cancer Registry	US	2010 to 2014	5,750	Traditional cutoffs Low: RS 0-17 Int: RS 18-30 High: RS 31-100 TAILORx cutoffs Low: RS 0-10 Int: RS 11-25 High: RS: 26-100	Traditional cutoffs Low: 3,266 (57) Int: 1,925 (33) High: 559 (10) TAILORx cutoffs NR	<50: 20% ≥50: 80% Pre/Post: NR	100% HR+ 100% HER2-	90% LN-, 10% LN (1-3)	ET, CET	Partially
Lynch et al, 2021[186]	Retrospective analysis of The Breast Cancer Bank of Tissue	Ireland	2016 to 2019	404	Low: age ≤50y & RS 0-15 or age >50y & RS 0-25 High: age ≤50y & RS 16-100 or age >50y & RS 26-100	Low: 235 (58) High: 169 (42)	Median: 54 Range: 20-75 Pre/Post: NR	100% HR+ 100% HER2-	100% LN-	ET, CET	Partially

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
Iles et al, 2022[187]	Retrospective analysis of the National Cancer Database	US	2010 to 2017	255,971	Low: RS 0-10 Int: RS 11-25 High: RS 26-100	NR	<40: 4% 40-69: 79% ≥70: 17% Pre/Post: NR	100% HR+ 100% HER2-	83% LN- 16% LN+	No CT, CT	No
Iorgulescu et al, 2019[214]	Retrospective analysis of the National Cancer Database	US	2010 to 2015	30,864	Low: RS 0-17 Int: RS 18-30 High: RS 31-100	Low: 3,591 (17) Int: 5,304 (25) High: 4,332 (21) No RS: 7,586 (36)	<50: 23% ≥50: 77% Pre/Post: NR	100% HR+ 100% HER2-	100% LN-	ET, CET	Partially
Wang et al, 2020[215]	Retrospective analysis of the National Cancer Database	US	2010 to 2014	848 males 110,898 females	Traditional cutoffs Low: RS 0-17 Int: RS 18-30 High: RS 31-100 TAILORx cutoffs Low: RS 0-10 Int: RS 11-25 High: RS: 26-100	Male, traditional Low: 496 (58) Int: 260 (31) High: 92 (11) Male, TAILORx Low: 294 (35) Int: 406 (48) High: 148 (17) Female, traditional Low: 65,935 (59) Int: 36,174 (33) High: 8,789 (8) Female, TAILORx Low: 25,929 (23) Int: 68,882 (62) High: 16,087 (15)	Male Mean: 61.9±10.4 Range: 26-88 Female Mean: 58.3±10.5 Range: 18-90	Male 100% HR+ 100% HER2- Female 100% HR+ 100% HER2-	Male 81% LN- 19% LN+ Female 83% LN- 17% LN+	ET, CET	Partially
Weiser et al, 2021a[216]	Retrospective analysis of the National Cancer Database	US	2010 to 2016	486,800	Low: RS 0-10 Int: RS 11-25 High: RS 26-100	NR	Mean: 62.2 ≤50: 19% >50: 81% Pre/Post: NR	100% HR+ 100% HER2-	84% LN- 11% LN (1)	No CT, CT	No
Weiser et al, 2021b[245]	Retrospective analysis of the National	US	2010 to 2016	28,591	Low: RS 0-11	NR	Mean: 59.8 ≤50: 22% >50: 78%	100% HR+	100% LN (1-3)	No CT, CT	No

