ONTARIO HEALTH TECHNOLOGY ASSESSMENT SERIES

Gene Expression Profiling Tests for Early-Stage Invasive Breast Cancer: A Health Technology Assessment

KEY MESSAGES

What Is This Health Technology Assessment About?

Breast cancer is a disease in which cells in the breast grow out of control. Treatment typically involves surgery, but there is a chance that the breast cancer can come back. One challenge is whether to recommend chemotherapy for breast cancer recurrence. While chemotherapy helps prevent cancer recurrence, it can cause negative side effects.

Gene expression profiling (GEP) tests analyze a sample of the breast cancer tissue to identify the presence or absence of certain genes in the cancer cell. This information may help physicians determine the likelihood that the cancer will return after surgery and can help guide decision-making about whether the patient may benefit from chemotherapy.

This health technology assessment looked at how safe and effective GEP tests are for people with early-stage invasive breast cancer. It looked at the cost-effectiveness and the budget impact of publicly funding GEP testing. It also looked at the experiences, preferences, and values of people with early-stage invasive breast cancer.

What Did This Health Technology Assessment Find?

Gene expression profiling tests can predict the recurrence of breast cancer in areas of the body other than the breast and patient survival. Some tests may also predict chemotherapy benefit. They also lead to changes in chemotherapy treatment decisions and generally increase physician confidence in treatment recommendations.

Compared with the current model of funding GEP tests through the out-of-country program, publicly funding GEP tests to be conducted in Ontario would cost an additional $1 million to $2 million annually, depending on how many additional people choose to receive the test.

Gene expression profiling tests are valued by people with breast cancer and physicians for the additional information they provide for treatment decision-making. Patients are satisfied with what they learn from GEP tests and feel their use can help reduce patients' decisional conflict and anxiety.

Published March 2020
Volume 20, Number 10

Ontario Health Quality

ACKNOWLEDGMENTS

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.

From CADTH, the qualitative review authors were Andrea Smith and Kelly Farrah.

The medical editor was Timothy Maguire. Others involved in the development and production of this report were Harrison Heft, Claude Soulodre, Kara Cowan, Elisabeth Smitko, Kathryn Schwarz, Sarah McDowell, Vivian Ng, Andrée Mitchell, Amy Lang, Nancy Sikich, and Irfan Dhalla.

We would like to thank the following individuals and organizations for lending their expertise to the development of this report:

- Yvonne Bombard, Li Ka Shing Institute of St. Michael’s Hospital
- Andrea Eisen, Sunnybrook Health Sciences Centre
- Ramzi Fattouh, St. Michael’s Hospital
- Yanping Gong, Kingston General Hospital
- Claire Holloway, Sunnybrook Health Sciences Centre
- Nicole Mittmann, Sunnybrook Health Sciences Centre
- Petros Pechlivanoglou, Hospital for Sick Children
- Aaron Pollett, Mount Sinai Hospital
- Maureen Trudeau, Sunnybrook Health Sciences Centre
- Ivan Tyono, Sunnybrook Health Sciences Centre
- William Wong, University of Waterloo
- Agendia N.V.
- Alberta’s Institute of Health Economics
- Ontario Health (Cancer Care Ontario) and McMaster University’s Program in Evidence-Based Care (PEBC)
- Genomic Health, Inc.
- Myriad Genetics, Inc.
- NanoString Technologies, Inc.

We also thank our lived experience participants, who generously gave their time to share their stories with us for this report.

Citation

Ontario Health (Quality). Gene expression profiling tests for early-stage invasive breast cancer: a health technology assessment. Ont Health Technol Assess Ser [Internet]. 2020 Mar;20(10):1–234. Available from: https://www.hqontario.ca/evidence-to-improve-care/health-technology-assessment/reviews-and-recommendations/gene-expression-profiling-tests-for-early-stage-invasive-breast-cancer

ABSTRACT

Background

Breast cancer is a disease in which cells in the breast grow out of control. They often form a tumour that may be seen on an x-ray or felt as a lump.

Gene expression profiling (GEP) tests are intended to help predict the risk of metastasis (spread of the cancer to other parts of the body) and to identify people who will most likely benefit from chemotherapy. We conducted a health technology assessment of four GEP tests (EndoPredict, MammaPrint, Oncotype DX, and Prosigna) for people with early-stage invasive breast cancer, which included an evaluation of effectiveness, safety, cost effectiveness, the budget impact of publicly funding GEP tests, and patient preferences and values.

Methods

We performed a systematic literature search of the clinical evidence. We 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. We also performed a literature survey of the quantitative evidence of preferences and values of patients and providers for GEP tests.

We performed an economic evidence review to identify published studies assessing the cost-effectiveness of each of the four GEP tests compared with usual care or with one another for people with early-stage invasive breast cancer. We adapted a decision-analytic model to compare the costs and outcomes of care that includes a GEP test with usual care without a GEP test over a lifetime horizon. We also estimated the budget impact of publicly funding GEP tests to be conducted in Ontario, compared with funding tests conducted through the out-of-country program and compared with no funding of tests in any location.

To contextualize the potential value of GEP tests, we spoke with people who have been diagnosed with early-stage invasive breast cancer.

Results

We included 68 studies in the clinical evidence review. Within the lymph-node–negative (LN−) population, GEP tests can prognosticate the risk of distant recurrence (GRADE: Moderate) and may predict chemotherapy benefit (GRADE: Low). The evidence for prognostic and predictive ability (ability to indicate the risk of an outcome and ability to predict who will benefit from chemotherapy, respectively) was lower for the lymph-node–positive (LN+) population (GRADE: Very Low to Low). GEP tests may also lead to changes in treatment (GRADE: Low) and generally may increase physician confidence in treatment recommendations (GRADE: Low).

Our economic evidence review showed that GEP tests are generally cost-effective compared with usual care.

Our primary economic evaluation showed that all GEP test strategies were more effective (led to more quality-adjusted life-years [QALYs]) than usual care and can be considered cost-effective below a willingness-to-pay of $20,000 per QALY gained. There was some uncertainty in our results. At a willingness-to-pay of $50,000 per QALY gained, the probability of each test

being cost-effective compared to usual care was 63.0%, 89.2%, 89.2%, and 100% for EndoPredict, MammaPrint, Oncotype DX, and Prosigna, respectively.

Sensitivity analyses showed our results were robust to variation in subgroups considered (i.e., LN+ and premenopausal), discount rates, age, and utilities. However, cost parameter assumptions did influence our results. Our scenario analysis comparing tests showed Oncotype DX was likely cost-effective compared with MammaPrint, and Prosigna was likely cost-effective compared with EndoPredict. When the GEP tests were compared with a clinical tool, the cost-effectiveness of the tests varied. Assuming a higher uptake of GEP tests, we estimated the budget impact to publicly fund GEP tests in Ontario would be between $1.29 million (Year 1) and $2.22 million (Year 5) compared to the current scenario of publicly funded GEP tests through the out-of-country program.

Gene expression profiling tests are valued by patients and physicians for the additional information they provide for treatment decision-making. Patients are satisfied with what they learn from GEP tests and feel GEP tests can help reduce decisional uncertainty and anxiety.

Conclusions

Gene expression profiling tests can likely prognosticate the risk of distant recurrence and some tests may also predict chemotherapy benefit. In people with breast cancer that is ER+, LN−, and human epidermal growth factor receptor 2 (HER2)–negative, GEP tests are likely cost-effective compared with no testing. The GEP tests are also likely cost-effective in LN+ and premenopausal people. Compared with funding GEP tests through the out-of-country program, publicly funding GEP tests in Ontario would cost an additional $1 million to $2 million annually, assuming a higher uptake of tests. GEP tests are valued by both patients and physicians for chemotherapy treatment decision-making.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OBJECTIVE

This health technology assessment evaluates the effectiveness, safety, and cost-effectiveness of gene expression profiling (GEP) tests for people with early-stage invasive breast cancer. It also evaluates the budget impact of publicly funding GEP tests and the experiences, preferences, and values of people with early-stage invasive breast cancer.

BACKGROUND

Health Condition

Breast cancer is a disease in which cells in the breast grow out of control, eventually forming a tumour. Environmental, lifestyle, and genetic factors influence a person's risk of developing breast cancer. These risk factors may include obesity, physical inactivity, alcohol consumption, age, hormone replacement therapy, dense breasts, genetic mutation, and a personal and/or family history of breast cancer.[1] Breast cancer is typically first detected as a lump or thickening of the breast that is discovered through self examination or through screening mammography. Diagnosis is made through tissue biopsy.

Different classifications for breast cancer can influence disease prognosis and treatment response. Classification can be based on cancer stage, visual examination of how abnormal cells look under a microscope (histological grade), the presence or absence of certain receptors (receptor status), molecular subtype, or specific gene expression.

The most common staging system for breast cancer is the TNM system from the American Joint Committee on Cancer (AJCC), which factors the size of the tumour and its extension (T), the lymph node involvement (N), and the metastasis (M).[2] The TNM system recognizes five stages of breast cancer[2]:

- Stage 0 (noninvasive): abnormal cells are present but have not spread to nearby tissue; also known as carcinoma in situ
- Stage 1 (invasive): cancer is present, but is contained in the area where the first abnormal cells began to develop
- Stage 2 (invasive): cancer is growing, but is still contained in the breast or growth has only extended to the nearby lymph nodes
- Stage 3 (invasive): 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
- Stage 4 (metastatic): cancer has spread to other distant areas of the body such as the liver, lung, bones, or brain

The absence of lymph node involvement is known as lymph-node–negative (LN−) breast cancer, and the spread of breast cancer to nearby lymph nodes is known as lymph-node–positive (LN+) breast cancer. One or more lymph nodes may be affected in LN+ breast cancer. Sometimes only a small mass of tumour cells has spread to nearby lymph nodes. This is distinct from LN− or LN+ breast cancer. Cancer spread in lymph nodes is called isolating tumour cells (ITC) if it is < 0.2 mm and micrometastasis if it is > 0.2 and ≤ 2 mm.

The most commonly tested receptors in breast cancer cells are the estrogen receptor (ER), the progesterone receptor (PR), and the human epidermal growth factor receptor 2 (HER2). A tumour is positive for a receptor where testing reveals the presence of that receptor (e.g., a cell with the ER receptor is ER-positive, or ER+) and negative where testing reveals its absence (e.g., a cell that lacks the ER receptor is ER-negative, or ER−). A tumour that tests negative for all three receptors (ER, PR, and HER2) is known as triple-negative breast cancer.

Breast cancer can also be classified by its molecular subtype. Table 1 summarizes the typical features of each of the four major molecular subtypes.

Table 1: Characteristics of the Four Major Molecular Subtypes of Breast Cancer

Subtype	ER	PR	HER2	Ki-67 Level[a]	Prevalence, %	Prognosis
Luminal A	+	+	−	Low	30−70	Usually grows slowly over time and has the best prognosis of the 4 subtypes
Luminal B	+	+ or −	+ or −	High	10−20	Usually has worse prognosis than luminal A
HER2-enriched	−	−	+	Any	5−15	Usually grows faster than luminal A and B subtypes
Basal-like	−	−	−	Any	15−20	Often aggressive and has poorer prognosis than luminal A and B subtypes

Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor 2 receptor; PR, progesterone receptor.
[a]Ki-67 is a protein associated with cell proliferation.
Sources: Dai et al,[3] and Cho et al.[4]

Treatment for early-stage invasive breast cancer typically involves surgery—either a lumpectomy or a mastectomy. In a lumpectomy (also known as breast-conserving surgery, a partial mastectomy, or wide excision), only the area of the breast containing the cancer is removed, preserving the rest of the breast tissue; in a mastectomy, the entire breast is removed. If the breast cancer has spread to nearby lymph nodes, surgery may also remove the affected nodes. This may be done as part of the breast cancer surgery or as a separate operation (sentinel lymph node biopsy or axillary lymph node dissection).

Surgery is often followed by adjuvant therapies such as radiation therapy, hormone therapy, biological therapy, and/or chemotherapy. Different types of breast cancer respond differently to each type of adjuvant therapy. For example, hormone receptor–positive tumours (i.e., ER/PR+) respond better to hormonal therapy, while patients with HER2+ tumours benefit from therapies that target the HER2 protein.

Terminology

As a government agency, Ontario Health (Quality) can play an active role in ensuring that people of all identities and expressions can recognize themselves in what they read, see, or hear from us. We recognize that, although breast cancer statistics are divided into male and female populations, gender identities are individual, and not everyone identifies with the sex they were assigned at birth.

Wherever possible, we use gender-inclusive pronouns and terms in accordance with the Ontario Health (Quality) Guidance for Gender-Inclusive Language (updated November 23, 2018). Due to the broader reporting and availability of data, this report focuses on breast cancer statistics gathered for people who have been identified as female by the reporting agencies and study authors. When necessary for clarity, discussions of these data will use the gendered terms "woman" and "women."

Clinical Need and Target Population

Breast cancer is the second most common cancer and the second leading cause of death from cancer in Canadian women.[5] In 2017, an estimated 26,530 people (26,300 women and 230 men) were diagnosed with breast cancer, and there were an estimated 4,960 deaths (4,900 women and 60 men).[5] This represents about 25% of all new cancer cases and 13% of all cancer deaths in women. In Ontario in 2018, about 12,000 cases of female breast cancer were expected to be diagnosed.[6]

The overall estimated 5-year survival rate for breast cancer is about 85%.[5] Most breast cancer cases in Ontario are diagnosed in the early stages, either at stage 1 (43%) or stage 2 (38%).[6]

Breast cancer recurrence, which is highest during the second year post-diagnosis, is related to the characteristics of the original breast cancer, such as tumour size, tumour subtype, tumour grade, patient age, and number of affected lymph nodes.[7] The risk of recurrence in another area of the body (distant recurrence) after 5 years of adjuvant endocrine therapy is strongly correlated with the original tumour lymph node status and tumour grade, ranging from 13% to 34% for stage 1 to 19% to 41% for stage 2 breast cancer.[8]

Appropriate selection and administration of breast cancer therapies is important to prolonging survival. Given the adverse effects of the different approaches, the decision to recommend chemotherapy for people with early invasive breast cancer presents a significant challenge.

Current Testing Options

Standard practice after surgical treatment of early-stage breast cancer is to administer adjuvant chemotherapy and/or hormonal therapy to reduce the risk of distant metastasis according to clinical, histological, and molecular characteristics of the tumour. Prognostic tools include the PREDICT tool[9] and the Nottingham Prognostic Index (NPI).[10] The PREDICT tool is made freely available through the UK's National Health Service and considers age, mode of detection, tumour size, tumour grade, ER status, number of LNs, HER2 status, Ki-67 status, and general chemotherapy regimen to predict 5- and 10-year survival.[9] The NPI incorporates tumour size, the number of LNs, and tumour grade.[11] Prognosis worsens as the NPI value increases and cut-off points are used to categorize people into good, moderate, and poor prognostic groups. NPI+ is an adaption of the NPI test that considers the breast cancer's molecular subtype.[10] Adjuvant! Online (AOL),[12] a free online tool that prognosticates a person's 10-year risk of distant recurrence and survival based on age, tumour size, tumour grade, ER status, and LN status, is no longer available.

Immunohistochemistry (IHC)-based tests may also be used to prognosticate the risk of distant recurrence, such as the IHC4 test.[13] IHC4 is a prognostic tool that estimates distant recurrence at 10 years in postmenopausal people with ER+ breast cancer who have received 5 years of endocrine therapy.[13] IHC4 incorporates ER, PR, HER2, and Ki-67status. IHC4 can also be combined with clinicopathological factors such as tumour size, tumour grade, LN status, and type of endocrine therapy (tamoxifen vs. aromatase inhibitor) into a modified tool known as the IHC4 + clinical (IHC4+C) score.

In Ontario, AOL was previously used as the primary non-genetic prognostic tool to help predict distant recurrence and inform decisions around chemotherapy treatment. Since it is no longer available, oncologists in Ontario now use the non-genetic prognostic PREDICT tool.

Health Technology Under Review

Gene expression profiling tests are intended to prognosticate the risk of distant metastasis and to identify the people who are most likely to benefit from chemotherapy. The aim of the test is to provide more accurate prognostic information than other non-genetic clinicopathological prognostic tests about specific molecular features of a person's breast cancer that may indicate an increased likelihood of rapid growth, metastasis risk, and response to chemotherapy. The tests are typically performed after surgery, in conjunction with other available information such as tumour size and grade. They are typically used in people with ER+ and LN− tumours (and sometimes LN+ tumours if the number of involved LNs is low or if there are micrometastases). Gene expression profiling tests are typically not used for breast cancers with a higher recurrence risk (e.g., LN+ tumours with a large number of affected LNs) because chemotherapy is often recommended regardless of the cancer's molecular profile.

A tissue sample is required for testing, which is typically obtained after surgery. Depending on the type of GEP test, fresh frozen specimens or formalin-fixed, paraffin-embedded (FFPE) specimens may be used. An FFPE sample is first preserved by fixing it in formaldehyde (formalin), to preserve the proteins and vital structures within the tissue. It is then embedded in a paraffin wax block.

Once the tissue sample is prepared, a GEP test assesses the type and number of messenger ribonucleic acid (mRNA) transcripts in the sample. The number of mRNA transcripts produced by a specific gene provides a measure of the gene's expression. Since mRNA transcripts are translated into proteins by the cells, GEP tests ultimately provide information about the changes in cell protein composition, which causes changes in the properties and functions of cells. Analytical methods used to produce a gene expression profile include reverse transcription–polymerase chain reaction (RT-PCR), microarray, and multi-mode analysis systems (e.g., nCounter Dx Analysis System).

There are multiple sources of variability that may affect the reproducibility and reliability of GEP test results. Specimens must contain a sufficient percentage of cancer cells.[14] A different ratio of cancer cells to normal cells may change the resulting gene expression profile and molecular signature of a tumour.[15] The GEP test may also yield false results in rarely seen tumours such as breast cancers that show neuroendocrine differentiation and mixed morphologies.[16] RNA is also very unstable and is prone to degradation and quality concerns, so proper preparation and isolation is vital.[14] The use of different test platforms, protocols, and reagents can also lead to differences in reproducibility among tests.

Commercially Available Gene Expression Profiling Tests

Levels of gene expression can be processed and combined according to complex algorithms to obtain composite scores associated with the specific types of tumours tested.[14] Four commercially available GEP tests are: EndoPredict, MammaPrint, Oncotype DX, and Prosigna. Each test's characteristics are summarized in Table 2. Tests results are typically available 2 to 3 weeks after testing.

Table 2: Test Characteristics of EndoPredict, MammaPrint, Oncotype DX, and Prosigna (PAM50)

Description	EndoPredict	MammaPrint	Oncotype DX	Prosigna (PAM50)
Manufacturer	Myriad	Agendia	Genomic Health	NanoString Technologies
Testing location	Can be done locally or at a central laboratory in the United States	Central (1 laboratory in the Netherlands, 1 in the United States)	Central (1 laboratory in the United States)	Can be done locally
Genes, n	12 for molecular score (8 cancer-related, 4 reference) Tumour size and nodal status for EPclin score	70 informative genes (465 reference genes)	21 (16 cancer-related, 5 reference)	50 (50 genes used to determine molecular subtype, 46 of 50 genes used to calculate recurrence score, 22 reference)
Test sample	FFPE	FFPE or fresh tissue	FFPE	FFPE
Test method	RT-PCR	Microarray-based	RT-PCR	nCounter Dx Analysis System
Population	Early-stage invasive breast cancer ER+, HER2− status	Stage 1 or 2 invasive breast cancer and LN− Tumour size ≤ 5.0 cm All ages[17]	Stage 1, 2, or 3a invasive breast cancer ER+, HER2− status[18]	Stage 1 or 2 invasive breast cancer and LN− Stage 2 or 3a invasive breast cancer and LN+ HR+ status Postmenopausal people[19]
Result measurement	Molecular score (0–15) EPclin score (1–6)[20]	MammaPrint Index[21]	Recurrence Score (0–100)[22]	Risk of Recurrence (0–100)[23]
Categories for risk measurement	*Molecular score* Low: < 5 High: ≥ 5 *EPclin score*[20] Low: < 3.3 High: ≥ 3.3	Low: 0 to 1 High: −1 to 0[21]	*LN− and age > 50 y*[22] Low: ≤ 25 High: 26–100 *LN− and age ≤ 50 y*[22] Low: ≤ 15 Intermediate: 16–20 & 21–25 High: 26–100 *LN+*[22] Low:< 18 Intermediate: 18–30 High: ≥ 31 *Previous categories*[a] Low: < 18 Intermediate: 18–30 High: ≥ 31	*LN−* Low: 0–40 Intermediate: 41–60 High: 61–100 *LN+ (1–3 nodes)* Low: 0–40 High: 41–100[24]

Description	EndoPredict	MammaPrint	Oncotype DX	Prosigna (PAM50)
10-year distant recurrence risk	*EPclin score*[25] Low: < 10% High: > 10%	Low: 10% (95% CI 4%–15%) High: 29% (95% CI 22%–35%)[17]	*LN− 9-year distant recurrence* RS ≤ 10: 3.2% ± 0.7 RS 11–25, endocrine therapy: 5.5% ± 0.5 RS 11–25, chemoendocrine therapy: 5.0% ± 0.5 RS ≥ 26, chemoendocrine therapy: 13.2% ± 1.7 *LN+ 9-year distant recurrence* RS < 18: 17% (95% CI 12–24%) RS 18–30: 28% (95% CI 20–39%) RS ≥ 31: 49% (95% CI 35-64%)	Low: < 10% Intermediate: 10%–20% High: > 20%[24]

Abbreviations: CI, confidence interval; EPclin, EndoPredict clinical score; ER, estrogen receptor; FFPE, formalin-fixed paraffin-embedded; HER2, human epidermal factor receptor 2; HR, hormone receptor; LN, lymph node; RS, Recurrence Score; RT-PCR, reverse-transcription polymerase chain reaction.
[a]Previous Oncotype DX risk categories before publication of the TAILORx trial (Trial Assigning IndividuaLized Options for Treatment [Rx]).[26]

EndoPredict evaluates the expression of eight cancer-related genes and four reference genes (EP score). The EndoPredict clinical score (EPclin) also integrates tumour size and nodal status. People are categorized as low or high risk. In low-risk people, the 10-year risk of distant recurrence when treated with 5 years of endocrine therapy alone is 4% on average. For high-risk people, the 10-year risk of recurrence is greater than 10%. Up to 78% of people with node-negative disease receive a low-risk score.[25]

MammaPrint was the first GEP test to publish evidence on its use (in 2002). The 70 genes included in the test (identified from the approximately 25,000 protein-coding genes in the full human genome) are predictive of recurrence risk.[27] A low-risk result indicates that a person has, on average, a 10% chance of distant recurrence within 10 years without any additional adjuvant hormonal therapy or chemotherapy. People with a high-risk result have a 29% chance.[21]

Oncotype DX evaluates 16 cancer-related genes, which were selected out of 250 possible genes based on their prognostic ability and test consistency. The expression of each of the 16 genes is measured in triplicate and then normalized relative to a set of five reference genes. The test uses a proprietary algorithm to calculate the Recurrence Score (RS) and then categorizes people as having a low, intermediate, or high risk of distant metastasis. For low-risk people, the benefit of chemotherapy is likely to be small and will not outweigh the risks of side effects. For high-risk people, the benefits of chemotherapy are likely greater than the risks of side effects. The risk–benefit calculation was originally uncertain for LN− intermediate-risk people. However, based on the results of the TAILORx (Trial Assigning IndividuaLized Options for Treatment [Rx]) trial by Sparano et al,[26] Oncotype DX test results have now been changed to a two-category risk score (low RS: ≤ 25, high RS: 26–100) for people > 50 years of age with LN− breast cancer.

Prosigna, formerly known as PAM50 (Predictor Analysis of Microarray 50), evaluates 50 genes and can distinguish between the molecular subtypes of breast cancer (i.e., luminal A, luminal B, HER2-enriched, and basal-like). The tumour's gene expression profile is compared with each of the four molecular subtypes to determine the degree of similarity. The results are combined with a proliferation score and tumour size to establish the Risk of Recurrence (ROR) score.[23] The ROR score is correlated with the 10-year probability of distant recurrence, with risk groups categorized as low (< 10%), intermediate (10–20%), and high (> 20%) ROR.[24]

Regulatory Information

Gene expression profiling tests are considered laboratory-developed tests and therefore do not require Health Canada approval unless they are marketed as test kits. Prosigna (license number 93159) and EndoPredict (license number 100294) test kits have Health Canada approval as Class 3 medical devices. In the United States, laboratory-developed tests do not require Food and Drug Administration (FDA) approval; however, Prosigna and MammaPrint have FDA approval.