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
	Cancer Database				Int: RS 12-17 High: RS 18-100		Pre/Post: NR	100% HER2-			
Cheng et al, 2020[212]	Retrospective analysis of the SEER database	US	2010 to 2015	49,539	Low: RS 0-10 Int: RS 11-25 High: RS 26-100	Low: 11,164 (23) Int: 31,731 (64) High: 6,644 (13)	Median: 58 Range: 18-88 ≤50: 27% >50: 73% Pre/Post: NR	100% HR+ 100% HER2-	100% LN-	No CT/ unknown, CT	No
Cheng et al, 2021[213]	Retrospective analysis of the SEER database	US	2010 to 2015	31,731	RS 11-15 RS 16-20 RS 21-25	RS 11-15: 13,052 (41) RS 16-20: 11,868 (37) RS 21-25: 6,811 (21)	Median: 57 Range: 18-92 Pre/Post: NR	100% HR+ 100% HER2-	100% LN-	ET, CET	No
Gao et al, 2020[170]	Retrospective analysis of the SEER database	US	2010 to 2018	2,958	Low clinical risk (pT1 and grade I-II disease), high RS risk (RS ≤25); High clinical risk (pT2 & grade II-III disease), low RS risk (RS >25)	C-low, RS-high: 762 (26) C-high, RS-low: 2,196 (74)	Median: 59 Range: 19-91 Pre/Post: NR	100% HR+ 100% HER2-	NR	NR	No
Kantor et al, 2021[102]	Retrospective analysis of the SEER database	US	2010 to 2015	60,886	Low: RS 0-10 Int: RS 11-25 High: RS: 26-100	Low: 13,570 (22) Int: 39,240 (64) High: 8,076 (13)	<50: 27% ≥50: 83% Pre/Post: NR	100% HR+ 100% HER2-	71% LN- 24% LN (1-3)	ET, CET	Yes
Ma et al, 2021[171]	Retrospective analysis of the SEER database	US	2010 to 2016	5,054	RS 26-30 RS 31-100	RS 26-30: 2,561 (51); RS 31-100: 2,493 (49)	NR Pre/Post: NR	100% HR+ 100% HER2-	100% LN-	No CT, CT	No
Poorvu et al, 2019[185]	Retrospective analysis of Young Women's	US, Canada	2006 to 2016	509	Traditional cutoffs Low: RS 0-17	Traditional cutoffs Low: 199 (39) Int: 211 (41)	Median: 37 Range: 18-41 Pre/Post: NR	100% HR+ 100% HER2-	59% LN- 41% LN+	No CT, CT	Yes

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
	Breast Cancer Study				Int: RS 18-30 High: RS 31-100 TAILORx cutoffs Low: RS 0-10 Int: RS 11-25 High: RS: 26-100	High: 99 (19) TAILORx cutoffs Low: 54 (11) Int: 306 (60) High: 149 (29)					
MammaPrint®											
Knauer et al (2010)[178]	Retrospective analyses of pooled databases	Netherlands, Italy	NR	541	Low: 0-1 High -1-0	Low: 252 (47) High: 289 (53)	≤50: 43% >50: 57% Pre/Post: NR	90% ER+ 69% PR+ 89% HER2-	49% LN- 51% LN (1-3)	ET, CET	Yes
Mook et al (2008)[179]	Retrospective analysis of cohort	Netherlands, Italy	1994 to 2001	241	Low: 0-1 High -1-0	Low: 99 (41) High: 142 (59)	<50: 52% ≥50: 48% Pre/Post: NR	79% ER+ 63% PR+ 82% HER2-	100% LN (1-3)	ET, CET	Yes
Piccart et al, 2021[177]	RCT (MINDACT)	Netherlands, France, Germany, Belgium, Spain, Italy, UK, Slovenia, Switzerland	2007 to 2011	6,693	High clinical risk (assessed by Adjuvant! Online) and low genomic risk (MammaPrint® score 0-1) Low clinical risk and high genomic risk (MammaPrint® score -1-0)	C-high / G-low: 699 (57) C-low / G-high: 534 (43)	≤50: 33% >50: 67% Pre/Post: NR	100% HR+ 100% HER2-	79% LN– 21% LN (1-3)	ET, CET	Yes
Lopes Cardozo et al, 2022[241]	Retrospective analysis of RCT (MINDACT)	Belgium, France, Germany, Italy,	2007 to 2011	6,693	Ultralow: MPI> 0.355 Low: 0< MPI≤ 0.355	Ultralow: 1,000 (15) Low: 3,295 (49) High: 2,398 (36)	≤50: 33% >50: 67% Pre/Post: NR	100% HR+ 100% HER2-	79% LN– 21% LN (1-3)	ET, CET	Yes