Ontario and Canadian Context

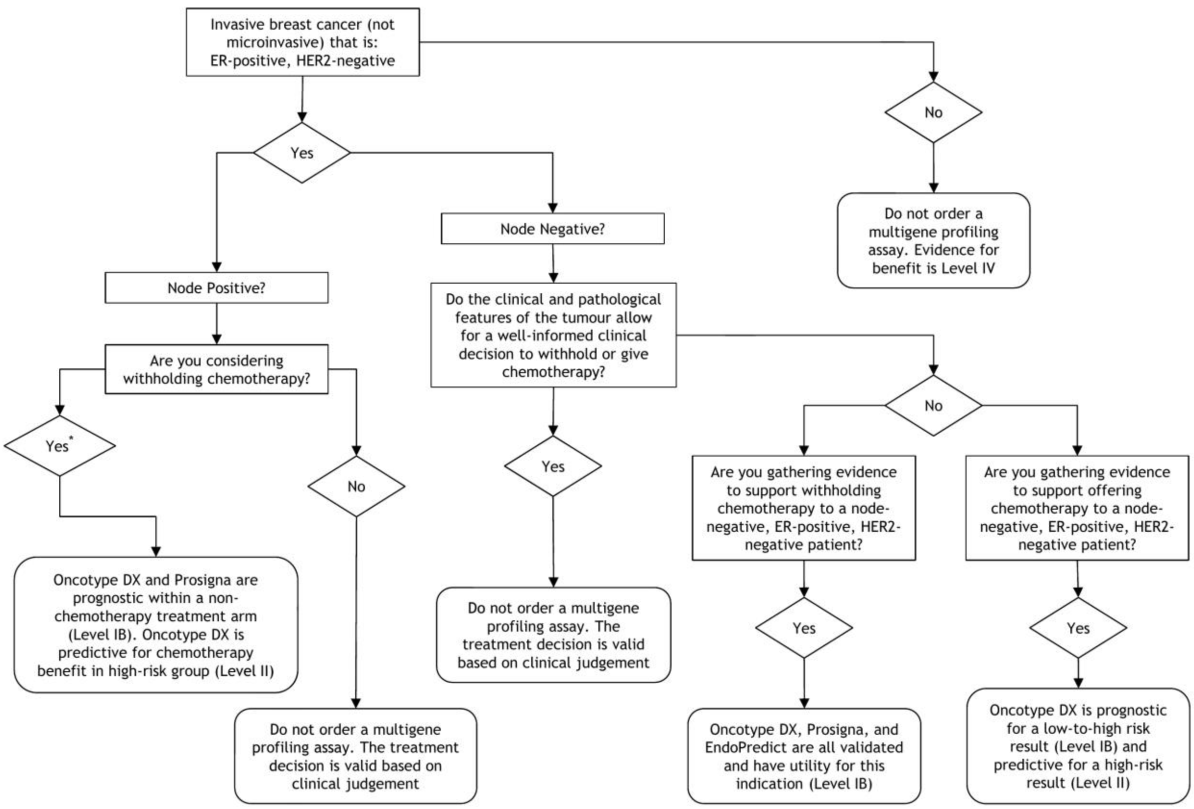
Ontario Context

The Ontario Ministry of Health publicly funds the GEP tests through the ministry's out-of-country program (phone communication, August 9, 2018). Oncotype DX is the predominant publicly funded test used for invasive breast cancer (not microinvasive) that is ER+ and HER2−. The test is performed by the manufacturer, Genomic Health Inc., in California. (Laboratories in the

United States that perform GEP tests are subject to federal regulatory standards called the Clinical Laboratory Improvement Amendments [CLIA].)[28] If GEP tests were to be performed in Ontario, laboratories would need to be provincially licensed to ensure quality standards are met.

From 2014 to 2017, there have been approximately 2,000 requests for Oncotype DX per year in Ontario (written communication, August 9, 2018). There have been approximately 50 requests for EndoPredict (only approved for out-of-country testing in Ontario in 2018), and none for other GEP tests (e.g., MammaPrint, Prosigna).

In 2016, Cancer Care Ontario (CCO)[a] performed a systematic review of GEP tests and made recommendations on clinical use[29,30] that form the basis for the ministry's eligibility criteria for out-of-country funding of GEP tests. The decision algorithm developed by CCO for the clinical use of GEP tests is presented below (Figure 1).



*In practice, this usually refers to micrometastatic (N1mi) disease

Figure 1: Cancer Care Ontario's Decision Algorithm for GEP Testing

Source: Chang et al.[29,30]
Note: **Level IB** evidence is defined as at least two category B studies (either randomized controlled trials designed to address a treatment intervention that is not the tumour biomarker or assay, or studies that prospectively enroll and follow patients, collect tumour samples, and then use archived tumour tissue retrospectively to evaluate the tumour biomarker or assay) with consistent results. **Level II** evidence is defined as one category B study, or multiple category B studies with inconsistent results, or at least two category C studies (prospective observational registry studies that prospectively enroll patients in a registry and collect, process, and archive tumour specimens, but treatment and follow-up are standard of care, and archived tumour tissue is used retrospectively

[a] Cancer Care Ontario is now Ontario Health (Cancer Care Ontario).

to evaluate the tumour biomarker or assay). Level IV evidence is defined as any number of category D studies (retrospective studies). **Level IV** evidence is insufficient for determining clinical utility.

Canadian Context

In Canada, Oncotype DX is the most common publicly funded GEP test. It is publicly funded in all 10 provinces (Genomic Health, personal communication, March 1, 2019). In most, the Oncotype DX test is publicly funded for the LN− population only; however, a few provinces will publicly fund the test for some LN+ people based on specific eligibility criteria (Genomic Health, written communication, March 1, 2019). Prosigna is publicly funded in Alberta, British Columbia, and Ontario. EndoPredict is publicly funded only in Ontario, and MammaPrint is not yet publicly funded in any province or territory. Appendix 1 summarizes the public funding status and eligibility criteria of EndoPredict, MammaPrint, Oncotype DX, and Prosigna.

Guidelines

Numerous international guidelines recommend the use of GEP tests to prognosticate distant recurrence in early-stage invasive breast cancer for adjuvant chemotherapy treatment decision-making (see Appendix 2 for a summary of guideline recommendations), including the American Society of Clinical Oncology,[31,32] the National Comprehensive Cancer Network (United States),[33] the St. Gallen International Expert Consensus,[34] the European Society of Medical Oncology,[35] and the National Institute for Health and Care Excellence (NICE).[36]

In general, the guidelines offer stronger recommendations for the use of GEP tests for early-stage invasive breast cancer that is ER/PR+, HER2−, and LN−. Recommendations for the use of GEP tests for the LN+ population are generally weaker, noting the more limited amount of evidence for this population. Recommendations on the use of GEP tests for chemotherapy benefit were typically restricted to specific GEP tests with evidence. NICE was the only guideline that we found that also considered the cost-effectiveness of different types of GEP tests in their recommendations.[36]

Health Technology Assessments and Systematic Reviews

A number of health technology assessments and systematic reviews have been conducted on GEP tests in recent years (see Appendix 3 for a summary of English-language HTAs). In addition, l'Institut national d'excellence en santé et services sociaux (INESSS) in Quebec has also published two HTAs in French on Oncotype DX and EndoPredict.[37,38] All the recent systematic reviews differ slightly in their population and outcomes of interest, types of GEP tests evaluated, and study eligibility criteria. Some have also included in their evaluation other GEP tests (e.g., Breast Cancer Index, or BCI) or non-genetic tests such as the IHC4.

In 2016, CCO published a systematic review on the clinical validity and utility of EndoPredict, MammaPrint, Oncotype DX, and Prosigna.[29,30] As part of their standard biennial guideline review and update process, CCO updated their literature search in 2018 to include more recent published evidence.[39] Due to the recency of CCO's updated review and its alignment with our clinical research questions, we decided to undertake an update of their work to include the most recent relevant clinical literature for four GEP tests (Oncotype DX, EndoPredict, MammaPrint, and Prosigna).

During the development of this HTA, we were also made aware of a similar ongoing HTA on GEP tests (examining only Oncotype DX and Prosigna) in Alberta.[40] Although the scope of

Alberta's Institute of Health Economics' HTA differs slightly from ours, we collaborated with them in an effort to share knowledge and avoid duplication of effort.

Expert Consultation

We engaged with experts in the specialty areas of medical oncology, pathology, breast cancer surgery, health services research, and health economics to help inform our understanding of the health technology, refine our methodologies, and contextualize the evidence.

CLINICAL EVIDENCE

Research Questions

- What are the effectiveness and safety of four gene expression profiling (GEP) tests (EndoPredict, MammaPrint, Oncotype DX, and Prosigna) for people with early-stage invasive breast cancer?
- What is the comparative effectiveness between GEP tests (EndoPredict, MammaPrint, Oncotype DX, and Prosigna) for people with early-stage invasive breast cancer?

Methods

Clinical Literature Search

To update the CCO 2018 report, we performed a clinical literature search on November 28, 2018, to retrieve studies published from January 1, 2018, until the search date. We used the Ovid interface to search the MEDLINE and Embase databases.

A medical librarian used the CCO literature search,[29,30] slightly modified for increased comprehensiveness. The final search strategy was peer-reviewed using the PRESS Checklist.[41]

We created database auto-alerts in MEDLINE and Embase and monitored them for the duration of the assessment period. We also performed a targeted grey literature search of clinical trial registries. The grey literature search was updated on May 2–3, 2019. See Appendix 4 for our literature search strategies, including all search terms.

Eligibility Criteria

Studies

Inclusion Criteria

- 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

Exclusion Criteria

- Retrospective nonrandomized (cohort) studies
- Nonsystematic reviews, narrative reviews, abstracts, editorials, letters, case reports, and commentaries
- Animal and in vitro studies

Participants

Inclusion Criteria

- People with early-stage invasive breast cancer of any age or receptor or LN status

Exclusion Criteria

- People with advanced invasive breast cancer
- People with only a specific subtype of breast cancer

Interventions

Inclusion Criteria

- Four commercially available GEP tests
  - EndoPredict
  - MammaPrint
  - Oncotype DX
  - Prosigna (PAM 50)
- Head-to-head comparative studies including two or more of the included GEP tests

Exclusion Criteria

- GEP tests not listed above
- No GEP test

Outcome Measures

- Prognostic ability (i.e., the degree to which GEP tests can accurately predict the risk of an outcome and discriminate people with different outcomes)
  - Freedom from distant recurrence (i.e., freedom from distant recurrence, second primary cancer, or death)
  - Disease-free survival (i.e., time from diagnosis or start of treatment until disease recurrence or death from any cause)
  - Overall survival (i.e., time from diagnosis or start of treatment until death due to any cause)
- Predictive ability (i.e., the degree to which GEP tests can identify people who will benefit most from chemotherapy)
  - Freedom from distant recurrence
  - Disease-free survival
  - Overall survival
- Clinical utility:
  - Changes in treatment management (i.e., changes in the recommendation or use of chemotherapy based on GEP test results)
  - Physician confidence in treatment recommendations

- Safety
    - Adverse events directly related to GEP testing

Literature Screening

A single reviewer conducted an initial screening of titles and abstracts using Covidence[42] 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. The reviewer also examined reference lists for any additional relevant studies not identified through the search.

Data Extraction

We extracted relevant data on study characteristics and risk-of-bias items using a data form to collect information on the following:

- Source (e.g., citation information, study type)
- Methods (e.g., study design, study duration and years, participant allocation, allocation sequence concealment, blinding, reporting of missing data, reporting of outcomes, whether the study compared two or more groups)
- Outcomes (e.g., outcomes measured, number of participants for each outcome, number of participants missing for each outcome, outcome definition and source of information, unit of measurement, upper and lower limits [for scales], time points at which the outcomes were assessed)

Statistical Analysis

We undertook a narrative summary of the results due to the heterogeneity[43] of patient populations and the reported endpoints of outcomes within studies. Results of the studies were stratified first by lymph node status (LN− or LN+) and presented by GEP test.

Critical Appraisal of Evidence

We assessed the risk of bias using the Cochrane Risk of Bias tool[44] for randomized controlled trials, the Prediction Model Risk of Bias Assessment Tool (PROBAST)[45] for prognostic studies, and the Risk of Bias Assessment Tool for Nonrandomized Studies (RoBANS)[46] for nonrandomized predictive ability or clinical utility studies for the included studies (Appendix 5).

We evaluated the quality of the body of evidence for each outcome according to the *Grading of Recommendations Assessment, Development, and Evaluation* (GRADE) *Handbook*.[47] The body of evidence was assessed based on the following considerations: risk of bias, inconsistency, indirectness, imprecision, and publication bias. The overall rating reflects our certainty in the evidence. This overall rating differs from CCO's Tumour Marker Utility Grading System (used in their 2016 review), which considers only the study design, the number of studies, and the consistency of results.

For interventional studies, the GRADE approach specifies that RCT evidence starts at high-quality and observational evidence at low quality.[47] In contrast, for prognostic studies, high-quality, prospective, longitudinal cohort studies provide high confidence.

Results

Clinical Literature Search

The database search of the clinical literature yielded 237 citations published from January 1, 2018, until November 28, 2018. We identified seven studies from the literature search, 13 from reference lists and experts, and two[48,49] from auto-alerts. We included an additional 46 studies from the CCO 2016 review[29,30] and the subsequent CCO 2018 update.[39] In total, we included 68 relevant studies in our review. Figure 2 presents the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram for the clinical literature search.

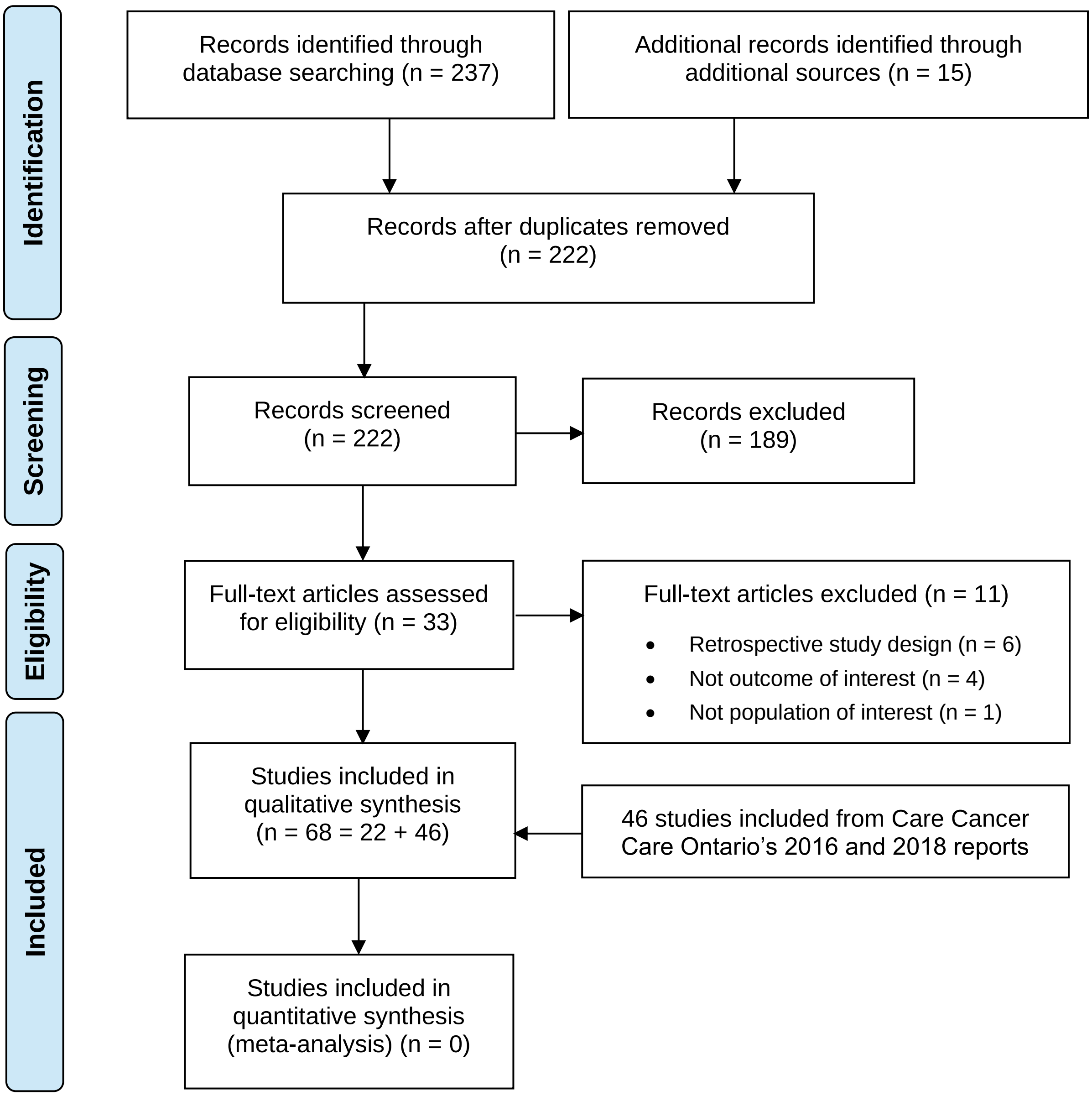


Figure 2: PRISMA Flow Diagram—Clinical Search Strategy

Source: Adapted from Moher et al.[50]

Characteristics of Included Studies

Table A7 (Appendix 5) summarizes the characteristics of the included studies. In general, the patient populations within the studies were variable. Studies typically included only ER/PR+, HER2− patients.

Hormone receptor status (ER/PR) was inconsistently reported within the studies (Table A7). Some studies reported that participants' tumours were all HR+ (specific ER or PR status was not reported), while others explicitly included people with tumours that were only ER+. Of the studies that reported ER or PR status, almost all included people with tumours that were > 95% ER+.

In some studies, there were also a small percentage of people with missing baseline characteristics. A few studies also accepted a small number of HER2+ patients or people with a larger number of positive lymph nodes (i.e., more than three positive LNs). Breast cancer characteristics (e.g., cancer stage, tumour grade, tumour size) varied throughout the studies. Most of the studies only included post-menopausal people, with few studies including either a combination of pre- or postmenopausal people, or did not report participants' menopausal status at all. For the prognostic studies, the treatment received was also variable (some studies included people who only had endocrine therapy, while others included groups who had endocrine therapy, chemoendocrine therapy, or chemotherapy).

The majority of the studies were from the United States, with other studies representing other countries such as Canada, the United Kingdom, Germany, France, Spain, Denmark, Israel, Mexico, Hong Kong, and Turkey. For the outcome of physician confidence in treatment recommendations, all physicians surveyed within the included studies were oncologists. Some studies were funded by test manufacturers.

Only two included studies were RCTs: Sparano et al[26] (the TAILORx trial for Oncotype DX), and Cardoso et al[51] (the MINDACT trial for MammaPrint). Both were noninferiority trials that evaluated the benefit of chemotherapy. The other nonrandomized studies were either retrospective reanalyses of other breast cancer treatment RCTs (sometimes a combination of different cohorts would be included) or prospective cohort studies. We also included studies based on retrospective analyses of cancer registries that met our inclusion criteria of prospective patient enrolment and prospective tumour specimen collection (e.g., National Cancer Database [NCDB] or the Surveillance, Epidemiology, and End Results [SEER] database in the United States). Due to the variety of sources of study participants (e.g., cohorts from previous RCTs, cancer registries), there may be substantial patient overlap between some studies.

We found a limited number of comparative studies that evaluate multiple GEP tests. We did not find any comparative studies that evaluated the performance of all four GEP tests within the same study population. The TransATAC study[52] is the closest fully comparative study that has been conducted thus far. The authors evaluated EndoPredict, Oncotype DX, and Prosigna.

We did not find any studies that reported on adverse events directly related to GEP testing.

Risk of Bias in the Included Studies

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.

Patient selection and analysis were the main areas for risk of bias. The methods used for patient recruitment and selection was unclear in some studies (e.g., whether consecutive patients were enrolled). People who were excluded based on insufficient samples or test failures may be systematically different from people who were included in the study.

Incomplete or selective reporting was another source for risk of bias as some studies did not report all prespecified analyses or subgroups. The method and process of randomization was also generally unclear within the RCTs.

Prognostic studies, unlike therapeutic interventions, do not test the influence of treatment on outcomes. Randomization is therefore irrelevant for prognostic studies. 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.[53] 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.

In contrast, predictive studies evaluate the ability of a test to affect outcomes (e.g., recurrence and survival) through prospective use of the test to guide treatment decisions. Therefore, studies that randomize chemotherapy guided by the test or by standard clinical practice are ideal. Observational studies that report on good clinical outcomes for low-risk people whose treatment (no chemotherapy) was guided by the test could support the avoidance of chemotherapy. The predictive ability of GEP tests has only been evaluated in two RCTs.[26,51] Nonrandomized studies are therefore at increased risk of bias due to other potential confounding factors.

Lymph-Node–Negative Population

Prognostic Ability

Twenty studies have evaluated the prognostic ability of GEP tests for distant recurrence (Table 3). Information on survival (overall or disease-free) is more limited. All four GEP tests were shown to be prognostic within the LN− population despite the clinical heterogeneity of patients within studies.

Studies Examining a Single GEP Test

Flipits et al[54] found that 10-year distant recurrence-free rates (0–10 years) was improved among people with low versus high EPclin scores. Similar results were observed for late recurrence (5–15 years; hazard ratio [HR] 4.52, 95% CI 2.65–7.72, $P < .001$).

For Oncotype DX, Nitz et al[55] found that nodal status, tumour grade, tumour size, continuous Ki-67, PR, IHC4, and RS were univariate prognostic factors for disease-free survival. The impact of Oncotype DX's RS on disease-free survival was particularly pronounced in people with intermediate Ki-67 (10% to 40%) tumours. Paik et al[56] found that RS provided significant prognostic power that was independent of age and tumour size ($P < .001$). It was also prognostic for overall survival ($P < .001$) and can be used as a continuous function to prognosticate distant recurrence.

Dowsett et al[57] found that Prosigna's ROR score added significant prognostic ability beyond clinical treatment score ($\Delta LR\chi^2 = 33.9$, $P < .001$). Liu et al[58] found that ROR did not have significant univariate effect on distant recurrence-free survival when high ROR was compared with low or intermediate ROR. Adjusting for patient and tumour characteristics, higher ROR was associated with worse recurrence-free survival. The intrinsic breast cancer subtype also had a significant prognostic effect on distant recurrence-free survival.[58]

Ohnstad et al[59] found that among the ER/PR+, HER2−, LN− population using Prosigna with no adjuvant treatment, 53.7% of people had a low ROR. The 15-year breast cancer-specific survival for this population was 96.3%. People with intermediate risk had reduced survival compared with those with low risk ($P = .005$). In contrast, no difference in survival between the low- and intermediate-risk groups was seen for people who received tamoxifen only.

Studies Comparing Different GEP Tests

Sestak et al (2018)[52] found that EndoPredict, Oncotype DX, and Prosigna performed similarly during the first 5 years of follow-up. There were differences in freedom from distant recurrence for Prosigna and EndoPredict during years 5 to 10, which suggests that they may be valuable for decision-making for extended endocrine treatment. Similarly, Sestak et al[60] compared Prosigna, Oncotype DX, and IHC4 and found that Prosigna's ROR score was also the strongest molecular prognostic factor in the late follow-up period ($\chi2 = 16.29$; $P < .001$), whereas IHC4 ($\chi2 = 7.41$) and RS ($\chi2 = 5.55$) were only weakly prognostic in this period.

Buus et al[61] compared EndoPredict with Oncotype DX and found that EndoPredict's EP and EPclin scores were highly prognostic for distant recurrence in endocrine-treated patients with ER+, HER2− disease. The authors also reported that EPclin provided more prognostic information than RS, which was partly but not entirely because EPclin integrates molecular data with nodal status and tumour size. EPclin also performed better than RS for late recurrence (years 5–10).

Table 3: Prognostic Ability of GEP Tests in a Lymph-Node–Negative Population

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
EndoPredict						
Buus et al, 2016[61]	179	EPclin < 3.3	10	94.2 (91.7–96.0)	NR	NR
	199	EPclin ≥ 3.3		71.5 (66.1–75.7)		
				EPclin ≤ 3.3 vs. > 3.3: HR 5.99 (3.94–9.11)		
Filipits et al, 2019[62]	907	EPclin < 3.3	10	95.5 (94.0–97.1)	NR	NR
	259	EPclin ≥ 3.3		87.0 (82.6–91.7)		
				EPclin ≤ 3.3 vs. > 3.3: HR 3.48 (2.18–5.56) *P* < .0001		
Sestak et al, 2018[52]	429	EPclin < 3.3	10	93.4 (90.3–95.5)	NR	NR
	162	EPclin ≥ 3.3		77.9 (70.2–83.8)		
				HR 2.14 (1.71–2.68)		
MammaPrint						
Drukker et al, 2013[63]	95	MP low/ AOL low	5	95.3 (90.9–100)	94.3 (89.5–99.3)	NR
	171	MP high/ AOL high		89.8 (85.1–94.8)	88.7 (83.8–93.8)	
	124	MP low/ AOL high		98.4 (96.1–100)	97.6 (94.9–100)	
	37	MP high/ AOL low		100 (100–100)	94.6 (87.6–100)	
Drukker et al, 2014[64]	219	Low risk	5	97.0 (94.7–99.4)	NR	NR
	208	High risk		91.7 (87.9–95.7)		
				Between groups: *P* = .03		
Esserman et al, 2017[65]	652	Ultra-low risk Low risk High risk	20	NR	High risk vs. ultralow risk breast-cancer-specific survival: adjusted HR 4.7 (1.38–16.22) Low risk vs. ultralow risk breast-cancer-specific survival: adjusted HR 4.54 (1.40–14.80)	NR

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
van de Vijver et al, 2002[66]	115	Low risk	10	NR	NR	96.7 ± 2.3
	180	High risk				49.6 ± 6.1
van't Veer et al, 2017[67]	159	Low risk	10	NR	0.93 (0.88–0.96)	NR
	57	High risk			0.85 (0.75–0.91)	
Oncotype DX						
Buus et al, 2016[61]	201	RS < 18	10	94.7 (91.8–96.5)	NR	NR
	78	RS 18–30		85.7 (79.4–90.2)		
	22	RS ≥ 31		74.9 (61.7–84.2)		
				RS < 18 vs. 18–30: HR 3.04 (1.68–5.50), *P* < .001 RS 18–30 vs. ≥ 31: HR 5.84 (2.99–11.40), *P* < .001		
Dowsett et al, 2010[68]	514	RS < 18	9	96 (93–97)	NR	88 (NR)
	228	RS 18–30		88 (82–92)		84 (NR)
	131	RS ≥ 31		75 (66–83)		73 (NR)
Mamounas et al, 2010[69]	509	RS < 18	10	95.7 (93.7–97.7)	NR	NR
	234	RS 18–30		92.8 (89.0–96.6)		
	280	RS ≥ 31		84.2 (78.8–89.6)		
				50-point RS change: HR 2.16 (1.26–3.68)		
Nitz et al, 2017[55]	248	RS ≤ 11	10	NR	NR	99.2 (98.0–100.0)
	156	RS 12–25				98.3 (97.0–99.5)
	61	RS ≥ 26				96.7 (94.4–99.0)
						RS ≤ 11 vs. ≥ 26: HR = NR; *P* < .05
						RS 12–25 vs. ≥ 26: HR = NR; *P* < .05
Paik et al, 2004[70]	340	RS < 18	10	93.2 (90.4–96.0)	NR	NR
	147	RS 18–30		85.7 (79.7–91.7)		
	180	RS ≥ 31		69.5 (62.9–76.4)		

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Petkov et al, 2016[71]	7,281	RS ≤ 11	5	NR	99.6 (99.4–99.8)	NR
	26,462	RS 12–25			99.3 (99.2–99.4)	
	6,391	RS ≥ 26			96.4 (95.6–97.0)	
					Between groups: $P < .01$	
Sestak et al, 2018[52]	374	RS ≤ 17	10	94.1 (90.9–96.2)	NR	NR
	156	RS 18–31		83.3 (76.0–88.5)		
	61	RS ≥ 32		72.8 (58.5–82.7)		
				HR 0.59 (0.49–0.71), $P < .05$		
Stemmer et al, 2017[72]	304	RS ≤ 10	5	99.0 (96.9–99.7)	NR	100.0 (100.0–100.0)
	1,037	RS 11–25		98.7 (97.8–99.2)		99.6 (98.1–99.8)
				Between groups: P = NS RS ≤ 10 vs. 11–25: HR = NR, P = NS		NR
	880	RS ≤ 17	5	99.2 (98.3–99.6)	NR	100.0 (100.0–100.0)
	733	RS 18–30		97.0 (95.5–98.0)		99.1 (98.1–99.6)
	188	RS ≥ 31		91.4 (86.3–94.6)		93.8 (89.1–96.5)
				Between groups: P <.001 RS ≤ 17 vs. ≥ 31: adjusted HR 0.17 (0.08–0.39), $P < .05$ RS ≤ 17 vs. 18–30: adjusted HR 0.50 (0.23–1.03), P = NS	NR	NR
Prosigna						
Filipits et al, 2014[54]	448	ROR ≤ 40	> 5	NR	NR	NR
	292	ROR 41–60		NR		
	179	ROR > 60		NR		
				ROR < 26 vs. 27–68: HR 4.03, $P < .002$ ROR 27–68 vs. ≥ 69: HR 4.74, $P < .001$		
Gnant et al, 2014[73]	487	ROR ≤ 40	10	96.6 (94.4–97.9)	NR	NR
	335	ROR 41–60		90.4 (86.3–93.3)		
	225	ROR > 60		84.3 (78.4–88.6)		

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Laenkholm et al, 2018[74]	361	ROR ≤ 40	10	95.0 (92.0–97.1)	NR	NR
	178	ROR 41–60		92.7 (89.4–95.2)		
	95	ROR ≥ 61		82.2 (78.0–86.0)		
				ROR ≤ 40 vs. ≥ 61: HR NR, *P* < .001 ROR 41–60 vs. ≥ 61: HR NR, *P* = NS		
Sestak et al, 2015[60]	983	ROR ≤ 26	10	98.0 (96.8–98.7)	NR	NR
	344	ROR 27–68		91.0 (87.0–93.8)		
	128	ROR ≥ 69		88.5 (81.0–93.2)		
				ROR ≤ 26 vs. 27–68: HR 3.75 (2.19–6.41) ROR 27–68 vs. ≥ 69: HR 5.49 (2.92–10.35)		
Sestak et al, 2018[52]	318	ROR ≤ 26	10	97.0 (94.2–98.4)	NR	NR
	178	ROR 27–68		85.9 (79.2–90.6)		
	95	ROR ≥ 69		67.6 (56.2–76.6)		
				HR 0.39 (0.30–0.51), *P* < .05		

Abbreviations: AOL, Adjuvant! Online; CI, confidence interval; EPclin, EndoPredict clinical score; GEP, gene expression profiling; HR, hazard ratio; MP, MammaPrint; NR, not reported; NS, not significant; ROR, Risk of Recurrence; RS, Recurrence Score.

Predictive Ability

Studies Examining a Single GEP Test

We found one study evaluating the predictive ability of MammaPrint and five studies for Oncotype DX (Table 4). The MINDACT[51] RCT evaluated the chemotherapy benefit for people who are primarily clinically high risk for recurrence as determined by MammaPrint. The authors found no statistical difference for freedom from distant recurrence and overall survival between the chemotherapy and no chemotherapy groups for clinically high-risk people who were MammaPrint low risk. This suggests that people who would have otherwise been candidates for adjuvant chemotherapy were able to forgo chemotherapy based on the low-risk MammaPrint result. At 5 years, people classified as clinically high risk and MammaPrint high risk had the lowest rate of recurrence-free survival in the study (90.6%), whereas people classified as clinically low risk and MammaPrint low risk had the highest rate (97.6%). People with discordant results had about 95% recurrence-free survival. They identified a discordance rate (the difference between clinical risk and MammaPrint risk) of 32%, suggesting that tumour characteristics are important factors in treatment decision-making.

The TAILORx RCT[26] evaluated the chemotherapy benefit of Oncotype DX and found that endocrine therapy was not inferior to chemoendocrine therapy for people who are ER/PR+, HER2−, and LN−. The 9-year rate of freedom from distant recurrence in people with a RS of 11 to 25 was about 95%, irrespective of chemotherapy use. Exploratory analyses indicated that chemotherapy was associated with some benefit for people ≤ 50 years of age with a RS of 16 to 25. Similar results for the predictive ability of Oncotype DX was found within the nonrandomized studies.[56,72,75,76]

Table 4: Predictive Ability of GEP Tests in a Lymph-Node–Negative Population

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
MammaPrint						
Cardoso et al, 2016[51]	592	MP low/AOL low	5	97.6 (96.9–98.1)	92.8 (91.7–93.7)	98.4 (97.8–98.9)
	336	MP high/AOL low		94.8 (92.4–96.4)	90.3 (87.3–92.6)	97.2 (95.5–98.3)
	224	MP low/AOL high		95.1 (93.8–96.2)	91.4 (89.7–92.8)	97.6 (96.6–98.3)
	254	MP high/AOL high		90.6 (89.0–92.0)	85.3 (83.4–87.0)	94.7 (93.4–95.7)
				MP high vs. MP low: adjusted HR 1.49 (1.05–2.13)		
Oncotype DX						
Geyer et al, 2018[76]	66	RS ≤ 10: ET	10	98.0 (95.0–100.0)	NR	NR
	110	RS ≤ 10: CET		95.0 (90.0–99.0)		
	103	RS 11–25: ET		95.0 (90.0–99.0)		
	168	RS 11–25: CET		94.0 (90.0–98.0)		
	35	RS ≥ 26: ET		62.0 (48.0–81.0)		
	87	RS ≥ 26: CET		88.0 (81.0–95.0)		
				RS ≤ 10: ET vs. CET: adjusted HR 0.84 (0.29–2.44), *P* = NS		
				RS 11–25: ET vs. CET: adjusted HR 1.64 (0.74–3.85), *P* = NS		
				RS ≥ 26: ET vs. CET: adjusted HR 3.70 (1.61–8.33), *P* < .001		

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Ibraheem et al, 2019[75]	29,412	RS 11–17: ET	5	NR	NR	97.4 (NR)
	1,534	RS 11–17: CET				97.5 (NR)
	16,013	RS 18–25: ET				96.4 (NR)
	7,133	RS 18–25: CET				97.1 (NR)
	2,085	RS 26–30: ET				94.0 (NR)
	3,845	RS 26–30: CET				95.8 (NR)
						RS 11–17: ET vs. CET: adjusted HR 1.03 (0.65–1.64), *P* = NS
						RS 18–25: ET vs. CET: adjusted HR 1.27 (1.00–1.61), *P* = .052
						RS 26–30: ET vs. CET: adjusted HR 1.47 (1.04–2.08), *P* = .029
Paik et al, 2006[56]	353	RS < 18: CT	10	98.7 (96.2–99.5)	NR	NR
	134	RS 18–30: CT		99.4 (98.4–99.8)		
	164	RS ≥ 31: CT		99.7 (99.5–99.9)		

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Sparano et al, 2018[26]	1,619	RS ≤ 10: ET	5	99.3 ± 0.2	94.0 ± 0.6	98.0 ± 0.4
	3,339	RS 11–25: ET		98.0 ± 0.3	92.8 ± 0.5	98.0 ± 0.2
	3,312	RS 11–25: CET		98.2 ± 0.2	93.1 ± 0.5	98.1 ± 0.5
	1,389	RS ≥ 26: CET		93.0 ± 0.8	87.6 ± 1.0	87.6 ± 1.0
	1,619	RS ≤ 10: ET	9	96.8 ± 0.7	84.0 ± 1.3	93.7 ± 0.8
	3,339	RS 11–25: ET		94.5 ± 0.5	83.3 ± 0.9	93.9 ± 0.5
	3,312	RS 11–25: CET		95.0 ± 0.5	84.3 ± 0.8	93.8 ± 0.5
	1,389	RS ≥ 26: CET		86.8 ± 1.7	75.7 ± 2.2	89.3 ± 1.4
				RS 11–25: ET vs. CET: HR 1.10 (0.85–1.41), *P* = NS	RS 11–25: ET vs. CET: HR 1.08 (0.94–1.24), NS	HR 0.99 (0.79–1.22), *P* = NS
				RS 11–15: ET vs. CET: HR 1.08 (0.64–1.82) *P* = NS	RS 11–15: ET vs. CET: HR 0.95 (0.75–1.22), NS	
				RS 16–20: ET vs. CET: HR 0.95 (0.63–1.43) *P* = NS	RS 16–20: ET vs. CET: HR 1.04 (0.84–1.29), NS	
				RS 20–25: ET vs. CET: HR 1.27 (0.85–1.90) *P* = NS	RS 20–25: ET vs. CET: HR 1.32 (1.01–1.71), *P* < .05	
				RS 11–17: ET vs. CET: HR 1.00 (0.67–1.49) *P* = NS	RS 11–17: ET vs. CET: HR 1.01 (0.82–1.23), NS	
				RS 18–25: ET vs. CET: HR 1.16 (0.84–1.60) *P* = NS	RS 18–25: ET vs. CET: HR 1.16 (0.96–1.40), NS	
Stemmer et al, 2017[72]	473	RS 18–25: ET	5	98.0 (96.2–99.0)	NR	NR
	89	RS 18–25: CET		96.4 (89.1–98.8)		
	86	RS 26–30: ET		94.2 (86.6–97.5)		
	85	RS 26–30: CET		95.0 (87.0–98.1)		

Abbreviations: AOL, Adjuvant! Online; CET, chemoendocrine therapy; CI, confidence interval; CT, chemotherapy; ET, endocrine therapy; GEP, gene expression profiling; HR, hazard ratio; MP, MammaPrint; NR, not reported; NS, not significant; ROR, Risk of Recurrence; RS, Recurrence Score.

Changes in Treatment Recommendations

Thirteen studies evaluated the change in treatment management within a LN− population (Table 5). Albanell et al (2012)[77] examined the clinical factors that may influence changes in treatment after a GEP test. They found that a higher tumour grade (*P* = .007) and a high proliferative index (Ki-67) (*P* = .023) were significantly associated with a greater chance of changing from hormone therapy to chemotherapy, while PR+ status (*P* = .002) was associated with a greater probability of changing from chemotherapy to hormone therapy. The Recurrence Score was also significantly associated with the likelihood of change from hormone therapy to chemotherapy (*P* < .001) and vice versa (*P* < .001).

Table 5: Changes in Treatment Recommendations in a Lymph-Node–Negative Population

Author, Year	Patients, n	No CT to CT, n (%)	CT to No CT, n (%)	Total Treatment Change, n (%)
MammaPrint				
Kuijer et al, 2017[78]	660	38/660 (6) Unsure to CT: 110/660 (17)	156/660 (24) Unsure to no CT: 173/660 (26)	194/377 for those with a clear recommendation prior to testing (51; 95% CI 46–56), *P* < .001
Oncotype DX				
Albanell et al, 2012[77]	107	12/107 (11)	22/107 (21)	34/107 (32; 95% CI 26–34)
Albanell et al, 2016[79]	527	53/527 (10)	115/527 (22)	168/527 (32)
Bargallo et al, 2015[80]	62	6/62 (10)	11/62 (17)	17/62 (27)
de Boer et al, 2013[81]	101	12/71 (17)	12/30 (40)	24/101 (24)
Dieci et al, 2018[82]	124	5/124 (4)	10/124 (8)	15/124 (12)
Eiermann et al, 2013[83]	244	28/244 (11)	45/244 (18) CT to CET: 1/244 (1)	74/244 (30; 95% CI 24.6–36.5)
Levine et al, 2016[84]	979	Unsure or no CT: 143/979 (15)	Unsure or CT: 365/979 (38)	508/979 (52)
Lo et al, 2010[85]	89	3/89 (3) No CT to no CT or CT: 3/89 (3)	20/89 (23) CT to no CT or CT: 2/89 (2)	28/89 (31)
Loncaster et al, 2017[86]	136	0/136	82/136 (60)	82/136 (60)
Ozmen et al, 2016[87]	165	10/165 (6)	41/165 (25)	51/165 (31)
Prosigna				
Hequet et al, 2017[88]	194	25/194 (13)	9/194 (5)	34/194 (18), *P* < .001
Wuerstlein et al, 2016[89]	198	22/198 (11)	5/198 (3)	27/198 (14)

Abbreviations: CI, confidence interval; CET, chemoendocrine therapy; CT, chemotherapy.

Physician Confidence in Treatment Recommendations

Six studies reported physician confidence in treatment recommendations for Oncotype DX (four studies) and Prosigna (two studies) for LN− breast cancer. Study results are presented in Table 6. In general, about 40% to 80% of physicians reported increased confidence after the use of a GEP test.

Table 6: Physician Confidence in Treatment Recommendations in a Lymph-Node–Negative Population

Author, Year	Physician Confidence in Treatment Recommendations
Oncotype DX	
Albanell et al, 2012[77]	Increased for 60% of physicians No change for 33% of physicians Decreased for 7% of physicians
Albanell et al, 2016[79]	Increased for 33.0%–60.2% of physicians No change for 33.0%–52.4% of physicians Decreased for 6.8%–14.9% of physicians
Eiermann et al, 2013[83]	45% increased confidence for lymph-node–negative cases
Lo et al, 2010[85]	Increased confidence in 68 cases (76%)
Prosigna	
Hequet et al, 2017[88]	Increased for 39% of physicians No change for 51% of physicians Decreased for 11% of physicians
Wuerstlein et al, 2016[89]	Increased for 88% of physicians No change for 10% of physicians Decreased for 2% of physicians

Lymph-Node–Positive Population

Prognostic Ability

Seven of the 19 studies we examined evaluated Oncotype DX. In general, GEP tests were prognostic among the LN+ population (Table 7); however, the results were weaker compared with the LN− population.

Studies Examining Prognostic Ability for a Single GEP Test

Albain et al[90] did not report individual results by Oncotype DX RS group but found that the continuous RS was highly significant for a 50-point difference with HR 2.64 (95% CI 1.33–5.27, $P = .006$). The HR for RS was not constant over time: in the first 5 years, the HR was 5.55 (95% CI 2.32–3.28, $P < .001$), but for those surviving beyond 5 years, the RS was no longer prognostic (HR 0.86, 95% CI 0.27–2.74, $P = .80$). The authors noted that the prognostic effect persisted over the entire study period.

The prognostic ability of EndoPredict was also assessed in mixed LN status populations. Dubsky et al[91] found that EndoPredict was significantly more prognostic compared with clinical parameters alone ($P < .001$). Fitzal et al[92] found that the risk of local recurrence for high-risk lesions was higher than for low-risk lesions (HR 1.31, 95% CI 1.16–1.48). In a LN+ population (1–3 nodes), Filipits et al (2019)[62] found that EndoPredict was prognostic for distant recurrence and, in multivariable analysis, EP and EPclin scores remained significant predictors for distant recurrence.

Studies Comparing Different GEP Tests

Dowsett et al (2013)[57] compared Prosigna to Oncotype DX and used a likelihood ratio value ($\Delta LR\chi^2$) to quantitatively measure the relative amount of information provided by one score compared with another. Prosigna's ROR score added significant prognostic information beyond clinical parameters in all LN+ people ($\Delta LR\chi^2 = 33.9$, $P < .001$) and more information was added by Prosigna's ROR than by Oncotype DX's RS. In addition, more patients were correctly scored as high risk and fewer as intermediate risk by Prosigna than by Oncotype DX.

Martin et al[93] compared EndoPredict to Prosigna (PAM50) and found a 20% discrepancy between risk categorizations. However, the distant recurrence rate between discrepant people was non-significant. EndoPredict low-risk patients were found to have a better outcome than low-risk Prosigna patients. EndoPredict's EPclin risk classification proved a superior predictor of freedom from distant recurrence when compared with Prosigna's ROR cut-offs of < 29, 29–65, and > 65 ($P = .04$), but not for ROR cut-offs of < 18, 18–65, > 65 ($P = .09$).

Buus et al[61] found EndoPredict's EPclin score was more prognostic than Oncotype DX's RS, in particular during late recurrence (years 5–10; $LR\chi^2$: EPclin = 59.3, $LR\chi^2$: RS = 5.6).

Sestak et al (2018)[52] found that EndoPredict, Oncotype DX, and Prosigna provided significant prognostic information for LN+ people, with EndoPredict and Prosigna being more prognostic than Oncotype DX. However, the prognostic ability of all three GEP tests was weaker for the LN+ population compared with the LN− population. EndoPredict provided the most prognostic value for late recurrence (years 5–10; HR 1.87, 95% CI 1.27–2.76), followed by Prosigna (HR 1.65, 95% CI 1.08–2.51). Oncotype DX did not provide prognostic information for late distant recurrence on its own or in combination with clinical parameters.

Table 7: Prognostic Ability of GEP Tests in a Lymph-Node–Positive Population

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
EndoPredict						
Buus et al, 2016[61]	94	EPclin < 3.3	10	78.7 (68.1–86.1)	NR	NR
	141	EPclin ≥ 3.3		63.6 (71.1–54.8)		
				EPclin ≤ 3.3 vs. > 3.3: 1.78 (1.04–3.04)		
Dubsky et al, 2013[91]	832	EPclin < 3.3	10	98.2 (95.64–99.85)	NR	NR
	870	EPclin ≥ 3.3		87.69 (82.86–92.52)		
Filipits et al, 2019[62]	159	EPclin < 3.3	10	95.6 (92.2–99.1)	NR	NR
	377	EPclin ≥ 3.3		75.8 (71.0–80.9)		
				HR 4.70 (2.27–9.71), *P* < .0001		
Sestak et al, 2018[52]	43	EPclin < 3.3	10	94.4 (79.1–98.6)	NR	NR
	140	EPclin ≥ 3.3		69.7 (60.7–77.0)		
				HR 1.69 (1.29–2.22)		
MammaPrint						
van de Vijver et al, 2002[66]	115	Low risk	10	NR	NR	92.0 ± 4.8
	180	High risk				59.5 ± 6.3
Oncotype DX						
Buus et al, 2016[61]	49	RS < 18	10	74.9 (66.1–81.8)	NR	NR
	19	RS 18–30		65.2 (52.8–75.1)		
	9	RS ≥ 31		51.4 (30.8–68.6)		
				RS < 18 vs. 18–30: HR 1.60 (0.94–2.71), *P* = .08		
				RS 18–30 vs. ≥ 31: HR 2.85 (1.49–5.45), *P* = .002		
Dowsett et al, 2010[68]	159	RS < 18	9	83 (76–88)	NR	74 (NR)
	95	RS 18–30		72 (61–80)		69 (NR)
	52	RS ≥ 31		51 (36–65)		54 (NR)

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Gluz et al, 2016[94]	223	RS ≤ 11	5	93.6 (90.8–96.4)	NR	99.1 (98.5–100)
Nitz et al, 2017[55]	680	RS 12–25		94.3 (92.8–95.8)		97.2 (96.0–98.5)
	78	RS > 25		84.2 (80.6–87.8)		93.3 (90.8–95.8)
				Between groups: *P* < .001		Between groups: *P* < .001
King et al, 2016[95]	22	RS < 18	2	NR	NR	100 (78–100)
	29	RS 18–30				100 (78–100)
	50	RS ≥ 31				80 (69–93)
						50-point RS change: adjusted HR 20.58 (1.89–224.2)
						10-point RS change: adjusted HR 1.83 (1.14–2.95)
Mamounas et al, 2017[96]	386	RS < 18	10	Locoregional recurrence: 96.8 (94.1–98.5)	NR	NR
	364	RS 18–30		Locoregional recurrence: 94.9 (91.6–97.2)		
	315	RS ≥ 31		Locoregional recurrence: 92.1 (87.9–95.3)		
				50-unit increment in RS: adjusted HR 2.69 (1.28–5.26) *P* = .008		
Penault-Llorca et al, 2018[97]	209	RS < 18: ET or CET	5	93.7 (89.4–96.3)	90.8 (86.0–94.1)	99.0 (96.2–99.8)
	159	RS 18–30: ET or CET		87.3 (81.0–91.6)	84.9 (78.3–89.6)	95.6 (90.9–97.9)
	162	RS ≥ 31: ET or CET		69.3 (61.5–75.8)	64.6 (56.7–71.4)	85.6 (79.1–90.2)
				Between groups: *P* < .001 50-point change in RS: HR 4.14 (2.67–6.43)	Between groups: *P* < .001 50-point change in RS: HR 3.28 (2.18–4.94)	Between groups: *P* < .001 50-point change in RS: HR 5.0 (3.01–8.28)
Petkov et al, 2016[71]	2,694	RS ≤ 17: ET or CET	5	NR	85.7 (76.2–91.6)	NR
	1,669	RS 18–30: ET or CET			97.7 (95.9–98.7)	
	328	RS ≥ 31: ET or CET			99.0 (98.0–99.5)	
Roberts et al, 2017[98]	3,790	RS ≤ 17: ET or CET	5	NR	98.8 ± 0.3	92.1 ± 0.8
	2,263	RS 18–30: ET or CET			97.3 ± 0.6	90.9 ± 1.0
	430	RS ≥ 31: ET or CET			88.5 ± 2.4	81.7 ± 2.8

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Sestak et al, 2018[52]	105	RS ≤ 17: ET	10	80.6 (70.5–86.5)	NR	NR
	58	RS 18–31: ET		70.9 (56.9–81.1)		
	20	RS ≥ 32: ET		62.0 (35.9–80.0)		
				HR 0.72 (0.54–0.95), *P* < .05		
Stemmer et al, 2017[99]	379	RS ≤ 17: ET or CET	5	96.8 (94.4–98.2)	NR	99.5 (97.9–99.9)
	258	RS 18–30: ET or CET		93.7 (89.9–96.1)		96.6 (93.3–98.3)
	72	RS ≥ 31: ET or CET		83.1 (72.1–90.0)		94.3 (85.6–97.8)
				RS ≤ 17 vs. ≥ 31: HR 0.19 (0.09–0.40), *P* < .05 (adjusted HR 0.23 [0.11–0.50], *P* < .05)		
	109	RS ≤ 10: ET	5	96.3 (90.5–98.6)		99.1 (93.7–99.9)
	379	RS 11–25: ET		95.4 (92.8–97.1)		98.6 (96.6–99.4)
Prosigna						
Sestak et al, 2013[100]	137	ROR ≤ 26	10	96.7 (91.4–98.8)	NR	NR
	160	ROR 27–68		92.2 (86.2–95.6)		
	260	ROR ≥ 69		79.1 (73.1–83.9)		
				ROR ≤ 26 vs. 27–30: HR 3.16 (1.04–9.61)		
				ROR 27–30 vs. ≥ 69: HR 7.94 (2.87–21.92)		
Filipits et al, 2014[54]	12	ROR ≤ 15	> 5	NR	NR	NR
	124	ROR 16–40		NR		
	191	ROR > 40		NR		
				ROR 27–68 vs. ≥ 69: HR 3.15, *P* = .02		

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
Gnant et al, 2014[73]	15	ROR ≤ 15	10	100 (100–100)	NR	NR
	143	ROR 16–40		93.6 (86.9–97.0)		
	273	ROR > 40		76.1 (69.9–81.2)		
Laenkholm et al, 2018[74]	359	Low ROR	10	96.5 (93.9–98.1)	NR	NR
	388	Intermediate ROR		88.5 (84.4–92.0)		
	648	High ROR		77.9 (74.2–81.4)		
				ROR low vs. intermediate: adjusted HR 0.39 (0.20–0.77), *P* < .05		
				ROR intermediate vs. high: adjusted HR 0.65 (0.44–0.96), *P* < .05		
Sestak et al, 2018[52]	15	ROR ≤ 26	10	100.0 (100.0–100.0)	NR	NR
	58	ROR 27–68		79.3 (65.5–81.1)		
	110	ROR ≥ 69		69.3 (58.7–77.8)		
				HR 0.64 (0.47–0.86), *P* < .05		
Jensen et al, 2018[101] Mixed LN	155	ROR 8–51: ET	10	NR	62 (43–76)	63 (45–76)
	148	ROR 52–71: ET			27 (14–43)	38 (22–54)
	157	ROR 72–100: ET			27 (15–41)	30 (17–43)

Abbreviations: CET, chemoendocrine therapy; CI, confidence interval; EPclin, EndoPredict clinical score; ET, endocrine therapy; GEP, gene expression profiling; HR, hazard ratio; NR, not reported; NS, not significant; ROR, Risk of Recurrence; RS, Recurrence Score.

Predictive Ability

We found limited evidence on the predictive ability of GEP tests for LN+ people (Table 8). We did not find any studies that examined the predictive ability of EndoPredict in a purely LN+ population. However, in a retrospective analysis of prospective studies, Sestak et al (2019)[102] found that people who received chemoendocrine therapy had significantly smaller increases in 10-year distant recurrence rates with increasing EPclin score compared with those receiving endocrine therapy alone, suggesting that EndoPredict may be able to predict chemotherapy benefit. The authors also observed a significant positive interaction between EndoPredict's EPclin score and treatment.