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
		Netherlands, Slovenia, Spain, Switzerland, UK			High: MPI≤0						
Dubsky et al, 2021[240]	Retrospective analysis of RCT (ABCSG-8)	Austria	1996 to 2004	658	Low: 0-1 High -1-0	Low: 512 (78) High: 146 (22)	Mean: 63.0 Post: 100%	100% HR+ 100% HER2-	69% LN- 28% LN (1-3) 3.0% LN (>3)	ET	Yes
Opdam et al, 2022[189]	Retrospective analysis of RCT (IKA)	Netherlands	1982 to 1994	135	Ultralow: MPI> 0.355 Low: (0< MPI≤ 0.355 High: MPI≤0	Ultralow: 23 (17) Low: 59 (44) High: 53 (39)	<55: 7% 55-64: 45% ≥65: 48% Post: 100%	100% HR+ 100% HER2-	59% LN- 41% LN (1-3)	ET, No ET	Yes
EndoPredict®											
Penault-Llorca et al. 2020[227]	Prospective observational study (ADENDOM)	France	2015 to 2016	201	Low: <3.32867 High: ≥3.32867	Low: 135 (67) High: 66 (33)	Median: 59 Range: 23-81 Pre/Post: NR	100% HR+ 100% HER2-	90.5% LN- 9.5% LN(1mi)	ET, CET	Yes
Prosigna®											
Fitzal et al, 2021[244]	Retrospective analysis of RCT (ABCSG-8)	Austria	1996 to 2004	1,034	Low: ROR <57 High: ROR ≥57	Low: 765 (74); High: 269 (26)	<65: 61% ≥65: 39% Post: 100%	100% HR+ 100% HER2-	70% LN- 30% LN (1-3)	ET	Yes
Pu et al, 2020[190]	Retrospective analysis of RCT (WHEL)	US	1995 to 2000	1,253	Low Medium High	NR	Median: 50 Range: 27-70 Pre: 22% Post: 78%	66.8% HR+ & HER2-	56% LN- 44% LN+	ET, CET	Yes
Mixed tests											
Ibraheem et al, 2020[188]	Retrospective analysis of the National Cancer Database	US	2010 to 2016	149,404 (144,357 received Oncotype DX®, 5,042 pairs of propensity matched	Oncotype Low 0-10; Int 11-25; High 26-100 MammaPrint® Low: 0-1	Oncotype DX® Low: 1,336 (16) Int: 15,625 (35); High: 15,513 (81) MammaPrint® Low, 2,908 (58);	Oncotype DX® <50: 22% ≥50: 78% Pre/Post: NR MammaPrint® <50: 24% ≥50: 76%	100% HR+ 100% HER2-	MammaPrint® 77% LN- 23% LN (1-3) Oncotype DX®	ET, CET	Partially

Author, year	Study design	Location	Study period	Total sample size	Risk category cut-offs	Risk categories, n (%)	Age, years Menopausal status (pre/ peri/post), %	Hormone receptor status, %	Lymph node status	Treatment	CoI[a]
				patients received Oncotype DX® or MammaPrint®)	High: -1-0	High risk, 2,134 (42)	Pre/Post: NR		83% LN-, 17% LN (1-3)		
Pérez Ramírez et al, 2020[225]	Prospective observational study (PREGECAM)	Spain	2012 to 2014	907 (467 received MammaPrint® 404 received Oncotype DX®)	MammaPrint® Low: 0-1 High: -1-0 Oncotype DX® Low: RS 0-17 Int: RS 18-30 High: RS 31-100	MammaPrint® Low: 297 (64) High: 170 (36) Oncotype DX® Low: 238 (54) Int: 168 (38) High: 34 (8)	Median: 54 Range: 18-77 Pre/Post: NR	100% HR+ 100% HER2-	MammaPrint® 100% LN- Oncotype DX® 100% LN-	ET, CET	Partially
Vallon-Christersson et al, 2019[231]	Retrospective analysis of SCAN-B	Sweden	2010 to 2015	423 patient subsample (received both Prosigna® and Oncotype DX®)	Prosigna® Low Medium High Oncotype DX® Low Int High	Prosigna® Low: 130 (31) Medium: 92 (22) High: 201 (48) Oncotype DX® Low: 129 (30) Int: 136 (32) High: 158 (37)	NR Pre/Post: NR	100% HR+ 100% HER2-	Prosigna® 100% LN (1-3) Oncotype DX® 100% LN (1-3)	ET	No