The MINDACT trial[51] examined the predictive ability of MammaPrint in a LN+ subgroup and found a significant improvement in distant metastasis-free survival for the MammaPrint low/AOL high group. However, the results for the MammaPrint high/AOL low group were too small to be analyzed.

Two studies evaluated the predictive ability of the Oncotype DX test based on retrospective analyses of cancer registry data.[75,99] Ibraheem at al[75] found a significant difference in overall survival between RS groups (*P* < .001). Stemmer et al[99] did not report the significance of their findings. In another study, Albain et al[90] found the risk of recurrence in years 1 to 10 for a 50-point RS change was HR 2.64 (95% CI 1.33–5.27, *P* = .006).

In a mixed LN population, Jensen et al[101] found that Prosigna's molecular subtypes (i.e., luminal A, luminal B, basal-like, and HER2-enriched) could predict chemotherapy benefit for high-risk people (defined by ROR score), but not for low-risk people. We did not find any subgroup data on the predictive ability of Prosigna within a purely LN+ population.

Table 8: Predictive Ability of GEP Tests in a Lymph-Node–Positive Population

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
MammaPrint						
Cardoso et al, 2016[51]	353	MP low/AOL high: CT	5	96.3 (93.1–98.1)	NR	NR
	356	MP low/AOL high: no CT		95.6 (92.7–97.4)		
				Adjusted HR 0.88 (0.42–1.82)		
Oncotype DX						
Albain et al, 2010[90]	367	RS < 18: CT	10	NR	HR 1.02 (0.54–1.93)	NR
		RS 18–30: CT			HR 0.72 (0.39–1.31)	
		RS ≥ 31: CT			HR 0.59 (0.35–1.01)	
Ibraheem et al, 2019[75]	5,203	RS 11–17: ET	5	NR	NR	96.5 (NR)
	1,889	RS 11–17: CET				97.7 (NR)
	2,328	RS 18–25: ET				92.7 (NR)
	2,567	RS 18–25: CET				96.0 (NR)
	286	RS 26–30: ET				85.5 (NR)
	890	RS 26–30: CET				92.2 (NR)
						Between groups: *P* < .001 RS 11–17: ET vs. CET: adjusted HR 1.59 (1.01–2.50), *P* = .044 Between RS 18–25 ET vs. CET: adjusted HR 1.89 (1.32–2.70), *P* = .001 Between RS 26–30 ET vs. CET: adjusted HR 2.00 (1.12–3.57), *P* = .018
Stemmer et al, 2017[99]	342	RS ≤ 17: ET	5	NR	NR	99.4 (97.7–99.9)
	27	RS ≤ 17: CET				100.0 (100.0–100.0)
	153	RS 18–30: ET				95.0 (89.8–97.6)
	102	RS 18–30: CET				98.9 (92.1–98.8)
	136	RS 18–25: ET	5	NR	NR	96.8 (91.7–98.8)
	62	RS 18–25: CET				100.0 (100.0–100.0)
	20	RS 26–30: ET				84.0 (57.9–94.6)

Author, Year	Patients, n	Risk Category	Time Period, y	Freedom From Distant Recurrence, % (95% CI)	Disease-Free Survival, % (95% CI)	Overall Survival, % (95% CI)
	40	RS 26–30: CET				97.1 (80.9–98.6)
Prosigna						
Jensen et al, 2019[101]	113	Continuous ROR: ET	< 5	NR	10-point ROR difference: 1.25 (1.10–1.41)	10-point ROR difference: 1.33 (1.19–1.49)
(Mixed LN)	347	Continuous ROR: CT			10-point ROR difference: 1.30 (1.19–1.43)	10-point ROR difference: 1.29 (1.10–1.52)
	113	Continuous ROR: ET	> 5	NR	10-point ROR difference: 0.95 (0.82–1.11)	10-point ROR difference: 1.00 (0.94–1.08)
	347	Continuous ROR: CT			10-point ROR difference: 1.18 (0.81–1.72)	10-point ROR difference: 1.12 (0.99–1.28)

Abbreviations: AOL, Adjuvant! Online; CET, chemoendocrine therapy; CT, chemotherapy; ET, endocrine therapy; GEP, gene expression profiling; HR, hazard ratio; MP, MammaPrint; NR, not reported; NS, not significant; ROR, Risk of Recurrence; RS, Recurrence Score.

Changes in Treatment Recommendation

Our literature search identified only three studies (all on Oncotype DX[80,82,83]) reporting on the impact of GEP tests for a purely LN+ population (Table 9). Among mixed LN populations, the impact of GEP tests on treatment changes were similar to those seen in the LN− group. The use of GEP tests generally led to changes in treatment of up to 40%.

Table 9: Changes in Treatment Recommendations in a Lymph-Node–Positive Population

Author, Year	Test	N	No CT to CT, n (%)	CT to No CT, n (%)	Total Treatment Change, n (%)
Bargallo et al, 2015[80]	Oncotype DX	34	2/34 (6)	11/34 (32)	13/34 (38)
Dieci et al, 2018[82]	Oncotype DX	126	5/126 (4)	20/126 (16)	25/126 (20)
Eiermann et al, 2013[83]	Oncotype DX	122	11/122 (9)	34/122 (28) CT to CET: 2/122 (2)	47/122 (39, 95% CI 29.9–47.8)
Torres et al, 2018[103]	Oncotype DX	67	3/67 (4)	21/67 (31)	24/67 (36, 95% CI 24–48)
Mixed LN Status					
Ettl et al, 2017[104]	EndoPredict	395	150/395 (38)	20/395 (5)	170/395 (43)
Fallowfield et al, 2018[105]	EndoPredict	149	28/149 (19)	27/149 (18)	55/149 (36.9)
Mokbel et al, 2017[106]	EndoPredict	120	9/120 (8)	8/120 (7)	17/120 (14)
Mokbel et al, 2018[107]	EndoPredict	120	9/41 (22)	28/79 (35)	37/120 (31)
Cusumano et al, 2014[108]	MammaPrint	453	68/453 (15)	75/453 (17)	143/453 (32)
Tsai et al, 2018[109]	MammaPrint	840	172 (38)	110 (29)	282 (34) OR 0.64 (0.50–0.82)
Wuerstlein et al, 2019[49]	MammaPrint	430	105/430 (24)	157/430 (37)	262/430 (61)
Curtit et al, 2019[48]	Oncotype DX	882	NR (13)	NR (61)	NR (44)
de Boer et al, 2013[81]	Oncotype DX	50	1/50 (2)	12/50 (24)	13/50 (26)
Evans et al, 2016[110]	Oncotype DX	193	Post CT: 47/193 (24)	NR	NR
Kuchel et al, 2016[111]	Oncotype DX	135	12/135 (9)	43/135 (32)	55/135 (41)
Leung et al, 2016[112]	Oncotype DX	146	3/146 (2)	24/146 (16)	34/146 (23; 95% CI 17–31) including 7/146 (5) changes in CT intensity
Loncaster et al, 2017[86]	Oncotype DX	65	0/65	45/65 (69)	45/65 (69)
Martinez del Prado et al, 2018[113]	Oncotype DX	401	9/401 (2)	133/401 (33)	142/401 (35)
Pestalozzi et al, 2017[114]	Oncotype DX	221	8/221 (4)	37/221 (17)	45/221 (20)
Voelker et al, 2018[115]	Oncotype DX	50	3/50 (6)	5/50 (10)	8/50 (16)

Abbreviations: CI, confidence interval; CT, chemotherapy; NR, not reported; OR, odds ratio.

Physician Confidence in Treatment Recommendations

Six studies were found that evaluated physician confidence in treatment recommendations among a mixed LN or LN+ population (Table 10). All studies showed that Oncotype DX testing increased physician confidence. Authors suggested that the confidence decrease among some physicians may be due to a hesitancy to withhold chemotherapy, particularly for people who are intermediate risk.

Table 10: Physician Confidence in Treatment Recommendations in a Lymph-Node–Positive Population

Author, Year	Test	Physician Confidence in Treatment Recommendations
Wuerstlein et al, 2019[49]	MammaPrint	The percentage of physicians with complete or high confidence increased overall from 68.6% to 85.1% Among low risk MammaPrint cases: 65.0%–83.4% increase Among high risk MammaPrint cases: 70.7%–84.9% increase
Bargallo et al, 2015[80]	Oncotype DX	66% physicians strongly agreed that they felt more confident 26% of physicians agreed that they felt more confident 8% of physicians neither agreed nor disagreed 0% of physicians disagreed or strongly disagreed
Dieci et al, 2018[82]	Oncotype DX	87% physicians agreed or strongly agreed that they were confident in treatment recommendations post-test
Eiermann et al, 2013[83]	Oncotype DX	Overall, 55% vs. 82% physicians felt absolute or high confidence 46% increased confidence for lymph-node–positive cases
Kuchel et al, 2016[111]	Oncotype DX	Increased from 49% to 81%
Torres et al, 2018[103]	Oncotype DX	Overall change in confidence: 64% pre-test vs. 88% post-test (*P* < .001) • 49% increased confidence • 40% no change • 11% decreased confidence For low Recurrence Score (*P* = .002) • 56% increased confidence • 34% no change • 10% decreased confidence For intermediate Recurrence Score (*P* = .10) • 39% increased confidence • 48% no change • 13% decreased confidence For high Recurrence Score (*P* = NS) • 50% increased confidence • 50% decreased confidence

Ongoing Studies

We searched ClinicalTrials.gov for relevant ongoing studies on EndoPredict, MammaPrint, Oncotype DX, and Prosigna. Potentially relevant ongoing studies are in Appendix 7. One study (the RxPONDER study; NCT01272037) evaluating Oncotype DX's ability to predict chemotherapy benefit for LN+ breast cancer is expected to be complete in 2022. The RxPONDER study is similar to the TAILORx trial,[26] but focuses specifically on the LN+ population to fill a gap in the current evidence. Cost-effectiveness research is also an integral component of RxPONDER.

The long-term results of the MINDACT trial (NCT00433589) are expected to be complete in 2020. From the ISRCTN (International Clinical Trials Registry Platform) trial registry, we also found the OPTIMA trial, which will evaluate the performance and cost-effectiveness of different prognostic tests (e.g., Oncotype DX, Prosigna, MammaPrint) to establish a method of selecting patients who are likely to benefit from chemotherapy.[116] The OPTIMA study is estimated to complete in September 2023.

Ongoing research is also being conducted to examine additional uses of GEP tests (other than for chemotherapy treatment in early-stage invasive breast cancer), such as for tailoring radiation therapy or use in ductal carcinoma in situ (stage 0 breast cancer, also known as DCIS, an early form of breast cancer that is noninvasive).

We searched PROSPERO for ongoing systematic reviews and found no other ongoing reviews.

Discussion

Prognostic Ability

Our results show that GEP tests can likely prognosticate the risk of distant recurrence, particularly for the ER/PR+, HER2−, LN− breast cancer population. The evidence for the LN+ population was more limited, and the results show weaker prognostic ability, but GEP tests may be prognostic in certain LN+ populations. Currently ongoing studies are exploring the prognostic ability of GEP tests for this population. Although the four GEP tests (EndoPredict, MammaPrint, Oncotype DX, and Prosigna) evaluate the expression of different genes that have little overlap and use different risk categories and cut-offs, they were all found to be able to use molecular factors such as cancer proliferation and invasiveness to categorize people into groups based on low or high risk of distant recurrence. The prognostic value of GEP tests is also reflected in Ontario and international guidelines that recommend their use for adjuvant chemotherapy decision-making in people with early-stage invasive breast cancer.

Our results on prognostic ability align with other recent health technology assessments and systematic reviews (see Appendix 3).

Predictive Ability

One of the strengths of our review is the inclusion only of prospective study designs (either prospective studies or retrospective analyses of prospective studies). We also included studies based on prospectively maintained large cancer registries. We excluded retrospective study designs because of the potential methodological limitations and biases (e.g., tumour specimen collection and analysis).

Currently, only MammaPrint and Oncotype DX have RCT evidence that show the predictive ability of each test. The evidence was primarily from nonrandomized studies. Low-risk groups in most nonrandomized studies showed high freedom from distant recurrence and survival without chemotherapy, which can support the potential to withhold chemotherapy for benefit. However, given the weaker quality of the evidence for predictive ability, this information should be interpreted by physicians in the context of other clinicopathological features for treatment decision-making.

Comparative Effectiveness

We found very limited evidence on the comparative effectiveness between the four GEP tests. The TransATAC study[52] includes only three GEP tests (EndoPredict, Oncotype DX, and Prosigna). There is no comparable MammaPrint study conducted within a similar patient population. The seminal MINDACT trial for MammaPrint[51] was specifically designed to evaluate the value of withholding chemotherapy for people with discordant clinical and genomic (MammaPrint) risk scores: either people who had high clinical risk and low genomic risk, or people who had high genomic risk and low clinical risk. No other study has specifically examined the use of GEP tests for the low clinical risk and high genomic risk population. However, due to the study design, the MINDACT trial included a patient population that was generally higher risk than the TransATAC population. Additional studies are required for conclusions to be drawn about which of the four GEP tests may be best for a specific population.

Recent Changes in the Oncotype DX Risk Categories

In response to the results of the TAILORx trial,[26] the official Oncotype DX categories have now been revised to a two-risk category (low RS ≤ 25 and high RS ≥ 26) for postmenopausal people who are LN− (from an original three-risk category—low RS < 18, intermediate RS 18–30, and high RS ≥ 31). Since TAILORx was recently published (2018), there is limited information on the impact of the new risk categories in practice. Many of the studies, including the comparative TransATAC study from Sestak et al (2018)[52] use the previous three-risk categories and are therefore likely less generalizable. Comparative studies that compare other GEP tests with the new TAILORx trial cut offs are needed to evaluate the impact of this change.

Changes in Treatment Decisions

Gene expression profiling tests were shown to greatly impact treatment decisions. There was generally a 20% to 50% change in treatment decisions pre- and post-testing, with changes both in recommending and in withholding chemotherapy (some studies showed a > 60% change). While recommendation changes were reported within the studies, only some studies reported on actual chemotherapy use (i.e., whether the chemotherapy was in fact administered as recommended). Details surrounding physician–patient discussions of treatment changes were often not described, and it is unclear if treatment decision-making involved other factors unrelated to the results of the GEP test.

Conclusions

In the LN− patient population, GEP tests are likely prognostic for freedom from distant recurrence (GRADE: Moderate) and may be prognostic for disease-free and overall survival (GRADE: Low). In the LN+ patient population, GEP tests may be prognostic for freedom from distant recurrence (GRADE: Low). They may also be prognostic for disease-free and overall survival (GRADE: Very Low), but we are very uncertain. Some GEP tests may predict chemotherapy benefit in the LN− population (GRADE: Low). They may also predict chemotherapy benefit in the LN+ population (GRADE: Very Low), but we are very uncertain. Gene expression profiling tests may lead to changes in treatment recommendations (GRADE: Low). The GEP tests may also increase physician confidence in treatment recommendations (GRADE: Very Low), but we are very uncertain.

ECONOMIC EVIDENCE

Research Question

What is the cost-effectiveness of four gene expression profiling (GEP) tests (EndoPredict, MammaPrint, Oncotype DX, Prosigna) compared with usual care or compared with one another for people with early-stage invasive breast cancer?

Methods

Economic Literature Search

We identified two systematic reviews assessing the cost-effectiveness of GEP tests versus usual care or one another in people with breast cancer through a scoping search.[117,118] The literature searches for the two reviews were conducted in April 2016[117] and March 2017.[118] We used the two systematic reviews to identify eligible studies published from inception until January 2016. We performed an economic literature search on December 4, 2018, to retrieve studies published from January 2016 until the search date. To retrieve relevant studies, we developed a search using a modified version of the clinical search strategy with an economic and costing filter applied. In addition to the databases used for the clinical search, we also used the Ovid interface in the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Health Technology Assessment database, and the National Health Service Economic Evaluation Database (NHSEED).

We created database auto-alerts in MEDLINE and Embase, and monitored them for the duration of the assessment period. We also 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. The grey literature search was updated on May 2–3, 2019. See Appendix 4 for our literature search strategies, including all search terms.

Eligibility Criteria

Studies

Inclusion Criteria

- English-language full-text publications
- Studies published between January 2016 and December 4, 2018
- Cost–utility, cost-effectiveness, cost–benefit, or cost-consequence analysis
- Budget impact analysis

Exclusion Criteria

- Reviews, letters/editorials, commentaries
- Abstracts, posters
- Unpublished studies

Population

- People with early-stage invasive breast cancer

Interventions

- EndoPredict
- MammaPrint
- Oncotype DX
- Prosigna (PAM50)

Outcome Measures

- Costs
- Quality-adjusted life-years (QALYs)
- Incremental costs
- Incremental effectiveness (i.e., incremental QALYs)
- Incremental cost per QALY

Literature Screening

A single reviewer conducted an initial screening of titles and abstracts using Covidence[42] 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. The reviewer also examined reference lists for any additional relevant studies not identified through the search.

Data Extraction

For all studies, including those identified through previous systematic reviews, we extracted relevant data on study characteristics and outcomes to collect information about the following:

- Source (e.g., citation information, study type)
- Methods (e.g., study design, analytic technique, perspective, time horizon, population, intervention[s], comparator[s])
- Outcomes (e.g., health outcomes, costs, incremental cost-effectiveness ratios)

Study Applicability and Limitations

We determined the usefulness of each identified study for decision-making by applying a modified quality appraisal checklist for economic evaluations originally developed by the National Institute for Health and Care Excellence (NICE) in the United Kingdom to inform the development of NICE's clinical guidelines.[119] We modified the wording of the questions to remove references to guidelines and to make it specific to Ontario. Next, we separated the checklist into two sections. In the first section, we assessed the applicability of each Canadian study to the research question (directly, partially, or not applicable). In the second section, we assessed the limitations (minor, potentially serious, or very serious) of the Canadian studies that we found to be directly applicable.

Results

Economic Literature Search

Our economic literature search yielded 113 citations published between January 2016 and December 2018. We identified 29 from other sources. We screened 86 records after removing duplicates. From this, we identified 12 economic evaluations that met our inclusion criteria.[86,113,120-128] We identified 46 additional eligible studies[129-174] published before 2016 from the reference lists of two previously published systematic reviews[117,118] and another primary economic evaluation in the diagnostics guidance report published by NICE.[118] Figure 3 presents the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram for the economic literature search.

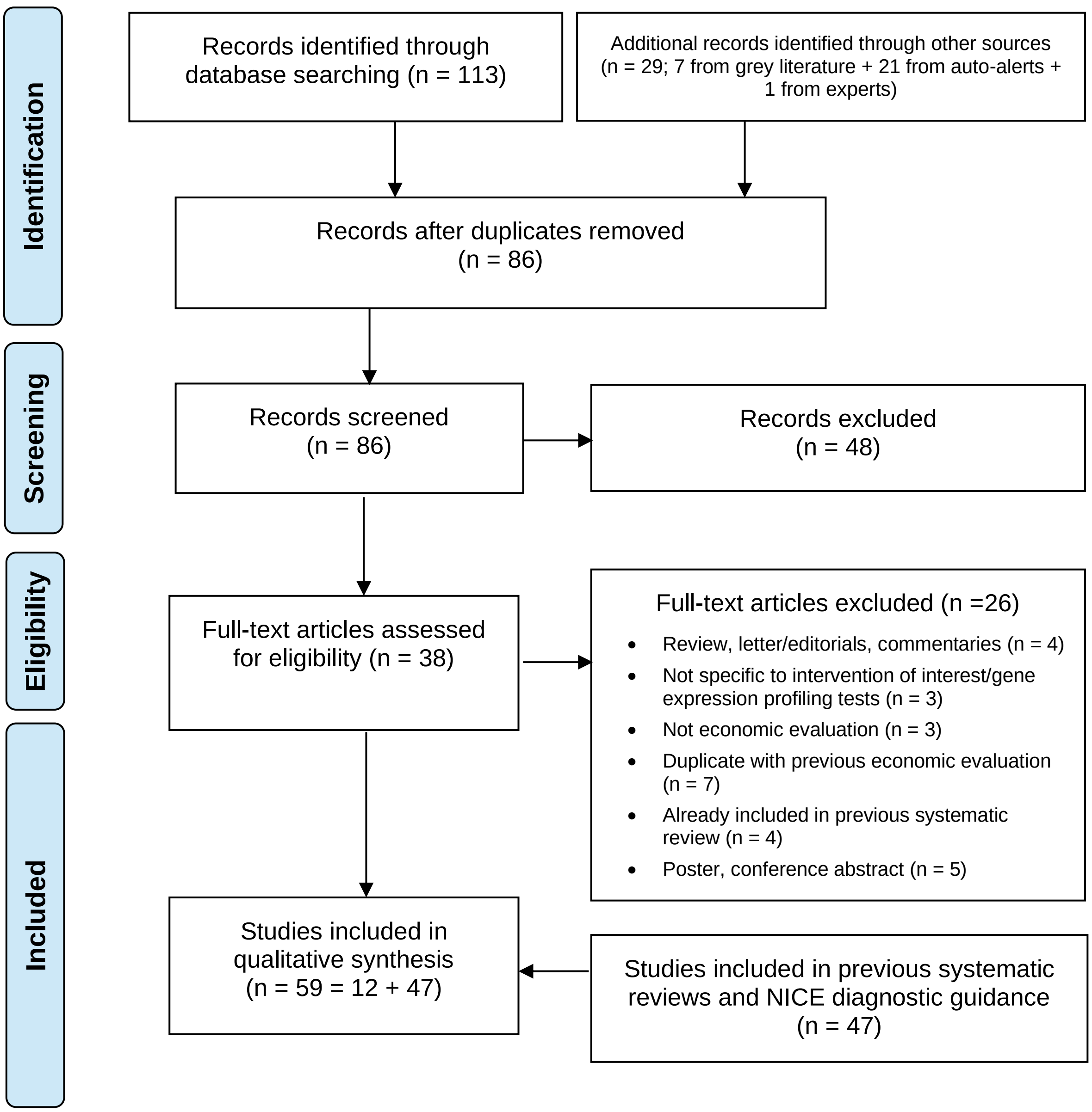


Figure 3: PRISMA Flow Diagram—Economic Search Strategy

Source: Adapted from Moher et al, 2009.[50]

Overview of Included Economic Studies

Of the 59 included studies, we identified 47 cost–utility or cost-effectiveness analyses, and 13 cost comparison or budget impact analyses.

Test Strategies Included in Economic Evaluations

All cost–utility and cost-effectiveness analyses used a model-based approach. Among these, four compared EndoPredict to no GEP test,[118,122,132,175] 11 compared MammaPrint to no GEP test,[118,134,135,139,150,157,160-162,166,172] 34 studies compared Oncotype DX to no GEP test,[118,120,123,128,129,131,133,136,137,140-149,151-155,158,159,166,168-173,175] and three compared Prosigna to no GEP test.[118,166,175] Additionally, three studies compared MammaPrint to Oncotype DX,[163,164,174] one study[166] compared MammaPrint, Oncotype DX, and Prosigna, and another one compared EndoPredict, Oncotype DX, and Prosigna.[175]

In addition to the cost–utility and cost-effectiveness analyses, we included 13 reports that compared the cost of using and not using GEP tests.[86,113,121,124-127,130,138,151,156,165,167] All of them were on Oncotype DX. One also assessed the impact of introducing all four tests.[124]

Study Population

For information about treatment of gender-related issues in the data, see note on Terminology in Background, above.

All cost–utility and cost-effectiveness analyses were conducted in people with early-stage breast cancer. The majority (two for EndoPredict, 9 of 11 for MammaPrint, 25 of 34 for Oncotype DX, and two for Prosigna compared the GEP test with no test, and four head-to-head comparisons between GEP tests) were conducted in people who have LN− breast cancer. Fourteen studies were conducted with people with LN+ breast cancer (alone or in combination with LN− breast cancer). Among these, three assessed EndoPredict, three assessed MammaPrint, 11 assessed Oncotype DX, and one assessed Prosigna.

Model Structure

The model-based cost–utility and cost-effectiveness analyses, at minimum, considered three health states: no recurrence/disease free, distant recurrence, and death. In the “no recurrence” health state, people either received adjuvant chemotherapy or they did not. Additional health states considered by some (but not all) analyses included local recurrence, chronic heart failure, and myeloid leukemia.

Perspectives and Time Horizons

The majority of cost–utility and cost-effectiveness analyses used a health care payer perspective. The time horizons ranged from 10 years to lifetime, with most analyses evaluating lifetime costs and outcomes.

Other Main Assumptions

The majority of included studies assumed that most or all people classified as high risk would accept chemotherapy (a majority of the evaluations used an uptake rate between 90% and 100%). Further, they assumed that most or all people classified as low risk would decline chemotherapy.

In addition to the ability to classify Risk of Recurrence, there has been evidence of the ability of GEP tests to predict the benefit of adjuvant chemotherapy, though the quality of evidence is, in general, poor. Most economic evaluations of Oncotype DX either explicitly assumed that Oncotype DX also predicted the benefit of chemotherapy or they used studies[56,90] assuming its predictive ability to populate their models, despite the poor quality of evidence.

Cost–Utility or Cost-Effectiveness Evidence on the Tests

EndoPredict

We identified one published economic evaluation conducted in the Canadian setting. Hannouf et al[175] estimated an incremental cost-effectiveness ratio (ICER) for EndoPredict of $36,274 versus no test for people with LN−, HER2−, early-stage breast cancer. Outside of Canada, two studies and the NICE diagnostics guideline assessed the cost-effectiveness of EndoPredict compared with no GEP test. A German study concluded that, compared with decision making based on clinical guidelines, EndoPredict, either in combination with clinical guidelines or used alone, was dominant (less costly, more QALYs gained).[132] We found a UK-based cost–utility analysis estimating that the ICER of EndoPredict compared with no GEP test was £26,836 per QALY for people classified as intermediate risk based on a clinical assessment such as the Nottingham Prognostic Index.[122] The German and UK evaluations were on a mixed population, including people with LN− and LN+ breast cancer. NICE concluded that EndoPredict was not cost-effective compared to no test for people with LN− breast cancer and a Nottingham Prognostic Index (NPI) of 3.4 or less.[118] However, the same evaluation for people with LN+ breast cancer suggested that EndoPredict may be cost-effective compared to no test.[118] Based on these studies, EndoPredict may be cost-effective in LN+ patients, the cost-effectiveness of EndoPredict compared with no GEP test in LN− patients was inconclusive due to conflicting results.

MammaPrint

We did not identify any published economic evaluations conducted in an Ontario or Canadian setting. Outside of Canada, 11 studies compared MammaPrint to no GEP test, including eight on LN− breast cancer only,[134,135,139,150,160-162,172] and three on both LN− and LN+ breast cancer.[118,157,166]

Seven of these studies concluded that, for people with LN− breast cancer, MammaPrint was cost-effective compared to usual care or Adjuvant! Online.[135,139,150,160-162,172] Bonastre and colleagues[134] assessed MammaPrint in node-negative breast cancer and found it was unlikely to be cost-effective compared with Adjuvant! Online.

One analysis comparing MammaPrint with a “no test and chemotherapy for all” strategy in both LN− and LN+ breast cancer patients supported its cost-effectiveness (ICER: about $1,900 CAD per QALY).[166] In contrast, Oestreicher et al[157] compared MammaPrint with usual care without GEP test in LN− and LN+ patients and suggested that MammaPrint decreased the lifetime cost but also had fewer QALYs. The results favoured usual care over MammaPrint.[157] In the NICE economic evaluation, the clinical parameters for MammaPrint were derived from the MINDACT trial.[51] The NICE diagnostics guideline concluded that GEP testing was not cost-effective compared with modified Adjuvant! Online.[118]

Oncotype DX

Oncotype DX is the most commonly assessed test, both within and outside Canada.

Economic Evidence on Oncotype DX in the Canadian Context

We found nine studies conducted in a Canadian context[129,137,141,142,153,158,168,169,175] (Table 11). All studies were model-based cost–utility analyses that compared Oncotype DX to no GEP test. Four reports[129,158,168,169] used similar or even the same model structure and, in these analyses, Adjuvant! Online was used as a clinical tool to assess risk and aid decision-making.

The study populations varied. Seven studies included only people with LN− breast cancer. Two analyzed both LN− and LN+ breast cancer.[142,153] As for menopausal status, Hannouf et al (2014)[142] focused specifically on people who are postmenopausal, while Hannouf et al (2012)[141] considered pre- and postmenopausal populations. The other studies did not specify the stage of menopause.

Eight analyses concluded that Oncotype DX is likely to be cost-effective compared with no test from a Canadian health care payer perspective.[129,137,141,142,153,158,168,169] However, the ICERs varied across subgroups (with the mixed population studies finding an ICER of $464 to $14,844 per QALY gained for people with LN+ breast cancer).[141,142,153] In general, the analyses suggested that the ICERs were lowest for people who are premenopausal or have a high clinical risk. In their analysis of Oncotype DX in both pre- and postmenopausal women, Hannouf et al[141] found that Oncotype DX dominated the no GEP test strategy for people who are premenopausal, but the ICER for those who are postmenopausal was approximately $60,000 per QALY gained. An analysis assessing the cost-effectiveness of Oncotype DX in combination with Adjuvant! Online compared with Adjuvant! Online alone estimated that the ICERs were $22,440, $2,526, and $1,111 for low-, intermediate-, and high-risk subgroups, respectively.[158]

Economic Evidence on Oncotype DX in Other Contexts

Most cost–utility and cost-effectiveness analyses showed that, compared to no GEP test, the use of Oncotype DX increased the total cost from a healthcare payer perspective.[123,131,133,136,140,143-149,151,152,154,155,159,166,170,171,173] All but four analyses[118,120,128,155] concluded that Oncotype DX was cost-effective regardless of lymph node status. The ICER varied across subgroups in and out of the Canadian context. In general, the analyses suggested lower ICERs in people who are premenopausal and in people with high clinical risk.[118,141,158,168,172] Wang et al[128] assessed the cost-effectiveness of Oncotype DX versus PREDICT, a clinical assessment tool, and found that Oncotype DX is not cost-effective for people with low clinical risk. NICE conducted a model-based primary economic evaluation to assess the cost-effectiveness of all tests relevant to the UK setting. The clinical parameters for Oncotype DX were derived from the TransATAC study.[52] NICE concluded that Oncotype DX was not cost-effective compared with no test for people with LN− breast cancer and an Nottingham Prognostic Index (NPI) of 3.4 or less.[118]

Table 11: Results of Economic Literature Review—Summary

Author, Year Country of Publication	Analytic Technique, Study Design, Perspective, Time Horizon	Population	Intervention(s) and Comparator(s)	Results: Health Outcomes	Results: Costs	Results: Cost-Effectiveness
Davidson 2013 Canada[137]	• CUA and CEA • Markov model • Health care system perspective • Lifetime horizon	50-year-old women with ER+, PR−, LN−, HER2− early-stage breast cancer	• Oncotype DX • No test	Oncotype DX: 17.72 LY, 17.26 QALYs No test: 17.41 LY, 16.94 QALYs LYG: 0.31 LY Δ QALY: 0.32 Discount rate: 5%	2010 CAD Oncotype DX: $15,395 No test: $13,027 Δ cost: $2,373 Discount rate: 5%	$6,995.37/LY $6,630.38/QALY
Hannouf 2019 Canada[175]	• CUA • Markov model • Canadian public health care system perspective • Lifetime horizon	Women with HR+, LN−, HER2− early-stage breast cancer	• EndoPredict • Oncotype DX • Prosigna • No test	*Δ QALY* EndoPredict: 0.08 Oncotype DX: 0.04 Prosigna: 0.06	*Δ Cost* EndoPredict: $2,720 Oncotype DX: $3,496 Prosigna: $2,992	EndoPredict: $36,274/QALY Oncotype DX: $74,911/QALY Prosigna: $48,525
Hannouf 2014 Canada[142]	• CUA • Markov decision model (decision tree and Markov decision model: based on an actual study cohort) • Canadian public health care system perspective • Lifetime horizon	n = 161 Median age (range), y: 61 (50–89) Postmenopausal women (defined as age ≥ 50 y) People with HR+, LN 1–3, early-stage breast cancer (stage II/III)	• Oncotype DX • Canadian clinical practice guiding strategy	Oncotype DX: 15.81 QALYs Canadian clinical practice strategy: 15.73 QALYs Δ QALY: 0.08 Discount rate: 5%	2012 CAD Oncotype DX: $49,129.20 Canadian clinical practice strategy: $49,093.00 Δ cost: $36.20 Discount rate: 5%	$464 per QALY gained
Hannouf 2012 Canada[141]	• CUA • Markov decision model (decision tree and Markov decision model: based on an actual study cohort) • Canadian public health care system perspective • Lifetime horizon	Total N = 498[a] *Median age (range), y* Premenopausal women: 44 (29–49) Postmenopausal women: 62 (52–88) People with HR+, LN–, early-stage breast cancer	• Oncotype DX • Canadian clinical practice guiding strategy	*Δ QALY* Premenopausal women: 0.05 Postmenopausal women: 0.062 Discount rate: 5%	2010 CAD *Δ Cost* Premenopausal women: −$50 Postmenopausal women: $3,700 Discount rate: 5%	Premenopausal women: dominant[b] Postmenopausal women: ~$60,000/QALY

Author, Year Country of Publication	Analytic Technique, Study Design, Perspective, Time Horizon	Population	Intervention(s) and Comparator(s)	Results		
				Health Outcomes	Costs	Cost-Effectiveness
HQO/Medical Advisory Secretariat 2010 Canada[129]	• CUA • Markov decision model • Ontario Ministry of Health perspective • Lifetime horizon	50-year-old Ontario women diagnosed with ER+, LN–,HER2/neu– early-stage breast cancer	• Oncotype DX assay • AOL	No Oncotype DX, only AOL: 13.34 QALYs *Oncotype for AOL risk groups* High: 14.04 QALYs Intermediate/high-risk: 14.42 QALYs All: 14.64 QALYs *Δ QALY using Oncotype for AOL risk groups* High: 0.70 QALYs Intermediate/high: 1.08 QALYs All: 1.30 QALYs	No Oncotype DX: $13,298 *Oncotype for AOL risk groups* High: $13,660 Intermediate/high: $13,961 All: $17,466 *Δ cost using Oncotype DX for AOL risk groups* High: $362 Intermediate/high: $663 All: $4,169	*ICER using Oncotype DX for AOL risk groups* High: $518/QALY Intermediate/high: $795/QALY All: $23,983/QALY
Lamond 2012 Canada[153]	• CUA • Decision analytic model (decision tree and Markov decision model) • Third-party direct payer perspective • 25-year horizon	Women with early-stage, endocrine-sensitive breast cancer undergoing adjuvant chemotherapy or no chemotherapy	• Oncotype DX • No test	*Δ QALY* LN–: 0.27 LN+: 0.06 Combined: 0.18 Discount rate: 3%	2010 CAD *Δ Cost* LN–: $2,585 LN+: $864 Combined: $1,852 Discount rate: 3%	LN–: $9,591/QALY LN+: $14,844/QALY Combined: $10,316/QALY

Author, Year Country of Publication	Analytic Technique, Study Design, Perspective, Time Horizon	Population	Intervention(s) and Comparator(s)	Results		
				Health Outcomes	Costs	Cost-Effectiveness
Paulden 2013 Canada[158]	• CUA and CEA • Decision analytic model (decision tree and Markov decision model) • Ontario public payer perspective • Lifetime horizon	50-year-old women with LN–, HR+, HER2/neu– early-stage breast cancer	• Oncotype DX assay • AOL	No Oncotype DX, only AOL: 11.063 QALYs *Oncotype DX for AOL risk groups* High: 11.276 QALYs Intermediate: 11.193 QALYs Intermediate/high: 11.407 QALYs Low: 11.147 QALYs Low and high: 11.361 QALYs Low and intermediate: 11.278 QALYs All: 11.492 QALYs Discount rate: 5%	2012 Canadian dollars No Oncotype DX, only AOL: $13,860 *Oncotype DX for AOL risk groups* High only: $14,090 Intermediate only: $14,190 Intermediate and high: $14,420 Low only: $15,750 Low and high: $15,990 Low and intermediate: $16,080 All: $16,320	*ICER using Oncotype DX for AOL risk groups* High only: $1,111/QALY Intermediate only: dominated by AOL high risk only strategy[c] Intermediate and high: $2,526/QALY Low only: dominated by AOL high risk–only strategy and by intermediate/high-risk strategy[c] Low and high: dominated by intermediate/high-risk strategy[c] Low and intermediate: dominated by intermediate/high-risk strategy[c] All: $22,440/QALY
Tiwana 2013 Canada[168]	• CUA and CEA • Decision analytic model (decision tree and Markov decision model) • Alberta public payer perspective • Lifetime horizon	50-year-old women diagnosed with LN–, HR+, HER2/neu– early-stage breast cancer	• Oncotype DX assay • AOL	*Δ QALY using Oncotype DX for AOL risk groups (no Oncotype DX as reference)* High risk only: 0.228 Intermediate only: 0.161 Intermediate and high risk: 0.389 Low risk only: 0.109 Low and high risk: 0.337 Low and intermediate risk: 0.270 All: 0.498	2012 Canadian dollars *Δ cost using Oncotype DX for AOL risk groups (no Oncotype DX as reference)* High risk only: $340 Intermediate only: $130 Intermediate and high risk: $460 Low risk only: $1,430 Low and high risk: $1,770 Low and intermediate risk: $1,550 All: $1,890	*ICER of using Oncotype DX for AOL risk groups (no Oncotype DX as reference)* High risk only: $1,476/ QALY Intermediate only: $764/QALY Intermediate and high risk: $1,182/QALY Low risk: $13,116/QALY Low and high risk: $5,234/QALY Low and intermediate risk: $5,749/QALY Oncotype DX for all: $3,789/QALY

Author, Year Country of Publication	Analytic Technique, Study Design, Perspective, Time Horizon	Population	Intervention(s) and Comparator(s)	Results		
				Health Outcomes	Costs	Cost-Effectiveness
Tsoi 2010 Canada[169]	• CUA and CEA • Decision analytic model (decision tree and Markov decision model) • Health care payer perspective • Lifetime horizon (to a maximum age of 100 years)	50-year-old woman with LN–, HR+, HER2– early-stage breast cancer	• Oncotype DX assay • AOL	Oncotype DX: 13.638 QALYs, 13.997 LYs AOL only: 13.573 QALYs; 13.933 LYs Δ QALY (AOL as reference): 0.065 QALY LYG (AOL as reference): 0.064 Discount rate: 5%	2008 CAD Oncotype DX: $19,747 AOL only: $15,645 Δ cost (AOL as reference): $4,102 Discount rate: 5%	$63,064/ QALY $63,911/LYG

Abbreviations: Δ, incremental; AOL, Adjuvant! Online; CEA, cost-effectiveness analysis; CUA, cost–utility analysis; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor (estrogen receptor or progesterone receptor); ICER, incremental cost-effectiveness ratio; LN, lymph node; LY, life-year; LYG, life-year gained; PR: progesterone receptor; QALY, quality-adjusted life-year; RS, Recurrence Score.

[a]109 premenopausal and 389 postmenopausal women in Manitoba from January 1, 2000, to December 31, 2002.

[b]Dominant: a strategy is dominant over another strategy when it has both more QALYs and less cost.

[c]Dominated: a strategy is dominated by another strategy when it has both with less QALYs and higher cost.

Prosigna

Hannouf et al[175] evaluated the cost-effectiveness of Prosigna in the Canadian setting and reported an ICER of $48,525 per QALY for people with early-stage invasive, LN−, HER2− breast cancer. Outside of Canada, we identified two studies.[118,166] Based on a cost–utility analysis of the OPTIMA trial, Prosigna (test to help guide the chemotherapy decision) was dominant compared with chemotherapy for all populations (treatment for all). However, there was uncertainty in the risk classification results and for long-term outcomes.[166] In addition, the NICE diagnostic guideline concluded that Prosigna was not cost-effective compared to no test for people with LN− breast cancer and an NPI of 3.4 or less.[118] The same analysis suggested that it may be cost-effective in people with LN+ breast cancer.

Comparison Between Tests

We have identified limited evidence on the relative effectiveness of the different GEP tests. Hannouf et al[175] suggested that EndoPredict and Prosigna were cost-effective compared to Oncotype DX in Canadian setting. MammaPrint was found to be cost-effective compared with Oncotype DX[175] and even dominant in analyses from the Netherlands[164] and the United States.[174] While Stein et al[166] compared three GEP tests (Oncotype DX, Prosigna, and MammaPrint) with chemotherapy for all strategies in the UK setting and concluded that Oncotype DX and Prosigna had lower costs and were more effective compared with MammaPrint, though all strategies had similar health outcomes (mean QALYs of: 7.89, 7.88, and 7.87 for Oncotype DX, Prosigna, and MammaPrint, respectively).

Applicability and Limitations of the Included Cost–Utility Analyses

Appendix 8 provides the results of the quality appraisal checklist for economic evaluations applied to the included studies. Of nine studies conducted using a Canadian perspective, four were considered directly applicable to our research question.[137,141,153,175] We assessed the limitations of these studies (Appendix 8, Table A15) and found that all of them had minor limitations. No study was identified that fully answered our research question with the most recent evidence. Eight of the nine Canadian studies assessed only the cost-effectiveness of Oncotype DX compared to usual care while the ninth study compared EndoPredict, Oncotype DX, and Prosigna versus usual care. No Canadian studies compared MammaPrint to usual care or compared different GEP tests to one another.

Budget Impact Analysis

Thirteen analyses assessed the cost of introducing GEP tests.[86,113,121,124-127,130,138,151,156,165,167] All 13 studies examined Oncotype DX, with one also examining EndoPredict, MammaPrint, and Prosigna.[124] The results varied considerably. An Ontario analysis compared the 20-month direct treatment cost for 998 people with breast cancer who either received Oncotype DX or did not.[126] This analysis concluded that while using Oncotype DX led to a 23% decrease in chemotherapy, it also led to an additional $3 million in direct health care costs. In contrast, six analyses suggested that introducing Oncotype DX was cost-saving.[86,125,127,130,156,165] One analysis compared the four GEP tests.[124] It suggested that Oncotype DX was the only test associated with cost savings to health care payer and society.

Discussion

Our review of economic evaluations identified 47 cost–utility or cost-effectiveness analyses. We used two previous systematic reviews to identify eligible studies published prior to January 2016. We then searched for studies published after this date. To our knowledge, this is the most comprehensive economic evidence review on this topic. The economic evaluations identified assessed cost-effectiveness across a variety of different settings. There was also variation in patient populations, health states included, and tests evaluated.

Most included analyses concluded that introducing GEP tests increased costs, led to better health outcomes, and were cost effective compared with no GEP testing. However, we interpret these results with caution due to limited evidence for important subgroup populations, variability in study designs, and limited applicability to the Ontario setting.

The prognosis of people with early-stage invasive breast cancer depends on factors such as menopausal and lymph node status. Few analyses considered these factors explicitly (i.e., through use of subgroups). Most analyzed a mixed population that included pre- and postmenopausal women (the recent TAILORx trial indicates that the interpretation of Oncotype DX scores is age-dependent[26]). Additionally, few analyses considered people with LN+ breast cancer (there is limited evidence assessing the impact of GEP tests on LN− and LN+ women using the same model, assumptions, and perspective).

There was variability in the comparators and health states included in the analyses. Although most of the analyses compared GEP tests to no test or to usual care, there was no clear standard of usual care. The most commonly used comparison, Adjuvant! Online, is no longer available for use in clinical practice. In addition, various health states were included in each of the analyses (e.g., some captured complications after chemotherapy such as chronic heart failure or myeloid leukemia). Gene expression profile tests are used to aide in adjuvant chemotherapy decisions. Therefore, the cost and utility parameters related to chemotherapy and how these parameters are modelled may influence the results.

We examined the applicability of economic evidence to the Ontario practice and the assumptions of included studies. The applicability of previous analyses to the Ontario or Canada settings is limited. First, limited evidence exists comparing the cost-effectiveness of GEP tests other than Oncotype DX in the same population in Ontario. Second, the patterns of chemotherapy uptake for low- and high-risk people considered in the economic evaluations were different from patterns observed in a recent Ontario study by Levine et al.[84] This study reported that 79% of people classified as high risk by Oncotype DX received adjuvant chemotherapy. In contrast, the economic evaluations described above assumed a higher adjuvant chemotherapy uptake rate (a majority of the evaluations used uptake rates between 90% and 100%). Third, the comparators and treatment regimens included in some of the studies do not represent current clinical practice. For example, Adjuvant! Online was included in four Canadian studies but it is no longer available for use. All analyses assessed the cost-effectiveness of Oncotype DX classifying people into low-, intermediate-, and high-risk categories (RS < 18, 18–30, and >30, respectively). The recent TAILORx trial[26] showed that, for people older than 50, a Recurrence Score between 0 and 25 indicates no chemotherapy benefit, while a Recurrence Score of ≥ 26 suggests chemotherapy benefit. The previously published evaluations cannot be reinterpreted to reflect this change in practice. Fourth, there is newly published clinical evidence for all four GEP tests (Oncotype DX, MammaPrint, Prosigna, and EndoPredict).[26,51,52] These data need to be incorporated into an economic evaluation.

Conclusions

Our review summarizes the current economic evidence on GEP tests, highlighting studies relevant to the Canadian context. Although most of the included analyses concluded that GEP tests are cost-effective, the economic evidence did not examine important subgroup populations, had variability in study design, and had limited generalizability to Ontario setting. Owing to these limitations, we conducted a primary economic evaluation.

PRIMARY ECONOMIC EVALUATION

Research Question

What is the cost-effectiveness of four gene expression profiling (GEP) tests (EndoPredict, MammaPrint, Oncotype DX, Prosigna) compared with usual care and one another, from the perspective of the Ontario Ministry of Health, in people with early-stage invasive, ER+, HER2–negative breast cancer?

Methods

The information presented in this report follows the reporting standards set out by the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement.[176]

Analysis

We conducted a cost–utility analysis using a Markov state-transition model. We conducted a reference case analysis and sensitivity analyses. Our reference case analysis adhered to the Canadian Agency for Drugs and Technologies in Health (CADTH) guidelines[177] when appropriate and represents the analysis with the most likely set of input parameters and model assumptions. Our sensitivity analyses explored how the results are affected by varying input parameters and model assumptions.

Because of limited clinical evidence on relative effectiveness between GEP tests and the different baseline risk levels of the study population on which the tests were evaluated, we included a cost–utility analysis of comparison between GEP tests as scenario analysis.

Target Population

In our reference case analysis, our target population was people who are postmenopausal, 58 years old, who have been diagnosed with early-stage invasive, LN−, ER+, HER2− breast cancer.[84,126] This population was chosen based on Cancer Care Ontario (CCO)[b] recommendations and clinical practice in Ontario.[29,30,84]

The CCO recommendations to use Oncotype DX, Prosigna, or EndoPredict for people with breast cancer were based on a 2016 systematic review of the clinical validity and clinical utility evidence.[29] We further limited our reference case analysis to LN− people because GEP tests are routinely funded in Ontario only for people who are LN− or who have micrometastasis.[29] We also conducted a subgroup analysis for people with LN+ breast cancer. Another important subgroup was people who are premenopausal, who are generally younger and with different risk profiles, and thus receive different treatment strategies.

In Ontario, other breast cancer populations are also eligible for and have requested GEP tests, but there is limited clinical evidence for utility. We assumed that the cost-effectiveness results are generalizable to all people with breast cancer.

---

[b] Cancer Care Ontario is now Ontario Health (Cancer Care Ontario).

Perspective

We conducted this analysis from the perspective of the Ontario Ministry of Health. This perspective includes all direct medical costs (e.g., outpatient care, inpatient care, physician billing, etc.).

Test Strategies

Our strategies of interest are the four commercially available GEP tests along with usual care:

- EndoPredict
- MammaPrint
- Oncotype DX (cut-off value: Recurrence Score of 25)
- Prosigna (formerly the PAM50 test)
- Usual care without a GEP test

We defined usual care as clinical decision-making without the use of a GEP test for people who are eligible. Usual care may include the use of clinical judgment, clinical practice guidelines, and clinical decision tools (i.e., the PREDICT tool, a free online program developed by the National Health Service in the United Kingdom and the University of Cambridge[9]). We assumed that clinical assessment results would not be used alongside GEP test results to make adjuvant chemotherapy decisions, nor as a triage to decide whether or not a person with early-stage invasive, ER+, HER2− breast cancer should receive a GEP test. Although people receiving usual care may have different risks, there is no risk classification for usual care. Therefore, we modelled a hypothetical cohort of people with average risk.

Discounting and Time Horizon

We used a lifetime horizon (50 years) and a 1-month cycle length in our reference case analysis. In accordance with the CADTH guidelines,[177] we applied an annual discount rate of 1.5% to both costs and quality-adjusted life-years (QALYs). We used varied discount rates (0%, 3%, and 5%) in the sensitivity analyses.

Model Structure

In collaboration with the Institute of Health Economics, we adapted the previous decision analytic model of Paulden et al,[158] who described the model in detail. The authors originally used the model to evaluate the cost-effectiveness of Oncotype DX in combination with Adjuvant! Online. More recently, it is being used to evaluate the cost-effectiveness of Oncotype DX versus Prosigna.[178]

For the present analyses, we retained the basic structure of the model, but we updated it to include all strategies of interest and recent, contextually relevant, clinical evidence. We used the Markov state transition model shown in Figure 4 to determine the incremental cost per QALY gained using a particular GEP test versus usual care or another test.[158] While the structure of the model is similar between comparators, we varied the probabilities of receiving adjuvant chemotherapy, natural history, utility, and cost parameters to reflect the data on each GEP test.

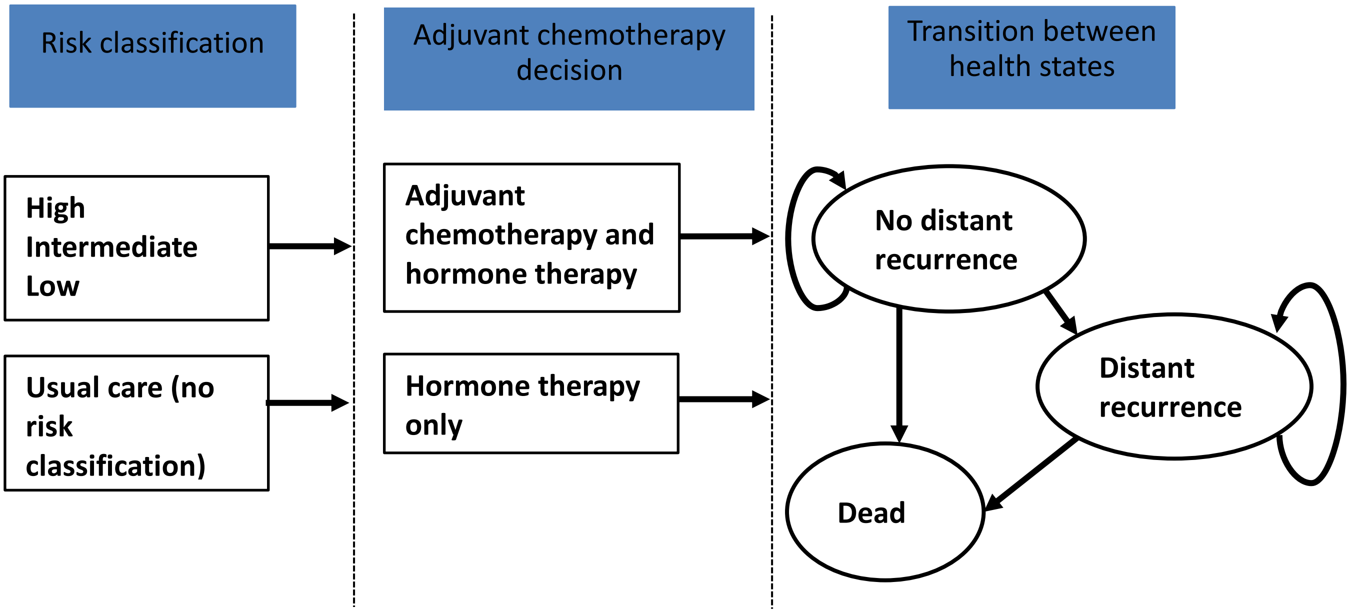


Figure 4: Markov Model Structure

Source: Adapted from Paulden et al, 2013.[158]

Risk Classification and Adjuvant Chemotherapy Decision

For Oncotype DX, EndoPredict, and MammaPrint, people who receive GEP tests are classified into two categories (high and low risk). For Prosigna, people who receive GEP tests are classified into three categories (high, intermediate, and low risk). There is no risk classification for people receiving usual care.