Key: CET=chemoendocrinetherapy; CoI=conflicts of interest; ET=endocrine therapy; HER2-=human epidermal growth factor receptor 2 negative; HR+=hormone receptor positive; Int=intermediate; IQR=interquartile range; LN-=lymph node negative; LN (1-3)=lymph node positive (1-3 nodes); MPI=MammaPrint® Index; NR=not reported; RCT=randomised controlled trial; SEER=Surveillance, Epidemiology, and End Results; UK=United Kingdom; US=United States of America.

[a]Conflicts of interest recorded as yes when the study was funded or facilitated by, or if any authors reported links to, the manufacturers of the test(s) being studied, partially when the study was funded or facilitated by, or if any authors reported links to, other pharmaceutical or cancer diagnostic companies, and no when authors reported no conflicts of interests

Table A.3 Risk of bias among prognostic studies (PROBAST)

Author, year	Risk of Bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
Collin et al, 2019[238]	High	Low	Low	Low	Low	Low	Low	High	Low
Dubsky et al, 2021[240]	High	Low	Low	Low	Low	Low	Low	High	Low
Fitzal et al, 2021[244]	High	Low	Low	Low	Low	Low	Low	High	Low
Ibraheem et al, 2020[188]	High	Low	Low	Low	Low	Low	Low	High	Low
Iles et al, 2022[187]	High	Low	Low	Low	Low	Low	Low	High	Low
Kantor et al, 2021[102]	High	Low	Low	High	Low	Low	Low	High	Low
Lopes Cardozo et al, 2022[241]	High	Low	Low	Low	Low	Low	Low	High	Low
Lynch et al, 2021[186]	High	Low	Low	High	Low	Low	Low	High	Low
Mook et al (2008)[179]	High	Low	Low	High	Low	Low	Low	High	Low
Nitz et al, 2019[239]	High	Low	Low	High	Low	Low	Low	High	Low
Opdam et al, 2022[189]	High	Low	Low	Low	Low	Low	Low	High	Low
Pece 2022[176]	High	Low	Low	High	Low	Low	Low	High	Low
Poorvu et al, 2020[185]	High	Low	Low	High	Low	Low	Low	High	Low
Pu et al, 2020[190]	High	Low	Low	High	Low	Low	Low	High	Low
Sparano et al, 2020[184]	High	Low	Low	High	Low	Low	Low	High	Low
Vallon-Christersson et al, 2019[231]	High	Low	Low	High	Low	Low	Low	High	Low
Wang et al, 2020[215]	High	Low	Low	High	Low	Low	Low	High	Low

Woodward et al, 2020[237]	High	Low	Low	High	Low	Low	Low	High	Low

Key: PROBAST=Prediction Model Risk of Bias Assessment Tool.

Table A.4 Risk of bias among nonrandomized studies (RoBANS)

Author, year	Selection of participants	Confounding variables	Measurement of exposure	Blinding of outcome assessments	Incomplete outcome data	Selective outcome reporting
Cognetti et al, 2021[224]	High	High	Low	Low	Low	Low
Dieci et al, 2019[223]	High	High	Low	Low	Low	Low
Gao et al, 2020[170]	High	High	Low	Low	Low	High
Knauer et al (2010)[178]	High	High	Low	Low	Low	High
Penault-Llorca et al, 2020[227]	Low	High	Low	Low	Low	Low
Pérez Ramírez et al, 2020[225]	Low	High	Low	Low	Low	Low
Zambelli et al, 2020[222]	Low	High	Low	Low	Low	Low
Cheng et al, 2020[212]	High	High	Low	Low	High	High
Cheng et al, 2021[213]	High	High	Low	Low	High	High
Iorgulescu et al, 2019[214]	High	High	Low	Low	High	High
Ma et al, 2021[171]	High	High	Low	Low	High	High
Weiser et al, 2021a[216]	High	High	Low	Low	High	High
Weiser et al, 2021b[245]	High	High	Low	Low	High	High

Key: RoBANS=Risk of Bias Assessment for Nonrandomized Studies.

Table A.5 Risk of bias among randomized controlled trials (Cochrane Risk of Bias Tool)

Author, year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Incomplete outcome data	Selective reporting	Other bias	Overall
Kalinsky et al, 2021[70]	Unclear	Low	Low	Low	Low	Low	Unclear
Sparano et al, 2019[67]	Unclear	Low	Low	Low	Low	Low	Unclear
Piccart et al, 2021[177]	Unclear	Low	Low	Low	Low	Low	Unclear

Table A.6 Summary of findings including certainty of evidence for (LN+ and LN-) Prognostic studies assessing Oncotype DX® and *Prosigna®*.

Patients or population: LN+ and LN- Intervention: Oncotype DX Comparison: Prosigna®					
Outcome	**N participants**	**N studies and designs**	**Summary effect estimate**	**Interpretation of effect**	**GRADE**
Prognostic LN+					
Distant recurrence	681	Retrospective analysis of RCT n=3 (Sestak et al, 2013[191] (n=230) & Sestak et al, 2018[183] (n=183) and Dowsett et al, 2013[193] (89% of patients HER2) negative (n=268))	Prosigna® was more prognostic than Oncotype DX®. Likelihood ratios (p value) Prosigna® 0-5 years: 1.96 (.20) 5-10 years: 4.78 (.03) Oncotype DX® 0-5 years: 4.01 (.05) 5-10 years: 0.38 (.50) HR (95% CI) Prosigna®: 1.58 (1.16-2.15) Oncotype DX®: 1.39 (1.05-1.85) Proisgna®: 23.1 (0.001) Oncotype DX®: 9.1 (0.03)	Could not differentiate as no direct comparison data presented.	⊕ Very low[a]
Overall survival	423	Prospective observational study n=1 (Vallon-Christersson et al, 2019)[231]	Study indicates Prosigna® has greater prognostic ability than Oncotype Kaplan-Meier significance Prosigna® p-value=0.00001 Oncotype DX® p-value=0.005	Could not differentiate as no direct comparison data presented	⊕⊕[a,b] Low
Prognostic LN-					
Distant recurrence	1,945	Retrospective analyses of RCTs n=3 (Dowsett 2013 (n=739)[193], Sestak 2013[191] (n=615) and Sestak 2018[183] (n=591))	Greater prognostic information presented by Prosigna®. Agreement between studies. Likelihood ratio (p-value) Prosigna®	Could not differentiate as no direct comparison data presented	⊕[a] Very low

Patients or population: LN+ and LN- Intervention: Oncotype DX Comparison: Prosigna®					
Outcome	**N participants**	**N studies and designs**	**Summary effect estimate**	**Interpretation of effect**	**GRADE**
			0-5 year: 8.61 (0.008) 5-10 year: 13.85 (0.001) Oncotype DX® 0-5 year: 6.84 (0.008) 5-10 year: 6.84 (0.008) Prosigna® 53.7 (0.001204) Oncotype DX® 22.9 (0.001204) HR (95% CI) Prosigna®: 2.56 (1.96-3.35) Oncotype DX®: 1.69 (1.40-2.03)		
Explanations : [a] Downgrade one level due to imprecision. Study numbers very low. No direct statistical comparison of tests. [b] Downgrade one level due to high risk of bias. Number of samples were of insufficient quality.					

Table A.7 Summary of findings including certainty of evidence for (LN+ and LN-) Prognostic studies assessing Oncotype DX® and *EndoPredict®.*