Based on the risk category, clinical judgment, and patient preference, people may or may not receive adjuvant chemotherapy. A Markov model was used to predict the lifetime QALYs and costs based on the adjuvant chemotherapy decision.

Health States in the Markov Model

The Markov model includes the following health states:

- No distant recurrence, on chemotherapy
- No distant recurrence in the first year, on hormone therapy
- No distant recurrence in the second and subsequent years
- Distant recurrence
- Dead

No Distant Recurrence, On Chemotherapy

People in this health state have no distant recurrence and receive adjuvant chemotherapy for 6 months. People may have fatal or nonfatal toxicities during adjuvant chemotherapy. Adjuvant chemotherapy impacts quality of life (decreased utility due to chemotherapy-related toxicities), costs (increased cost due to treatment and chemotherapy-related toxicities), and transition probabilities (lower risk of distant recurrence compared with those who do not accept chemotherapy). People in this health state continue to be free from distant recurrence or die. Owing to the low probability of developing distant recurrence while on adjuvant chemotherapy, we assumed people do not transition from this state directly to the distant recurrence state. This assumption is consistent with previous analyses.[158]

No Distant Recurrence in the First Year, on Hormone Therapy

People in this health state have no distant recurrence and have accepted hormone therapy, but not adjuvant chemotherapy, in the first year. It also applies to people who have finished their adjuvant chemotherapy in the first year after breast cancer diagnosis. People in this health state may continue to be free from distant recurrence, develop distant recurrence, or die.

No Distant Recurrence in the Second and Subsequent Years

People in this health state have no distant recurrence and may stay free of distant recurrence, develop distant recurrence, or die. Mortality rate and quality of life are independent of the adjuvant chemotherapy decision. However, we assumed that those who received adjuvant chemotherapy continued to have a lower risk of distant recurrence through the second and subsequent years compared with those who did not receive adjuvant chemotherapy.

Distant Recurrence

People in this health state have cancer that has metastasized to areas away from where the cancer was first located. People with distant recurrence may stay in the same state or die. We assumed people with distant recurrence would not recover or transition back to the “no distant recurrence” health state.

Dead

At any point, individuals have a probability of death due to age-specific background mortality, chemotherapy toxicity, and distant recurrence.

Main Assumptions

The model’s main assumptions were as follows:

Model Structure, and Impact of Gene Expression Profile Tests

- GEP tests influence clinical outcomes only by changing the probability of people receiving adjuvant chemotherapy in the reference case analysis. In the scenario analysis, we assumed the tests influence the clinical outcome by changing the probability of people receiving adjuvant chemotherapy and predicting the benefits of chemotherapy in individual cases
- There is no risk classification in the usual care group. For those who are considered eligible based on hormone receptor, LN, HER2 status, there is no further triage process to decide who should or should not receive a GEP test. We modelled a hypothetical cohort of people with an average risk that represents those who are eligible for but have not received GEP tests
- People classified in the same risk category (e.g., high risk) by different GEP tests have a different risk of recurrence, based on the published literature[51,52,56,57,61]

- People classified in the same risk category (e.g., high risk) by different GEP tests have the same probability of receiving chemotherapy
- GEP test results have no direct impact on quality of life.[142,158,169] People who are high-risk and have no distant recurrence have the same utility as people who are low-risk and have no distant recurrence given they both receive the same treatment
- Each person gets one GEP test. We do not consider retesting after recurrence. Our model also excludes people who have multiple breast cancer episodes over their lifetime
- The first cycle in the model was for risk classification. Everyone entered the model without distant recurrence. We assume people start the treatment (adjuvant chemotherapy or hormone therapy) in the second cycle of the model. In the second cycle, they can be without distant recurrence (under adjuvant chemotherapy or hormone therapy), with distant recurrence (under hormone therapy only), or dead

Treatment

- Everyone is postmenopausal and receives the same hormone therapy regimen, tamoxifen, for 5 years for low-risk, and 10 years for intermediate- and high-risk people. People in the usual care group receive tamoxifen for 7 years
- The length of adjuvant chemotherapy is 6 months regardless of risk level. In the scenario analysis, we assessed the impact of risk-dependent adjuvant chemotherapy regimens on cost-effectiveness
- Patients receive additional chemotherapy treatment or radiation therapy if distant recurrence occurs[179]
- Adjuvant chemotherapy has a constant effect on the risk of distant recurrence, regardless of risk classification results or GEP test used
- Adjuvant chemotherapy has a temporary impact on utility and cost, and temporarily increases mortality risk due to chemotherapy toxicities

Transition Between Health States

- The probability of transitioning from the “no distant recurrence” health state to the “distant recurrence” health state depends on the GEP test given, the test result (i.e., risk level determined), and whether adjuvant chemotherapy was given. But while receiving adjuvant chemotherapy, the probability of distant recurrence is low and we assume a probability of 0.[158] All other transition probabilities (e.g., transitions to death) are identical across the different comparisons
- There was no local recurrence in our model. We conducted a scenario analysis to test the potential impact of local recurrence on the cost-effectiveness
- People do not recover from distant recurrence. No one in our model transitioned from the “distant recurrence” health state to “no distant recurrence” health state

Clinical Outcomes Parameters

We used several input parameters to populate the model:

- Variables to model the impact of the GEP tests on the adjuvant chemotherapy decision

- Variables to model the natural history of breast cancer and the impact of adjuvant chemotherapy
- Variables to capture health-related quality of life

Impact of Gene Expression Profile Tests on Adjuvant Chemotherapy Decision-Making

The GEP tests classify people into different recurrence risk categories. The variables to model the impact of tests on the treatment decision of breast cancer are described in Table 12 and were developed from the following sources:

- The probability of being classified in each risk category for each GEP test
  - Probabilities were obtained from clinical studies identified in our clinical evidence review
  - For EndoPredict and Prosigna, the risk classification probabilities were obtained from a secondary analysis of the TransATAC study, which reported the risk classification of 591 participants with ER+, LN−, and HER2− breast cancer.[52] All participants in TransATAC were postmenopausal women
  - For MammaPrint, we used the MINDACT study, which compared the clinical utility of MammaPrint with standard pathological criteria to identify people unlikely to benefit from adjuvant chemotherapy.[51] The MINDACT study reported participants' baseline characteristics and outcomes according to four risk groups created based on both clinical and genetic risk levels (i.e., high or low clinical risk and high or low genetic risk). We used the proportion of these four risk categories from the subgroup of hormone receptor–positive, HER2−, and LN− breast cancer. This subgroup had a similar proportion of low clinical risk people (69.0% in the subgroup of MINDACT) as the TAILORx trial (69.9%) when assessed with the same clinical tool
  - For Oncotype DX, the risk classification probabilities were taken from the TAILORx trial.[26] The TAILORx trial recruited 10,273 women, both pre- and postmenopausal, with hormone receptor–positive, HER2−, and LN− breast cancer. For those ≥ 50 years of age, Oncotype DX classified people into two categories based on RS of: ≤ 25 and ≥ 26. We used the postmenopausal subgroup information for risk classification after Oncotype DX test
  - Those in the usual care group were not classified by risk group
- The probability of receiving adjuvant chemotherapy given each risk category
  - Based on a report of 979 people in Ontario who received the Oncotype DX test through the Ontario out-of-country program from 2012 to 2013,[84] 79.3%, 32.9%, and 4.1% of people who were classified as high-, intermediate-, and low-risk, respectively, received adjuvant chemotherapy. We assumed the same proportion of people in each risk category received adjuvant chemotherapy regardless of GEP test
  - In the same report, oncologists whose patients did not receive a GEP test recommended adjuvant chemotherapy for 21.7% of people, did not recommend chemotherapy for 44.6%, and were "unsure" for 33.7%[84]
  - In our usual care arm, we estimated that 38% of the reference case population received adjuvant chemotherapy based on clinical decision-making. We assumed that all people with a recommendation for chemotherapy and half of the people with an "unsure" recommendation received chemotherapy

Table 12: Parameter Inputs, Impact of Gene Expression Profiling Tests Used in the Economic Model

Variable	Probability, %	Range in DSA, %	Distribution[a]	Reference
Probability of Being Classified in Each Risk Category, by Test				
EndoPredict				
Low risk	72.6	60–80	Beta (429, 162)	Sestak et al, 2018[52]
High risk	27.4	—	NA	
MammaPrint				
Clinical low risk	69.0	50–100	Beta (2,916, 1,309)	Cardoso et al, 2016[51]
Low genetic risk in clinical low risk group	84.5	75–95	Beta (2,464, 452)	
Low genetic risk in clinical high risk group	54.7	45–65	Beta (716, 593)	
Clinical low risk				
Genetic low risk[a]	58.3	40–80	NA	
Genetic high risk[a]	10.7	5–20	NA	
Clinical high risk				
Genetic low risk[a]	16.9	0–30	NA	
Genetic high risk[a]	14.0	0–20	NA	
Oncotype DX				
Recurrence score ≤ 25	84.7 (≤ 10: 17.8; 11–25: 66.9)	≤ 10: 15–20 11–25: 60–80	≤ 10: Dirichlet (1, 1141) 11–25: Dirichlet (1, 4,296)	Sparano et al, 2018[26]
Recurrence score ≥ 26	15.3	—	Dirichlet (1, 982)	
Prosigna				
Low risk	53.8	45–65	Dirichlet (1, 318)	Sestak et al, 2018[52]
Intermediate risk	30.1	20–35	Dirichlet (1, 178)	
High risk	16.1	—	Dirichlet (1, 95)	
Proportion of People Receiving Adjuvant Chemotherapy, by Test and Risk Category				
Low risk[b]	4.1	0–10	Beta (23, 542)	Levine et al, 2016[84]
Intermediate risk[c]	32.9	20–40	Beta (106, 216)	
High risk[d]	79.3	60–100	Beta (73, 19)	
Unknown risk (usual care)	38.0	0–100	NA	

Abbreviations: DSA, deterministic sensitivity analysis; NA, not applicable.
[a]Proportions of 4 risk categories were estimated by multiplying the proportion of clinical risks and their corresponding genetic risk proportions; for example, the proportion for genetic and clinical low risk (58.3%) was estimated by multiplying 69.0% of clinical low risk and 84.5% of low genetic risk when the clinical risk was low. The ranges used in the deterministic sensitivity analysis were estimated following the same strategy.
[b]Applicable to groups with Oncotype DX Recurrence Score ≤ 25, low risk from Prosigna and EndoPredict, and low genetic risk from MammaPrint.
[c]Applicable to Prosigna intermediate-risk group.
[d]Applicable to groups with Oncotype DX Recurrence Score ≥ 26, high risk from Prosigna and EndoPredict, and high genetic risk from MammaPrint.

Natural History and Impact of Treatment

We made two key assumptions related to the natural history of breast cancer and the impact of treatment. First, we assumed that people classified in the same risk category by different GEP tests have different recurrence risks. Second, we assumed that adjuvant chemotherapy has a constant effect on decreasing the recurrence risk.

The variables used to model the natural history of breast cancer and impact of treatment are detailed in Table 13 and include:

- The relative risk of developing distant recurrence for people receiving adjuvant chemotherapy versus those who do not
  - According to a meta-analysis involving over 100,000 women with early-stage breast cancer, the annual rates of developing a distant recurrence were 3.3% for those accepting chemotherapy using any anthracycline-based regimen versus 4.6% for those without chemotherapy.[180] Converting these annual rates to 10-year probabilities, we estimated a 10-year relative risk of 0.76 (28.5% with adjuvant chemotherapy vs. 37.6% without).[118] But a relative risk applies only to a specific follow up time, so we converted risk probabilities beyond the 10-year timeframe (5-year or 9-year probabilities reported) to 10-year probabilities before multiplying these risks by our 10-year relative risk of 0.76
- The probability of distant recurrence for different risk categories without adjuvant chemotherapy (hormone therapy only)
  - For EndoPredict and Prosigna, we used the 10-year probabilities of distant recurrence for each risk category reported in the TransATAC study (Table A17, Appendix 9)[52]
  - For MammaPrint, we used estimates from the MINDACT trial.[51] For three risk groups—clinical and genetic low risk, clinical high and genetic low risk, and clinical low and genetic high risk—we used the 5-year probabilities of distant recurrence-free survival after hormone therapy (Table A17, Appendix 9).[51] For the clinical and genetic high risk group, all patients in MINDACT trial received chemotherapy and there is no other source of information for the hormone therapy estimate, so we estimated the probability of distant recurrence without adjuvant chemotherapy by dividing the 10-year probability of distant recurrence after adjuvant chemotherapy by the relative risk of 0.76 (relative risk for distant recurrence for adjuvant chemotherapy versus no chemotherapy). As previously discussed, we first converted the 5-year probability to a 10-year probability before we used it with the relative risk
  - For Oncotype DX, we used the TAILORx trial, which reported the 9-year probabilities of distant recurrence-free survival for people with a Recurrence Scores of < 10 or 11–25.[26] However, we were unable to use the TAILORx trial for all people with a Recurrence Score of ≥ 26 because this population received chemotherapy. Instead, for those with a Recurrence Score of ≥ 26, we used the TransATAC study, which reported the 10-year probability of distant recurrence for high-risk category (≥ 31). (Table A17 in Appendix 9)
  - For usual care, because the TAILORx trial, the TransATAC study, and the MINDACT trial used different baseline risk levels, we estimated the expected risks of distant recurrence if no chemotherapy were given separately.[26,51,52] Our expected risk was the sum of products of each probability of being classified into

one risk category and its corresponding distant recurrence risk in 10 years if the people have not received adjuvant chemotherapy. Using this method, we were able to create comparisons that best represent the average patient in each study

- We converted all probabilities, reported or estimated, into 1-month probabilities for the analysis (Table 13).[52] In the probabilistic sensitivity analysis, we used the mean and standard errors reported in the respective studies. We used the Beta distribution to simulate the 5-, 9-, and 10-year probabilities described above and converted these into monthly probabilities (see Table A17 in Appendix 9).
- The probabilities of distant recurrence for different risk categories with adjuvant chemotherapy
    - We assumed that adjuvant chemotherapy has a constant effect to decrease the risk of distant recurrence. To obtain the probability of distant recurrence with adjuvant chemotherapy, we multiplied the relative risk of 0.76 by the probabilities of distant recurrence in 10 years without adjuvant chemotherapy.[118,180] We applied this method to all risk categories except the clinical- and genetic high-risk categories reported by MammaPrint
- The probability of death, no distant recurrence
    - We used the life expectancy for Ontario female residents from 2014 to 2016 from Statistics Canada to inform the baseline probability of death (background mortality) for all people in the model[181]
- The probability of death, chemotherapy toxicity-related
    - We applied an additional probability of death to people receiving adjuvant chemotherapy. The probability of dying from chemotherapy was estimated to be 0.35%.[158,182] We applied this probability to our estimate of people who were alive after the first cycle.
- The probability of death, distant recurrence
    - The median length of life expectancy for those with distant recurrence is 28 months, according to data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program in the United States. We converted this life expectancy to a monthly probability of death among patients with distant recurrence (Table 14)

Table 13: Monthly Probabilities of Distant Recurrence

Test	Adjuvant Chemotherapy[a]	Low Risk, % (Range in DSA)	Intermediate Risk, % (Range in DSA)	High Risk, % (Range in DSA)
EndoPredict	No	0.057 (0.034–0.088)	NA	0.208 (0.186–0.425)
	Yes	0.043 (0.026–0.066)		0.153 (0.137–0.302)
MammaPrint	No	Clinical low: 0.041 (0–0.086)	NA	Clinical low: 0.077(0.034–0.175)
		Clinical high: 0.105 (0.034–0.175)		Clinical high: 0.216 (0.114–0.534)
	Yes	Clinical low: 0.031 (0–0.064)	NA	Clinical low: 0.058 (0.026–0.130)
		Clinical high: 0.079 (0.026–0.130)		Clinical high: 0.159 (0.086–0.371)
Oncotype DX	No	Recurrence score ≤ 10: 0.030 (0–0.048)	NA	Recurrence score ≥ 25: 0.264 (0.186–0.425)
		Recurrence score 10–25: 0.052 (0.019–0.098)		
	Yes	Recurrence score ≤ 10: 0.023 (0–0.036)	NA	Recurrence score ≥ 25: 0.193 (0.137–0.302)
		Recurrence score 10–25: 0.040 (0.014–0.073)		
Prosigna	No	0.025 (0–0.043)	0.127 (0.088–0.186)	0.326 (0.186–0.425)
	Yes	0.019 (0–0.032)	0.094 (0.066–0.137)	0.235 (0.137–0.302)

Abbreviations: DSA, deterministic sensitivity analysis; NA, not applicable.
[a]No = hormone therapy only; Yes = hormone therapy and adjuvant chemotherapy.

Table 14: Monthly Probability of Death

Model Parameter	Mean, %	Value in DSA, %	Distribution	Reference
Monthly probability for those without distant recurrence	Age-specific probabilities for female population	—	—	Lifetable by age[183]
Monthly probability for those with distant recurrence[a]	3.5	2.5–4.9	—	Leone et al, 2017[184]
Probability of death due to chemotherapy toxicity	0.35	0.2–0.5	Beta (3, 845)	Ludwig et al, 1989,[182] Paulden et al, 2013[158]

Abbreviation: DSA, deterministic sensitivity analysis.
[a]Estimated based on the median overall survival of 28 months (95% confidence interval: 27–29). The range for deterministic sensitivity analysis was decided by the median survival ranging from 20 months to 40 months. We used gamma distribution for the survival time of patients with distant recurrence.

Health State Utilities

We assumed that GEP tests have no direct impact on quality of life. We considered utilities for the following health states: no distant recurrence and on adjuvant chemotherapy, no distant recurrence in the first year after diagnosis and on hormone therapy, no distant recurrence in the second and subsequent years, distant recurrence, and death (Table 15). We obtained the utility values from Lidgren et al,[185] who reported EQ-5D index values based on 361 consecutive women with breast cancer from the Karolinska University Hospital Solna (Sweden) (Table 15). Utilities were scored on a scale of 0 to 1, with 0 being dead and 1 being perfect health. The utilities were consistent with the previous analyses using the same model. Paulden et al[158] reported that in the first year, adjuvant chemotherapy has a temporary negative impact on utility because of its toxicity. We applied this utility value during the six months in which the patients received adjuvant chemotherapy.

We assumed that deaths from chemotherapy toxicity happened at the beginning of treatment and no QALY was assigned to these people. In our model, they lived for a month with the utility of 0.744 (the first cycle of risk classification and decision making).

Table 15: Utilities Used in the Economic Model

Health State/Treatment State	Utility, mean (SE)[a]	Value in DSA	Reference
No distant recurrence, on chemotherapy	0.620 (0.048)	0.5–0.8	Lidgren et al, 2007,[185] Paulden et al 2013[158]
No distant recurrence in the first year, on hormone therapy	0.744 (0.068)	0.6–0.9	Lidgren et al, 2007,[185] Paulden et al 2013[158]
No distant recurrence in the second and subsequent years	0.779 (0.018)	0.6–0.9	Lidgren et al, 2007,[185] Paulden et al 2013[158]
Distant recurrence	0.685 (0.029)	0.5–0.8	Lidgren et al, 2007,[185] Paulden et al 2013[158]
Dead	0		

Abbreviations: DSA, deterministic sensitivity analysis; SE, standard error of the mean.
[a]Beta distribution.

Cost Parameters

We reported costs in 2018 Canadian dollars. When 2018 costs were not available, we used the Ontario-specific health care component of Consumer Price Index to adjust costs.[186] We converted costs given in USD using the average conversion rate in 2018 (1 USD = 1.2957 CAD).[187]

Resources and costs considered in the reference case analysis include the following:

- Cost of GEP test
- Cost of treatment for
  - People who have no distant recurrence, are in the first year after diagnosis, and are on adjuvant chemotherapy and hormone therapy
  - People who have no distant recurrence, are in the first year after diagnosis, and are on hormone therapy only
  - People who have no distant recurrence and are receiving ongoing care in the second and subsequent years
  - People who have developed distant recurrence

Gene Expression Profiling Test Costs

We obtained the price of Oncotype DX from the Ontario Ministry of Health and from manufacturers. We obtained the prices of three tests, EndoPredict (Kim Slamka, Myriad Genetics, email communication, February 26, 2019), MammaPrint (Vicky Huerta Reyes, Agendia, e-mail communication, March 5, 2019), and Prosigna (Sam Wissa, NanoString, email communication, February 11, 2019), from the manufacturers. In our reference case, we assumed all tests are conducted in Ontario. For Oncotype DX, we assumed that the test price in Ontario would be 90% of the listed market price for the out-of-country program. The manufacturers of Prosigna, EndoPredict, and MammaPrint developed models with which these tests could be conducted in Ontario. We used the assumed price for the decentralized model in Canadian dollars from the manufacturers, rather than market prices. Table 16 summarizes the test costs used in the economic model. There is no test cost for people receiving usual care. We did not consider the setup cost of the GEP tests or labour costs because the estimated test prices in Ontario provided by manufacturers already included these considerations.

Table 16: Test Costs Used in the Economic Model

Test	Unit Cost (CAD)[a]	Value in DSA	Reference
EndoPredict	$2,964[b]	65%, 75%, 90%, 110%, 125%, and 150% of the price in the reference case analysis	EndoPredict[b]
MammaPrint	$3,757[c]		MammaPrint[c]
Oncotype DX	$4,869[d]		Ministry of Health[d]
Prosigna	$2,576[e]		Prosigna[e]

Abbreviation: DSA, deterministic sensitivity analysis.

[a]In 2018, 1 USD = 1.2957 CAD.[187]

[b]Based on cost including consumable kits and labour, plus commitment of 1,250 tests per year with associated potential rebates; not the actual price in Ontario (EndoPredict, Kim Slamka, Myriad Genetics, email communication, February 26, 2019).

[c]The price was $2,900 USD for the decentralized model for MammaPrint (MammaPrint, Vicky Huerta Reyes, Agendia, email communication, March 5, 2019).

[d]We assumed that the cost of test in Ontario would be 90% of the price paid through the out-of-country program. The price was $4,175 USD in the centralized model for Oncotype DX (Ministry of Health, e-mail communication, January 14, 2019).

[e]Based on cost including consumable kits and labour, and maintenance cost for 400 tests annually; not the actual price in Ontario (Prosigna, Sam Wissa, NanoString, presentation, February 11, 2019).

Treatment- and Care-Related Costs

Table 17 summarizes treatment- and care-related costs. We used three main sources:[126,158,188]

- For the first 20 months after breast cancer diagnosis, we used the Ontario costing analysis by Mittmann et al,[126] to estimate the costs for people who are recurrence-free. This analysis compared the cost and resource use among 998 people who had Oncotype DX testing and received either adjuvant chemotherapy and hormone therapy or hormone therapy alone. The analysis considered all relevant direct costs, including inpatient cost, emergency department and cancer clinic visits, rehabilitation, complex continuing care, home care, physician billing, chemotherapy medication, hormone therapy medication, and supportive drugs. The total cost per person receiving chemotherapy was $39,322.[126] The total cost per person receiving hormone therapy alone was $23,030
- To estimate the costs for ongoing care in people who have no distant recurrence in months 21 to 24 after diagnosis, we used an Ontario-based cost analysis on cancer survivors who were transitioned to primary care[188]

- To estimate the costs for ongoing care in people who have no distant recurrence after 2 years, and the costs for people who have a distant recurrence, including initial treatment, we used the ongoing care and end-of-life care costs from a previous Ontario analysis[158]

Table 17: Treatment Costs Used in the Economic Model

Variable	Mean Unit Cost	Value in DSA, %	Reference
Monthly Treatment Cost, Adjuvant Chemotherapy and Hormone Therapy, First 2 Years After Diagnosis[a]			
On chemotherapy	$3,614.68[b,c]	± 50%	Mittmann et al, 2018[126]
When chemotherapy is completed in the first year after diagnosis	$2,780.05[b,c]	± 50%	Mittmann et al, 2018[126]
Months 13 to 20	$422.90[b,d]	± 50%	Mittmann et al, 2018[126]
Months 21 to 24	$271.46[c,e]	± 50%	Mittmann et al, 2018[188]
Monthly Treatment Cost, Hormone Therapy Only, First 2 Years After Diagnosis[a]			
Months 1 to 12	$1,798.11[c,f]	± 50%	Mittmann et al, 2018[126]
Months 13 to 20	$274.71[c,f]	± 50%	Mittmann et al, 2018[126]
Months 21 to 24	$271.46[c,e]	± 50%	Mittmann et al, 2018[188]
Monthly hormone therapy	$11.97[c,g]	± 50%	Mittmann et al, 2018[126]
Monthly Cost (Excluding Hormone Therapy Medication Cost), Third and Subsequent Years After Diagnosis, Without Distant Recurrence[a]			
Year 3	$48.48[h]	—	Will et al, 2000[189] Paulden et al, 2013[158]
Year 4	$42.65[h]	—	
Year 5 and beyond	$36.83[h]	—	
Cost, Distant Recurrence[a]			
Initial treatment (1 time)	$8,374.08[h]	± 50%	Will et al, 2000[189] Paulden et al, 2013[158]
Ongoing treatment (per month)	$719.10[h]	± 50%	
End-of-life care (1 time, over last 3 months)	$22,088.05[h]	± 50%	
Cost, Chemotherapy-Related Death (Gamma Distribution)			
Death related to chemotherapy toxicity	$36,112.57 ($2,413.95)[i]	± 50%	Paulden et al, 2013[158]

Abbreviation: DSA, deterministic sensitivity analysis.