Patients or population: LN+ and LN- Intervention: Oncotype DX® Comparison: EndoPredict®					
Outcome	**N participants**	**N studies and designs**	**Summary effect estimate**	**Interpretation of effect**	**GRADE**
Prognostic LN+					
Distant recurrence	183	Retrospective analysis of RCT n=1 (Sestak et al, 2018)[183]	With low certainty of evidence, EndoPredict® presented with greater prognostic ability than Oncotype DX®. HR (95% CI) EndoPredict®: 1.69 (1.29-2.22) Oncotype DX®: 1.39 (1.05-1.85)	Could not differentiate as no direct comparison data presented	⨁[ab] Very low
Prognostic LN-					
Distant recurrence	1,519	Retrospective analysis of RCT n=2 (Sestak et al, 2018[183] (n=591) and Buus et al, 2016[192] (n=928))	Overall prognostic ability for all tests greater in LN- population, with low certainty of evidence. EndoPredict® presented with greater prognostic ability than Oncotype DX®. HR (95% CI) EndoPredict®: 2.14 (1.71-2.68) Oncotype DX®: 1.69 (1.40-2.03) Likelihood ratio EndoPredict®: 139.3 Oncotype DX®: 29.1	Could not differentiate as no direct comparison data presented	⨁[ab] Very low

Explanations
[a] Downgrade one level due to imprecision. Study numbers very low. No direct statistical comparison of tests.
[b] Downgrade one level due to high risk of bias. Number of samples were of insufficient quality.

Table A.8 Summary of findings including certainty of evidence for (LN+ and LN-) Prognostic studies assessing EndoPredict® and *Prosigna®.*

Patients or population: LN+ and LN-
Intervention: EndoPredict®
Comparison: Prosigna®

Outcome	N participants	N studies and designs	Summary effect estimate	Interpretation of effect	GRADE
Prognostic LN+					
Distant recurrence	183	Retrospective analysis of RCT n=1 (Sestak et al, 2018)[183]	With low certainty of evidence, EndoPredict® presented with greater prognostic ability than Prosigna®. HR (95% CI) EndoPredict®: 1.69 (1.29-2.22) Prosigna®: 1.58 (1.16-2.15)	Could not differentiate as no direct comparison data presented	⨁[ab] Very low
Prognostic LN-					
Distant recurrence	591	Retrospective analysis of RCT n=1 (Sestak et al, 2018)[183]	Overall prognostic ability for all tests greater in LN- population, with low certainty of evidence. Prosigna® presented with greater prognostic ability than EndoPredict®. HR (95% CI) EndoPredict®: 2.14 (1.71-2.68) Prosigna®: 2.56 (1.96-3.35)	Could not differentiate as no direct comparison data presented	⨁[ab] Very low

Explanations
[a] Downgrade one level due to imprecision. Study numbers very low. No direct statistical comparison of tests.
[b] Downgrade one level due to high risk of bias. Number of samples were of insufficient quality.

Table A.9 Summary of findings including certainty of evidence for (LN+ and LN-) Prognostic studies assessing Oncotype DX® and *MammaPrint®*.

Patients or population LN+ LN- Intervention: Oncotype DX® Comparison: MammaPrint®					
Outcome	**N participants**	**N studies and designs**	**Summary effect estimate**	**Interpretation of effect**	**GRADE**
Prognostic LN-mixed					
Overall survival	5,042	Retrospective analysis of prospective database with propensity score matching n=1 (Ibraheem et al, 2020)[188]	Oncotype DX® and MammaPrint® presented with similar prognostic results following propensity score matching. Oncotype DX® showed weaker prognostic ability in ethnic minority patients, but this was not tested for MammaPrint®. **5-Year Risk (95% CI)** MammaPrint®: Low risk: 3.4% (2.4%-4.7%) High risk: 9.3% (7.4%-11.7%) Adjusted HR: 2.25 (1.56-3.25) Oncotype DX®: Low risk (RS 0-10): 4.7% (3.0%-7.4%) Intermediate risk: 5.2% (3.9%-6.8%) High risk (RS >25): 12.4% (9.1%-16.8%) Intermediate vs. low adjusted HR: 1.04 (0.66-1.62) High versus low adjusted HR: 1.81 (1.05-3.09)	No meaningful difference	⨁[ab] very low
Explanations: [a] Downgrade one level due to imprecision. Study numbers very low. No direct statistical comparison of tests. [b] Downgrade one level due to risk of bias.					

Published by the Health Information and Quality Authority (HIQA).

For further information please contact:
Health Information and Quality Authority
George's Court
George's Lane
Smithfield
Dublin 7
D07 E98Y

+353 (0)1 8147400
info@hiqa.ie
www.hiqa.ie

© Health Information and Quality Authority 2024