[a]Costs were estimated based on Mittmann et al, 2018 (Appendix 9, Table A18).[126] For probabilistic sensitivity analysis, we estimated the parameters based on the gamma distributions summarized in Appendix 9, Table A21.

[b]For people in the first 20 months following diagnosis, without distant recurrence, the treatment cost (standard deviation) was $39,322 ($17,099) for those who received adjuvant chemotherapy.

[c]Converted from cost in 2014; $1 in 2014 CAD = 1.0735 in 2018 CAD.[186]

[d]Divided into 1 to 8 months and 9 to 12 months in the second year following diagnosis, because the source provided costs up to the first 8 months of the second year.

[e]Based on the cost estimate of $6,575 for 26 months of treatment for cancer survivors in primary care.[188]

[f]For people in the first 20 months following diagnosis, without distant recurrence, the treatment cost (standard deviation) was $23,030 ($11,951) for those who received hormone therapy only.

[g]Based on the cost estimate of $223 for 20 months for those receiving hormone therapy only.

[h]Cost for ongoing care is based on a previous economic evaluation[158] and converted from costs from 1995 CAD; $1 in 1995 CAD = $1.4226 in 2018 CAD.[158,186]

[i]Mean (standard error) of cost for death related to chemotherapy toxicity is based on a previous economic evaluation[158]; converted from costs from 2010: $1 in 2010 CAD = $1.1125 in 2018 CAD.[158,186]

Costs for People in the First Year After Diagnosis, Without Distant Recurrence

Twenty-month costs were obtained from the Ontario costing study.[126] To arrive at first-year estimates, we converted the Ontario costing numbers to monthly costs, assuming that:

- 100% of the chemotherapy and supportive drugs will be incurred in the first 6 months of chemotherapy treatment
- 90% of the inpatient cost, emergency department visit, cancer clinic, rehabilitation, complex continuing care, home care, and physician billing will be incurred in the first year after diagnosis and is evenly distributed over the year (Sunnybrook Health Sciences Centre, Andrea Eisen and Maureen Trudeau, email communication, March 14, 2019). The remaining 10% will be incurred in the first 8 months of the second year and is also evenly distributed
- Hormone therapy costs will be evenly distributed over 20 months

For people receiving adjuvant chemotherapy and hormone therapy, the monthly cost excluding medication cost was estimated to be $2,768.08. The monthly medication costs were $559.81, $286.79, and $11.97 for adjuvant chemotherapy, supportive drugs (given during course of chemotherapy), and hormone therapy, respectively. Treatment was sequential—hormone therapy did not start until after adjuvant chemotherapy completed. As such, the monthly costs in total for people receiving adjuvant chemotherapy and hormone therapy were $3,614.68 while the patients are receiving adjuvant chemotherapy, and $2,780.05 while they receive hormone therapy. For people receiving hormone therapy alone, the monthly cost excluding medication was estimated to be $1,798.11. Including medication cost, the monthly total would be $1,810.08 (Table A18, Appendix 9).

Cost for People in the Second Year After Diagnosis, Without Distant Recurrence

For the first 8 months of the second year, we used monthly costs derived from Mittmann and colleagues[188] as described above and in Table A18 (Appendix 9). The monthly cost for those who receive adjuvant chemotherapy in combination with hormone therapy and those who receive hormone therapy alone was $422.90 and $274.71, respectively.

For the last four months of the second year, we estimated a monthly cost of $271.46 using an Ontario-based cohort study of breast cancer survivors.[188] In this analysis, the total cost was $7,058 for 26 months per person when patients finished initial treatment and were transferred from oncologist-led care to primary care. Assuming the ongoing care would be evenly distributed, the monthly cost, excluding hormone therapy, was $271.46 after adjustment for inflation. We continued to cost hormone therapy at $11.97 per month in the second year after diagnosis.

Cost for People in the Third and Subsequent Years After Diagnosis, Without Distant Recurrence

For subsequent years, the cost included two components: hormone therapy and ongoing care. We assumed that the monthly cost of hormone therapy was constant but that the duration differed. We continued to use $11.97 as the cost per month for hormone therapy, was applied in the third to fifth, third to tenth, third to tenth, and third to seventh years in people who have no distant recurrence and are classified as low, intermediate, high, and unclear risk, respectively, based on usual care.[126] We added this to the monthly cost of ongoing care for people who have no distant recurrence, as reported by Paulden et al in 2013.[158]

Cost for Distant Recurrence

For people with distant recurrence, we included initial treatment costs (considering hospital- and outpatient clinic–related costs), ongoing treatment costs, and end-of-life care costs (Table 17). Consistent with Paulden et al,[158] we used estimates of direct health care costs in the lifetime management of 17,700 women with breast cancer.[189] We assumed there would be significant costs incurred at the end of life. Thus, for people who die after a distant recurrence, we added a 3-month end-of-life care cost. We did not include the cost of ongoing treatment for distant recurrence during this 3-month period.

Cost for Chemotherapy Toxicity–Related Death

We assumed a one time cost of $36,112.57 (standard error: $2,413.95) for chemotherapy toxicity–related death, as reported by Paulden et al in 2013.[158]

Analysis

Internal Validation

Formal internal validation was conducted by the secondary health economist. This included testing the mathematical logic of the model and checking for errors and accuracy of parameter inputs and equations.[177]

Reference Case Analysis

We conducted all analyses in Microsoft Excel. We calculated the mean costs and QALYs for each intervention assessed. We also calculated the mean incremental costs, incremental QALYs, and incremental cost-effectiveness ratios (ICERs) for each GEP test versus usual care. The ICER represents the incremental cost for each additional QALY gained.

We calculated the reference case of this analysis by running 10,000 simulations (probabilistic analysis) that simultaneously captured the uncertainty in all parameters that were expected to vary. We set distributions for variables within the model. The model variables and the corresponding distributions are listed in Appendix 9.

Sensitivity Analysis

The results of the probabilistic sensitivity analysis are presented on a cost-effectiveness acceptability curve. We assessed variability and uncertainty in the model using one-way sensitivity analyses. We varied specific model variables and examined the impact on the results. Tables 12–17 present the variables and ranges. We used a tornado diagram to plot the results of the one-way sensitivity analyses. In addition, we assessed the cost-effectiveness of several subgroups and scenarios. For these subgroups and scenarios, we also conducted probabilistic sensitivity analyses by running 10,000 simulations. These are described briefly below and in Table A20 (Appendix 9).

Scenario: Comparison Between GEP Tests

We were unable to conduct a multiway comparison between tests because of the different baseline risk levels among studies that we used to populate the model (TransATAC for EndoPredict and Prosigna, MINDACT for MammaPrint, and TAILORx for Oncotype DX). In our reference case analysis, the expected risk of 10-year distant recurrence was 11.1%, 8.9%, and

8.9% in TransATAC, TAILORx, and the subgroup of people with hormone receptor–positive, LN−, HER2− breast cancer in MINDACT, respectively.[26,51,52] In a scenario analysis, we conducted two 2-way comparisons. First, we compared EndoPredict and Prosigna, as our inputs for these tests were both drawn from the TransATAC study. Second, we compared MammaPrint and OncotypeDX, as the baseline risk levels in the MINDACT and TAILORx studies were similar. Additional head-to-head data are required to compare all tests against one another.

Scenario: Triage Test Before GEP Tests

In the reference case analysis, we assumed that no triage test was used. In this scenario analysis, we explored the cost-effectiveness of GEP tests when there were available clinical tools to classify people into clinical risk groups. We assessed only Oncotype DX and MammaPrint in this scenario analysis because of data availability. The TAILORx and MINDACT trials compared these GEP tests with a modified version of AOL.[26,51] We chose this modified AOL for several reasons—it was applied in both trials, it is simple to use in practice (considering ER status, HER2 status, tumour grade, nodal status, and tumour size), and the unmodified version of AOL is no longer available.

In this analysis, people receiving usual care received risk classification results from modified AOL and we assessed the cost-effectiveness of GEP tests when the tests were conducted for clinically low-risk patients only, clinically high-risk patients only, and all patients. We assumed that 15% of clinically low-risk and 70% of clinically high-risk patients would accept adjuvant chemotherapy if they had not received a GEP test, which was consistent with the NICE diagnostics guidelines.[118]

Subgroups Differing From Our Reference Case Target Population

*Premenopausal Population***.** We assessed the cost-effectiveness of GEP tests for premenopausal people (50 years of age). This subgroup differed from our reference case analysis in three ways:

- Background mortality is lower as this cohort is younger
- Hormone therapy follows a different protocol (tamoxifen alone for 10 years)[142]
- Recurrence scores for Oncotype DX testing is categorized differently for women who are ≤ 50 years of age compared with women who are > 50 years (Recurrence Score of 0–25 vs. 26–100, respectively). Depending on the potential benefit of chemotherapy, they would be further subdivided into groups with Recurrence Scores of 0–15, 16–20, 21–25, and 26–100. We used subgroup data from the TAILORx trial[26] to conduct this analysis. The risk classification and parameters related to risk classification for the probability of distant recurrence for other GEP tests were the same as for the reference case

*People With LN+ Breast Cancer***.** We explored the cost-effectiveness of GEP tests in people with LN+ breast cancer using a different set of transition probabilities. Notably, TAILORx did not include people with LN+ breast cancer. Thus, we used the LN+ subgroup from the TransATAC study to populate the model for EndoPredict, Oncotype DX, and Prosigna.[52] In addition, we used an Ontario-based study on the decision impact of Oncotype DX specifically for women with LN+ breast cancer to model the impact of GEP tests on adjuvant chemotherapy decision.[103] In general, even where the same GEP test classifies patients at the same risk level, physicians are more likely to provide adjuvant chemotherapy to people with LN+ than to people with LN− breast cancer.[84,103]

Scenarios: Gene Expression Profiling Test Use

The “Three Risk Categories Strategy” for Oncotype DX. People receiving Oncotype DX are divided into three risk groups (low, intermediate, and high). We used the risk classification information from the TransATAC study.[52] This scenario analysis compared Oncotype DX with EndoPredict and Prosigna, and usual care at a similar risk level, though its generalizability to the current Ontario practice may be limited.

Status Quo Comparison. We compared the GEP test strategies with a status quo strategy where people with breast cancer receive Oncotype DX through the out-of-country program.

Oncotype DX Varied Risk Classification. In our reference case analysis, we populated the model with data from the TAILORx trial.[84] In our scenario analysis, we used Ontario evidence to estimate the proportion of people who are classified in each of three risk categories (low, intermediate, and high based on recurrences score of <18, 18–30, and > 30, respectively).[84]

Various Rates of Chemotherapy Acceptance. In our reference case analysis, 4.1%, 32.9%, and 79.3% of low-, intermediate-, and high-risk people would accept chemotherapy, respectively, regardless of the GEP tests they receive. We considered two scenarios of chemotherapy acceptance:

- No low-risk people receive chemotherapy and all high-risk people receive chemotherapy
- Based on another evaluation,[118] 0%, 17%, and 74% of low-, intermediate-, and high-risk people receive chemotherapy, respectively, in the three-category test, and 7% and 77% of low- and high-risk people receive chemotherapy, respectively in the two-category test

Scenarios: Model Structure

Local Recurrence. We included local recurrence and assumed that 10.5% of people who have distant recurrence would experience local recurrence.[118,190] We did not consider the time spent with local recurrence; rather, we applied a one-time cost and disutility in the model to account for local recurrence.

The Ability to Predict Chemotherapy Benefit. In the reference case analysis, we assumed the test influences the clinical outcome by changing the probability of receiving adjuvant chemotherapy. There is evidence suggesting that some tests may be beneficial to people with breast cancer by predicting the response to adjuvant chemotherapy.[26,56,90] That is, people classified as high risk are more likely to receive chemotherapy and, therefore, have a greater relative risk reduction for distant recurrence. In this scenario, we considered the additional benefit of GEP testing by applying a different relative risk of distant recurrence for adjuvant chemotherapy versus no chemotherapy across different risk levels classified by Oncotype DX. We also considered the predictive benefit for EndoPredict.[102] We did not consider this scenario for MammaPrint because the information was available only for the subgroup of clinically high-risk and MammaPrint low-risk group of the MINDACT study population, and not for the patient group with early-stage invasive, hormone receptor–positive, HER2− breast cancer.

Scenarios: Additional

Different Chemotherapy Regimens for Different Risk Levels. In the scenario analysis, we assessed the impact of risk-dependent adjuvant chemotherapy regimens on cost-effectiveness. People who were classified as high risk received six chemotherapy cycles consisting of fluorouracil, epirubicin and cyclophosphamide (FEC) on day 1 of the first 3 cycles, and docetaxel on day 1 of the last 3 cycles. The treatment duration would be 6 months. People who were classified as low or intermediate risk received four chemotherapy cycles consisting of docetaxel and cyclophosphamide TC given on day 1 of the cycle. All people receiving chemotherapy also receive granulocyte colony–stimulating factor for 8 days every cycle.[30,191] For simplicity, we evenly distributed the cost of the four cycles of TC chemotherapy across 6 months of the first year after a breast cancer diagnosis. See Table A20 (Appendix 9) for a summary of all variables in the scenario analysis.

Results

We present the results of our primary economic evaluation with reference case analysis and sensitivity analyses.

Reference Case Analysis

All test strategies were more effective (led to more QALYs) than usual care (Tables 18–21). Prosigna was more effective and less costly than usual care (dominant). Oncotype DX, EndoPredict, and MammaPrint were more effective and more costly than usual care. The ICERs ranged from $1,490 to $19,793 per QALY. Due to variation in the baseline risks in the studies informing these analyses, the results for each test should not be compared with one another. Table A21 (Appendix 10) summarizes additional outcomes and cost breakdowns of our reference case analysis. Based on the cost breakdowns, the low incremental cost was attributable to the avoided adjuvant chemotherapy, which partially offset the increased cost of GEP testing.

Table 18: Reference Case Analysis Results, EndoPredict Versus Usual Care

Strategy	Mean Total Cost, $ (95% CrI)	Mean Incremental Cost, $ (95% CrI)	Mean Life Years (95% CrI)	Mean Life Years Gained (95% CrI)	Mean QALYs (95% CrI)	Mean Incremental QALYs (95% CrI)	ICER, $/QALY
Usual care[a]	46,960 (44,854–49,326)	—	19.81 (19.17–20.40)	—	15.33 (14.50–16.13)	—	—
EndoPredict	47,144 (45,024–49,474)	183 (−2,501 to 2,755)	19.95 (19.32–20.52)	0.14 (−0.69 to –0.99)	15.45 (14.62–16.26)	0.12 (−0.55 to 0.79)	1,490

Abbreviations: CrI, credible interval; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.
[a]Reference group. Usual care for EndoPredict represents the cost and outcomes for the study population in the TransATAC study,[52] if they had not received EndoPredict.

Table 19: Reference Case Analysis Results, MammaPrint Versus Usual Care

Strategy	Mean Total Cost, $ (95% CrI)	Mean Incremental Cost, $ (95% CrI)	Mean Life Years (95% CrI)	Mean Life Years Gained (95% CrI)	Mean QALYs (95% CrI)	Mean Incremental QALYs (95% CrI)	ICER, $/QALY
Usual care[a]	45,590 (44,025–47,315)	—	20.28 (19.90–20.62)	—	15.70 (14.97–16.43)	—	—
MammaPrint	46,494 (44,995–48,087)	905 (161–1,590)	20.32 (19.96–20.65)	0.05 (−0.00 to 0.11)	15.75 (15.02–16.48)	0.05 (0.00–0.10)	19,793

Abbreviations: CrI, credible interval; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.
[a]Reference group. Usual care for MammaPrint represents the cost and outcomes for the study population in the MINDACT trial,[51] if they had not received MammaPrint.

Table 20: Reference Case Analysis Results, Oncotype DX Versus Usual Care

Strategy	Mean Total Cost, $ (95% CrI)	Mean Incremental Cost, $ (95% CrI)	Mean Life Years (95% CrI)	Mean Life Years Gained (95% CrI)	Mean QALYs (95% CrI)	Mean Incremental QALYs (95% CrI)	ICER, $/QALY
Usual care[a]	45,557 (43,806–47,437)	—	20.29 (19.82–20.72)	—	15.71 (14.93–16.47)	—	—
Oncotype DX	46,243 (44,731–47,813)	686 (−134 to 1,446)	20.35 (19.97–20.72)	0.07 (−0.02 to 0.19)	15.78 (15.01–16.52)	0.07 (−0.01 to 0.16)	10,383

Abbreviations: CrI, credible interval; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.
[a]Reference group. Usual care for Oncotype DX represents the cost and outcomes for the study population in the TAILORx trial,[26] if they had not received Oncotype DX.

Table 21: Reference Case Analysis Results, Prosigna Versus Usual Care

Strategy	Mean Total Cost, $ (95% CrI)	Mean Incremental Cost, $ (95% CrI)	Mean Life Years (95% CrI)	Mean Life Years Gained (95% CrI)	Mean QALYs (95% CrI)	Mean Incremental QALYs (95% CrI)	ICER, $/QALY
Usual care[a]	46,960 (44,854–49,326)	—	19.81 (19.17–20.40)	—	15.33 (14.50–16.13)	—	—
Prosigna	46,630 (44,522–48,874)	−331 (−1,209 to 463)	19.99 (19.40–20.54)	0.18 (0.09–0.31)	15.48 (14.67–16.28)	0.15 (0.08–0.26)	Dominant[b]

Abbreviations: CrI, credible interval; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.

[a]Reference group. Usual care for Prosigna represents the cost and outcomes for the study population in the TransATAC study,[52] if they had not received Prosigna.

[b]Less costly and more effective than usual care.

Sensitivity Analysis

Probabilistic Sensitivity Analyses

In the reference case analysis, at a willingness-to-pay of $50,000 per QALY gained, the probability of each test being cost-effective compared to usual care was 63.0%, 89.2%, 89.2%, and 100% for EndoPredict, MammaPrint, Oncotype DX, and Prosigna, respectively (Table 22; see also Figures A1 and A2 in Appendix 10).

Table 22: Probability of Tests Being Cost-Effective Versus Usual Care[a] Under Different Willingness-to-Pay Values

Strategy[a]	Willingness-to-Pay Values, $/QALY, CAD 10,000	20,000	50,000	100,000
EndoPredict	0.590	0.613	0.630	0.636
MammaPrint	0.177	0.537	0.892	0.957
Oncotype DX	0.487	0.732	0.892	0.934
Prosigna	0.998	0.999	1	1

Abbreviation: QALY, quality-adjusted life-year.
[a]Usual care varies for each comparison.

Deterministic Sensitivity Analyses

We conducted deterministic sensitivity analyses to assess the robustness of our results. Our analyses suggested that the results were relatively robust when subject to variation in discount rates, age, and utilities.

The variables most influential to the ICER across all four analyses (each test compared with usual care) included:

- The proportion of patients receiving chemotherapy in the usual care group
- The relative risk for high-risk people of developing a distant recurrence for those receiving chemotherapy versus those not receiving chemotherapy
- The monthly chemotherapy cost for those with high risk
- The monthly chemotherapy cost for those receiving usual care
- The monthly treatment cost for high-risk people whose adjuvant chemotherapy has completed and are in their first year after diagnosis
- The monthly cost for those receiving hormone therapy only in the first year after diagnosis. We elaborate on a few parameters below.

Proportion of Patients Who Received Adjuvant Chemotherapy

The estimates of ICER for GEP tests compared to usual care were robust to the proportions of low-, intermediate-, and high-risk people receiving adjuvant chemotherapy. The ICERs for all GEP tests remained less than $50,000 per QALY in the one-way sensitivity analyses. However, the proportion of patients who received adjuvant chemotherapy in usual care was an influential factor for the cost-effectiveness of all four GEP tests. In the reference case, we assumed that 38% of people in the usual care group would receive adjuvant chemotherapy. As this proportion increased, the costs and QALYs increased for the usual care strategy. Table 23 shows the results when 60% of patients received adjuvant chemotherapy in the usual care strategy. In this analysis, all GEP tests were less costly than usual care. Oncotype DX and MammaPrint were less effective than usual care. EndoPredict and Prosigna were more effective and, therefore, dominant compared with usual care. When 80% of patients received adjuvant chemotherapy in the usual care strategy, all GEP tests were less costly and less effective compared to usual care.

Table 23: One-Way Sensitivity Analysis Results: 60% of Patients Accept Adjuvant Chemotherapy Without Test

Strategy	Average Total Cost, $	Incremental Cost, $	Average Total QALYs	Incremental QALYs	ICER ($/QALY)
Usual care[a,b]	50,227	—	15.43	—	—
EndoPredict	47,144	−3,083	15.45	0.02	Dominant[c]
Usual care[a,d]	48,926	—	15.78	—	—
MammaPrint	46,494	−2,432	15.75	−0.03	72,244[e]
Usual care[a,f]	48,895	—	15.79	—	—
Oncotype DX	46,243	−2,652	15.78	−0.01	208,787[e]
Usual care[a,b]	50,227	—	15.43	—	—
Prosigna	46,630	−3,597	15.48	0.05	Dominant[c]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.
[a]Reference group; usual care varies for each comparison.
[b]Usual care for EndoPredict and Prosigna represents the cost and outcomes for the study population in the TransATAC study,[52] if they had not received EndoPredict or Prosigna.
[c]Less costly and more effective than usual care.
[d]Usual care for MammaPrint represents the cost and outcomes for the study population in the MINDACT trial,[51] if they had not received MammaPrint.
[e]Usual care was more effective and more costly than the gene expression profiling test. The ICER should be interpreted as the incremental cost per QALY gained for usual care compared with the gene expression profiling test.
[f]Usual care for Oncotype DX represents the cost and outcomes for the study population in the TAILORx trial,[26] if they had not received Oncotype DX.

Other Variables Influential to Cost-Effectiveness of All GEP Tests

Figure A3 (Appendix 10) shows the impact of other variables. When the adjuvant chemotherapy cost or hormone cost increases, the ICERs of GEP tests compared to usual care would increase, but would still be under $50,000 per QALY at a cost increase of 25% (from $3,614 to $4,518 per month). The ICERs of GEP tests compared to usual care would decrease as the chemotherapy cost in usual care increases, and all GEP tests would be dominant compared to usual care if the chemotherapy cost in usual care increases by 25% (from $3,614 monthly to $4,518).

Individual Test Prices

Individual test prices influenced the results. For Oncotype DX, EndoPredict, and MammaPrint, lower test prices could make them dominant compared with usual care. A decrease by 20% for Oncotype DX (from $4,869, using our estimate that Ontario price would be 90% of the current market price, to $3,895), a 10% decrease for EndoPredict (from $2,964 to $2,667) and a 25% decrease for MammaPrint (from $3,758 to $$2,818), made them dominant strategies. Prosigna, which was dominant in the reference case analysis, became more costly than usual care when the cost of the test increased by 20%, from $2,576 (reference case) to $3,091. The ICER was $1,195 per QALY gained.

Scenario Analyses

Comparison Between Tests

We were able to make two head-to-head comparisons in the study population with the same or similar baseline risks. But results from both comparisons were with great uncertainty. The clinical parameters used in the analysis for EndoPredict and Prosigna came from the same study.[52] When compared with EndoPredict, Prosigna had a higher probability of being cost-effective across all willingness-to-pay values. The probability of EndoPredict being cost-effective increased slightly as the willingness-to-pay increased. When the willingness-to-pay is $50,000 per QALY gained, the probability of being cost-effective was 54.2% for Prosigna and 45.8% for EndoPredict (see Figure 5).

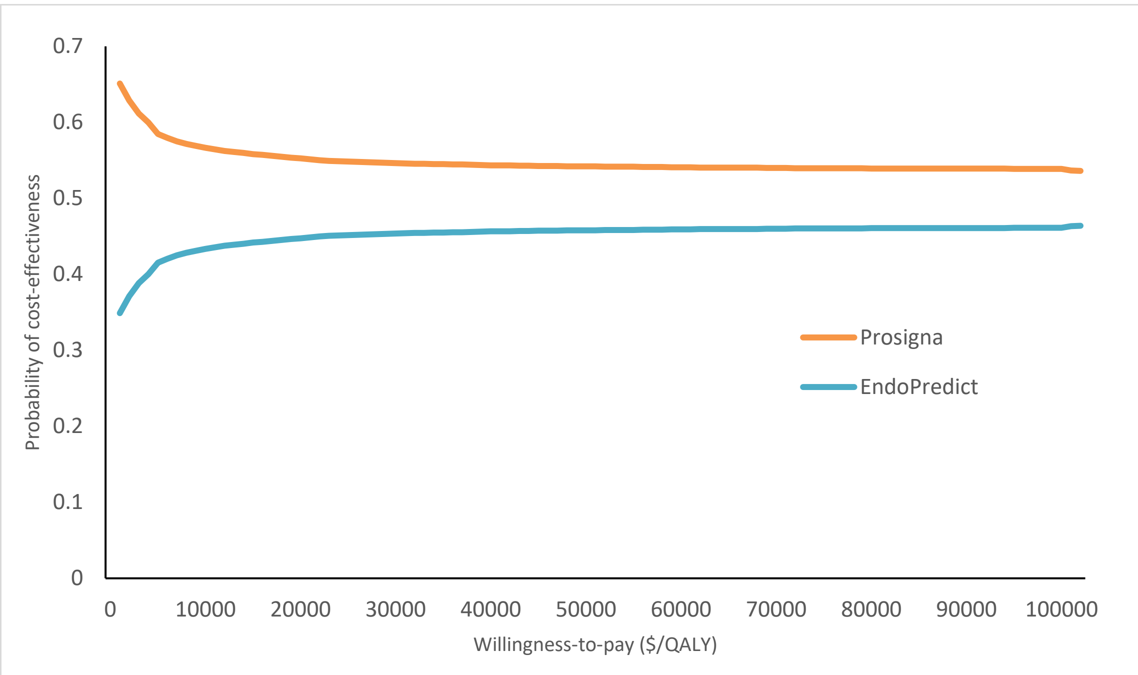


Figure 5: Cost-Effectiveness Acceptability Curve of Prosigna Versus EndoPredict

Abbreviation: QALY, quality-adjusted life year.

Because the TAILORx trial and the subgroup for hormone receptor–positive, LN−, and HER2− patients in the MINDACT trial had similar risk levels, it is feasible to compare Oncotype DX with MammaPrint.[26,51] When compared with MammaPrint, Oncotype DX had higher probabilities of being cost-effective across all willingness-to-pay values. The probability of MammaPrint being cost-effective increased slightly and then plateaued as the willingness-to-pay increased. With a willingness-to-pay of $50,000, the probability of being cost-effective was 56.2% for Oncotype DX and 43.8% for MammaPrint (see Figure 6).

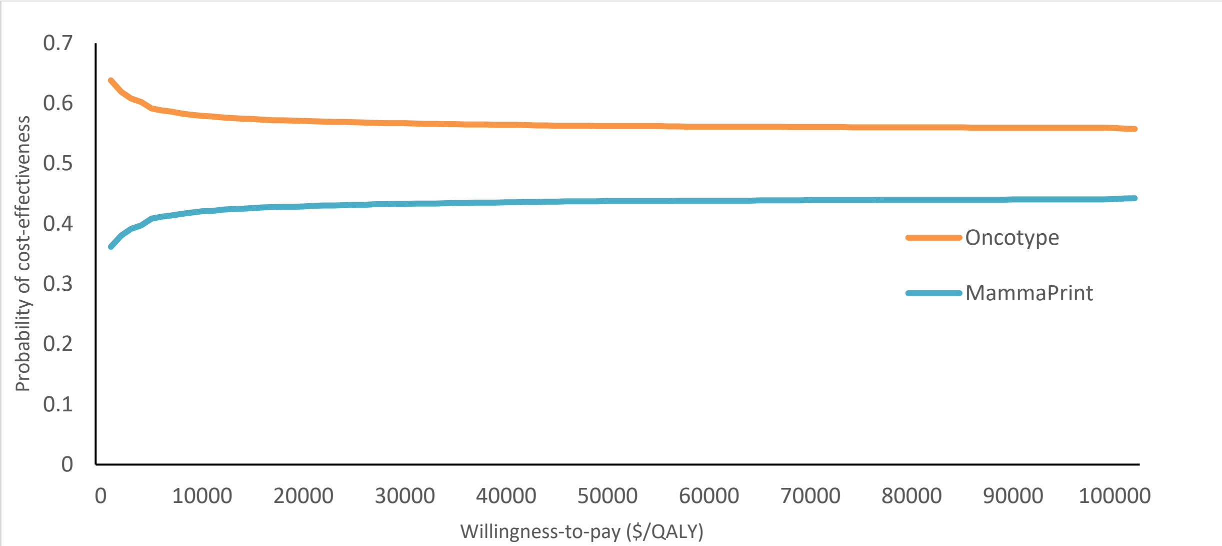


Figure 6: Cost-Effectiveness Acceptability Curve of Oncotype DX Versus MammaPrint

Abbreviation: QALY, quality-adjusted life year.

Triage Test for GEP Tests

When the GEP tests were compared with the modified AOL, MammaPrint was dominated and the ICER for Oncotype DX compared to usual care with the modified AOL strategy was $29,831 per QALY gained.

We then assessed the cost-effectiveness of Oncotype DX or MammaPrint versus no test for the two clinical risk groups defined by modified AOL. Considering low-risk people only, MammaPrint gained 0.02 QALYs and increased the cost by $3,635, for an ICER of $150,770 per QALY gained compared to usual care without MammaPrint. Meanwhile, Oncotype DX was more costly ($3,624 cost increase) and more effective (0.08 QALYs gained). The ICER for Oncotype DX compared to usual care without Oncotype DX was $44,960 per QALY gained.

Considering only high-risk people, MammaPrint was less costly ($1,110 saved) and less effective (0.12 QALYs lost) than usual care with no MammaPrint, and the ICER for no MammaPrint compared to MammaPrint was $9,170 per QALY gained. Oncotype DX was less costly ($2,831 saved) and less effective (0.004 QALYs lost) than usual care with no Oncotype DX, and the ICER for no Oncotype DX compared to Oncotype DX was $737,414 per QALY gained.

Subgroups Differing From Our Reference Case Target Population

Premenopausal Population

For all GEP tests, the ICERs were lower for premenopausal population compared with the reference case analysis ($917, $12,811, and $1,948 per QALY gained for EndoPredict, MammaPrint, Oncotype DX, respectively, when each was compared with usual care. Prosigna was still dominant; Table 24).

Table 24: Cost-Effectiveness of GEP Tests for Premenopausal Population

Strategy	Mean Total Cost, $ (95% CrI)	Mean Incremental Cost, $ (95% CrI)	Mean QALYs (95% CrI)	Mean Incremental QALYs (95% CrI)	ICER, $/QALY
Usual care[a,b]	50,495 (48,083–53,174)	—	17.93 (16.88–18.97)	—	—
EndoPredict	50,663 (48,246–53,346)	167 (−2,625 to 2,963)	18.12 (17.08–19.12)	0.18 (−0.78–1.12)	917
Usual care[a,c]	48,967 (47,166–50,909)	—	18.47 (17.55–19.35)	—	—
MammaPrint	49,896 (48,159–51,703)	928 (177–1,602)	18.54 (17.64–19.43)	0.07 (0.01–0.16)	12,811
Usual care[a,d]	49,288 (47,355–51,397)	—	18.36 (17.39–19.27)	—	—
Oncotype DX	49,924 (48,412–51,533)	636 (−658 to 1,789)	18.68 (17.79–19.53)	0.33 (0.00–0.70)	1,948
Usual care[a,b]	50,495 (48,083–53,174)	—	17.93 (16.88–18.97)	—	—
Prosigna	49,910 (47,607–52,421)	−585 (−1,527 to 221)	18.18 (17.15–19.21)	0.25 (0.12–0.42)	Dominant[e]

Abbreviations: CrI, credible interval; GEP, gene expression profiling; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.
[a]Reference group; usual care varies for each comparison.
[b]Usual care for EndoPredict and Prosigna represents the cost and outcomes for the study population in the TransATAC study,[52] if they had not received EndoPredict or Prosigna.
[c]Usual care for MammaPrint represents the cost and outcomes for the study population in the MINDACT trial,[51] if they had not received MammaPrint.
[d]Usual care for Oncotype DX represents the cost and outcomes for the study population in the TAILORx trial,[26] if they had not received Oncotype DX.
[e]Less costly and more effective than usual care.

People with LN+ Breast Cancer

In people with LN+ breast cancer, most of the tests remained likely to be cost-effective compared with usual care (Table 25). However, in this scenario, Oncotype DX, with three risk categories (low, intermediate, and high), was dominated by (more costly and less effective) usual care. MammaPrint was dominant compared with usual care.

Table 25: Cost-Effectiveness of GEP Tests for People With Lymph-Node–Positive Breast Cancer

Strategy	Mean Total Cost, $ (95% CrI)	Mean Incremental Cost, $ (95% CrI)	Mean QALYs (95% CrI)	Mean Incremental QALYs (95% CrI)	ICER, $/QALY
Usual care[a,b]	60,246 (56,348–64,414)	—	13.48 (12.38–14.58)	—	—
EndoPredict	62,839 (58,761–67,032)	2,593 (−1,994 to 7,151)	13.83 (12.73–14.91)	0.34 (−0.93 to 1.63)	7,520
Usual care[a,c]	54,261 (51,798–56,774)	—	15.19 (14.40–15.96)	—	—
MammaPrint	53,734 (51,008–56,593)	-527 (−2,319 to 1,310)	15.20 (14.38–16.00)	0.01 (−0.08 to 0.09)	Dominant[d]
Usual care[a,e]	60,246 (56,348–64,414)	—	13.48 (12.38–14.58)	—	—
Oncotype DX	60,761(56,532–65,238)	515 (−4,355 to 5,275)	13.40 (12.23–14.53)	−0.08 (−1.43 to 1.25)	Dominated[f]
Usual care[a,b]	60,246 (56,348–64,414)	—	13.48 (12.38–14.58)	—	—
Prosigna	63,766 (59,737–67,868)	3,520 (1,969–4,992)	13.68 (12.62–14.75)	0.19 (0.11–0.30)	18,331

Abbreviations: CrI, credible interval; GEP, gene expression profiling; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.
[a]Reference group; usual care varies for each comparison.
[b]Usual care for EndoPredict and Prosigna represents the cost and outcomes for the study population in the TransATAC study,[52] if they had not received EndoPredict or Prosigna.
[c]Usual care for MammaPrint represents the cost and outcomes for the study population in the MINDACT trial,[51] if they had not received MammaPrint.
[d]Usual care for Oncotype DX represents the cost and outcomes for the study population in the TAILORx trial,[26] if they had not received Oncotype DX.
[f] Oncotype DX was dominated by usual care—it was more costly and less effective than usual care.

Other Scenarios

Performing the Oncotype DX test in Ontario led to a cost-saving of $541 per person over the status quo of sending samples out-of-country for testing. Using Oncotype DX with three risk categories (low-, intermediate-, and high-risk results) cost $1,478 more than usual care and led to 0.06 QALYs gained. In this scenario, the ICER of Oncotype DX compared to usual care was $26,460 per QALY. This scenario analysis used parameters from the TransATAC study, so the cost and QALY results were estimated based on the same clinical source with the results for EndoPredict and Prosigna in the reference case analysis. It suggested that, compared with EndoPredict and Prosigna, Oncotype DX, was more costly (Oncotype DX: $48,438; EndoPredict: $47,144; Prosigna: $46,630), but less effective (Oncotype DX: 15.39 QALYs; EndoPredict: 15.45 QALYs; Prosigna: 15.48 QALYs).

The scenario of Oncotype DX risk classification based on published Ontario data, but also interpreted Oncotype DX with the three risk categories. Oncotype DX was dominated by usual care, with an incremental cost of $2,033 and a 0.04 QALY loss.

In scenarios including various rates of chemotherapy treatment, local recurrence, predictability of chemotherapy benefit, and risk-dependent chemotherapy regimens, the gene expression profiling tests remained likely to be cost-effective (ICERs below $50,000 per QALY or dominant compared with usual care). We have provided the results of these scenarios in Table A23 (Appendix 10).

Discussion

Our reference case analysis showed that, for people with early-stage invasive, ER+, LN−, HER2− breast cancer, GEP tests were likely to be cost-effective compared with usual care without GEP testing. Blok et al[117] summarized published economic evaluations on GEP tests and also concluded that GEP tests were cost-effective compared with no tests.[118] In contrast, NICE concluded that GEP tests are unlikely to be cost-effective compared with no GEP testing.[118] The ICERs comparing GEP tests to usual care depended on test price, assumptions about usual care, and, specifically for Oncotype DX, assumptions about how test results would effect decision-making by clinicians and people with breast cancer. Moreover, the difference between our analysis and NICE diagnostic guidelines could be explained with assumptions related to risk classification and usual care.

Our results were sensitive to variation in testing costs. In the reference case, the incremental cost of GEP testing versus usual care ranges from a cost savings of $331 to a cost increase of $905 per person. This range is relatively narrow compared with the price difference of the tests. In the reference case, the only test that led to cost savings was Prosigna. However, sensitivity analysis showed that if the other tests had comparable pricing, they may be cost saving too. Since GEP tests are not currently available in Ontario, prices were based on assumptions. These prices may vary depending on how testing is implemented in Ontario.

Our cost-effectiveness results depended on the assumptions made in our usual care scenario. If high quality evidence exists to inform the risk classification with clinical tools used in the Ontario setting, and the prognosis of these subgroups were decided by clinical tools, we would assess the GEP tests versus usual care with clinical tools in practice. Lacking this type of evidence, we assumed that people did not receive clinical risk classification in the usual care strategy or before GEP testing. Hence, we did not explicitly model the impact of clinical judgment and reasoning on risk classification and chemotherapy decisions. Instead, for our usual care strategy, we modelled a hypothetical cohort of people with an average risk of recurrence. We then assumed that a proportion of people would receive adjuvant chemotherapy. In practice, adjuvant chemotherapy decisions are likely based on several factors, including patient characteristics (e.g., age), clinical guideline recommendations, comorbidities, contradictions, or patient values and preferences. Thus, the usual care in the model may be limited utility in representing usual care without GEP testing, and the economic evaluation methodology may favour GEP testing strategies over usual care.

Clinician and patient decision-making around GEP test results may have an impact on the cost-effectiveness of GEP testing, especially for Oncotype DX. Our analysis for Oncotype DX assessed the cost-effectiveness of classifying people into two categories (RS ≤ 25 and ≥ 26) rather than the previous three. Using the three categories to interpret Oncotype DX and data from TransATAC, we compared Oncotype DX with EndoPredict, Prosigna, and usual care, both in LN− and LN+ breast cancer based on clinical parameters from the same study.[52] For people with LN− breast cancer (the target population in our reference case analysis), Oncotype DX was cost-effective compared to usual care, but more costly and less effective compared with the other two tests. This is similar to a recent economic evaluation to assess the cost-effectiveness of the three tests.[175] These results, however, may be insufficient to prove that Prosigna and EndoPredict are superior, considering the potential practice change related to Oncotype DX after the TAILORx trial. In our model, we used data from the TAILORx trial and assumed that people with a Recurrence Score of ≤ 25 would make the same decision as low-risk people. This means that a larger proportion of people who would be classified as intermediate-risk (RS of 18–30) under the three-category system would forgo adjuvant chemotherapy. Nevertheless, we lack sufficient data to determine the extent to which the practice has changed after this trial and after the introduction of the new two-category classification method.

Some of our methods differed from those of NICE, including our assumptions about usual care and the way Oncotype DX is used to classify patients. First, based on Ontario clinical practice, we compared GEP tests to usual care without any clinical tool in our reference case analysis. However, NICE assumed patients would first be evaluated using a clinical tool (i.e., the NPI). The NPI is not widely used in Ontario. Second, for Oncotype DX, we used the data from the TAILORx trial while NICE used the TransATAC study.[26] The TAILORx trial considers the use of Oncotype DX with two risk categories (as opposed to three risk categories in TransATAC), which is relevant to the current clinical practice. These differences may explain the discrepancy in results.[26]

We also assessed the cost-effectiveness of GEP tests in subgroups including premenopausal people and people with LN+ breast cancer. For the premenopausal population, all GEP tests were cost-effective. For people with LN+ breast cancer, MammaPrint, Prosigna, and EndoPredict were likely to be cost-effective compared to no test. There is no evidence assessing the effectiveness of Oncotype DX using the two-category classification system for people with LN+ breast cancer. There is, however, evidence assessing the effectiveness of Oncotype DX in LN+ breast cancer using the three-category system. Applying three risk categories to the data from TransATAC, we find that, compared to usual care, Oncotype DX led to a cost increase and a 0.08 decrease in QALYs, making it unlikely to be cost-effective. In contrast, a recently published study conducted from a Canadian perspective found that Oncotype DX was cost-effective in people with LN+ breast cancer.[192] This difference likely occurred because that study's analysis included predictive validity. We included predictive validity in a scenario analysis in people with LN− breast cancer; however, based on limited evidence, we chose not to include this in LN+ patients. Incorporating predictive validity would likely improve the cost-effectiveness of Oncotype DX in this population.

We conducted an analysis comparing Oncotype DX to MammaPrint, and EndoPredict to Prosigna, because the studies for these tests included people with similar baseline risk levels.[26,51,52] Our analysis showed that Oncotype DX with two risk categories might be cost-effective compared with MammaPrint, and Prosigna might be cost-effective compared with EndoPredict.

Our sensitivity analysis compared the cost-effectiveness of GEP tests to usual care using a modified AOL tool,[26,51] which classifies people as either low or high risk. We chose this modified version over other clinical tools because this version is still publicly available and because RCTs were conducted to compare this tool with two GEP tests.[26,51] We assessed the cost-effectiveness of GEP tests versus usual care with this tool, and versus no test for the clinically low- and high-risk groups, respectively, indicated by the modified AOL tool. Our results suggested that Oncotype DX with this tool may be cost-effective compared to the modified AOL tool alone, and it may be cost-effective compared to no Oncotype DX for the low- or high-risk groups. This analysis suggested Oncotype DX might provide further benefits to patients even after they have received an evaluation using a clinical tool. However, MammaPrint was unlikely be cost-effective versus this tool alone. MammaPrint was also unlikely to be cost-effective versus no MammaPrint for low- or high-risk groups. We were unable to compare EndoPredict or Prosigna with the same clinical tool due to limited data.

Strengths and Limitations

Our analysis had several strengths. First, we used a lifetime horizon and a monthly cycle, which allowed us to capture both costs and outcomes over the course of people's lives, and the large variation in costs and outcomes that occur in the period immediately after a breast cancer diagnosis. Second, we used recent Ontario treatment costs in our analysis.[126] Third, we compared the test strategies, usual care practices, and chemotherapy acceptance rates that are most relevant in Ontario. Fourth, we conducted comprehensive scenario analyses and subgroup analyses to examine the impact of variability and uncertainty on our results.

Our analysis had limitations as well, many of which are discussed above. First, we were unable to fully incorporate the impact of clinical judgement, patient preferences, and clinical tools on risk classification. While we conducted several scenario analyses to try to address this issue, the cost-effectiveness of the GEP tests compared to usual care could be overestimated. Second, limited data constrained our ability to compare GEP tests. Third, market prices for the tests in Ontario are not available. Hence, we used estimates provided by manufacturers. Different implementation strategies could affect the actual test prices. Fourth, we assumed GEP testing has no direct impact on the quality of life. However, test results may have a psychological impact such as heightened anxiety for people classified as high risk.[193]

Conclusions

We found that GEP tests were likely cost-effective compared with usual care in people with ER+, LN−, HER2− breast cancer. We also found that all GEP tests except Oncotype DX were likely cost-effective in people with LN+ breast cancer, and that all GEP tests were likely cost-effective in the premenopausal population. We are uncertain about the cost-effectiveness of GEP tests compared with each other.

BUDGET IMPACT ANALYSIS

Research Question

From the perspective of the Ontario Ministry of Health, what is the potential budget impact in Ontario of publicly funding gene expression profiling (GEP) tests for people with early-stage invasive breast cancer?

Methods

Analytic Framework

We estimated the budget impact of publicly funding gene expression profiling (GEP) tests using the cost difference between two scenarios: (1) the current clinical practice, in which GEP tests are funded through the out-of-country program by the Ontario Ministry of Health and are conducted outside of Canada, and (2) the anticipated clinical practice with public funding for GEP tests conducted in Ontario. Figure 7 presents the budget impact model schematic.

We conducted a reference case analysis and sensitivity analyses. Our reference case analysis represented the analysis with the most likely set of input parameters and model assumptions. Our sensitivity analyses explored how the results were affected by varying input parameters and model assumptions.

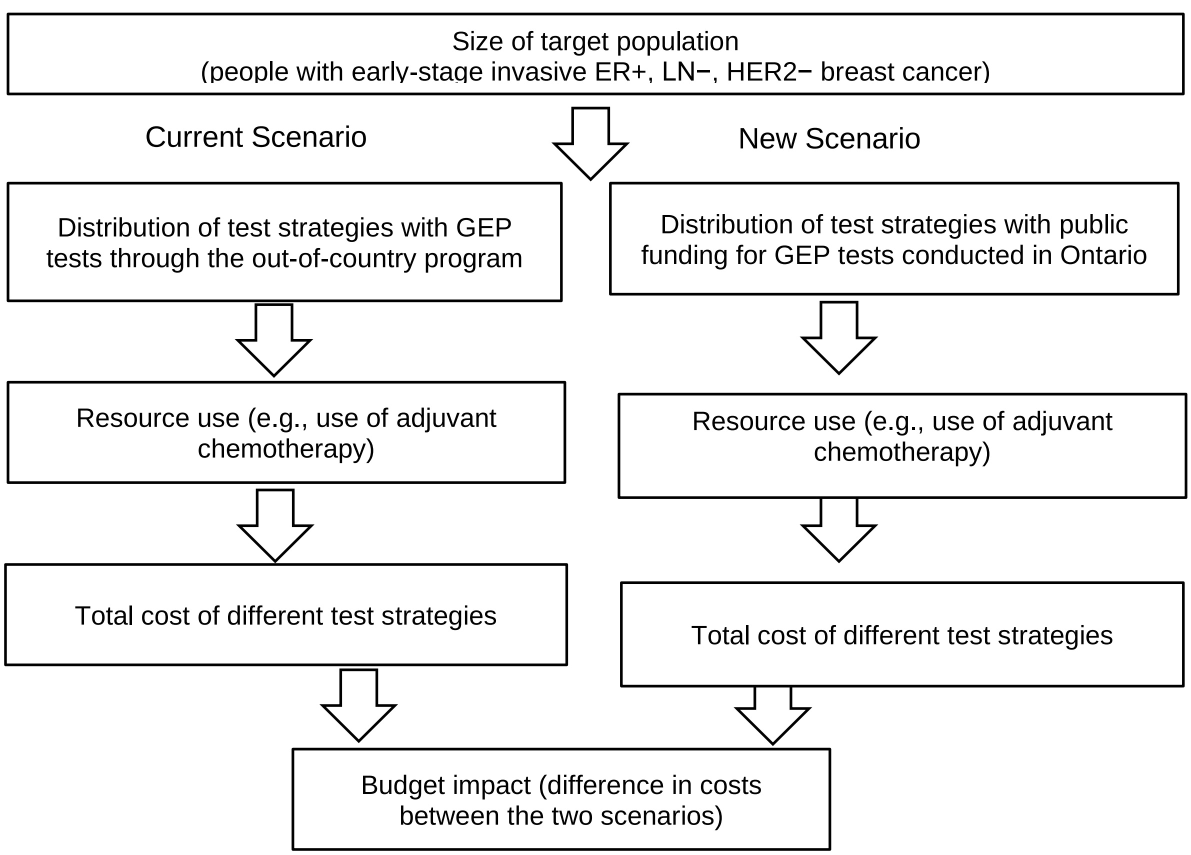


Figure 7: Schematic Model of Budget Impact

Key Assumptions

- In the current scenario, GEP tests in Ontario are funded through the out-of-country program. The number of people receiving GEP tests through other programs (i.e., private insurance, out-of-pocket) is negligible
- In the new scenario, all four GEP tests are publicly funded for people who have been diagnosed with breast cancer
- In the new scenario, where tests are conducted and publicly funded in Ontario, no additional resources are required to set up testing, from the perspective of Ministry of Health. Testing prices provided by the manufacturers include all relevant costs (i.e., human resources, platform, and consumable kits)
- Public funding will cause the uptake of all four GEP tests to increase
- There is no multiple testing. Each person tested receives a single GEP test
- In the new scenario, for those who receive a GEP test, 80% use Oncotype DX, 10% use EndoPredict, 5% use Prosigna, and 5% use MammaPrint. This assumption was based on the current uptake through the out-of-country program. In the 2017/2018 fiscal year, there were 2,030 requests for Oncotype DX in Ontario and 47 requests for EndoPredict (there were none in prior years). (Ministry of Health, email communication, January 14, 2019) There have been no requests for Prosigna or MammaPrint
- We have no evidence on the proportion of ER+, HER2− men diagnosed with early-stage breast cancer, so we assumed that the proportion of eligible men is the same as the proportion of eligible women

Target Population

Our target population includes people with early-stage, ER+, HER2− breast cancer. We considered the budget impact of funding GEP tests for people with LN− cancer only in the reference case, but also included people with LN+ breast cancer in the scenario analysis.

Table 26 summarizes our target population estimates. GEP tests are only applicable to newly diagnosed cases of breast cancer, and not to prevalent cases. To estimate the annual volume of eligible people, we used the Ontario population size,[194] the incidence rate of breast cancer,[6,195] the proportion of people with stage 1 or stage 2 breast cancer,[6] and the proportion of people who are estrogen and/or progesterone receptor–positive and HER2−.

In Ontario in 2018, 11,762 cases of female breast cancer were expected to be diagnosed.[6] Most breast cancer cases in Ontario are diagnosed at stage 1 (42.9%) or stage 2 (38.3%).[6] We assumed that 65% of people with stage 1 or stage 2 cancer are diagnosed with ER+ and HER2− breast cancer. Additionally, we assumed that all stage 1 cancer and 60% of stage 2 cancer are LN− (Sunnybrook Health Sciences Centre, Andrea Eisen and Maureen Trudeau, email communication, March 14, 2019). Therefore, if GEP tests are funded only for people with LN− breast cancer, 65% of stage 1 and 39% of stage 2 breast cancer patients would be eligible. If tests are funded for people with LN− or LN+ breast cancer, 65% of people with stage 1 or stage 2 breast cancer will be eligible.

To estimate the number of eligible women from 2019 to 2023, we first obtained the Ontario female population projection based on data from the Ontario Ministry of Finance.[194] Based on a stabilized trend of breast cancer incidence rate in recent years,[195] we assumed the incidence rate would be constant over time, and projected the number of newly diagnosed breast cancer cases eligible for GEP testing from 2019 to 2023.[194-196]

We used the same strategy to estimate the number of eligible men. In 2017, 230 men with breast cancer were diagnosed in Canada.[197] We used the Canadian population size in 2017[29] to estimate the proportion of men diagnosed with breast cancer. We then used the Ontario male population size to estimate the number expected to be diagnosed in Ontario.[194] Assuming the same percentages for stage 1 and stage 2 breast cancer for men as for women, we calculated the total number of men eligible for GEP testing in the next five years.

Table 26: Target Population for GEP Tests

Target Population	2018	2019	2020	2021	2022	2023
Ontario female population (all age groups)	7,334,472	7,442,811	7,545,191	7,641,047	7,730,043	7,819,328
Ontario male population (all age groups)	7,108,094	7,217,398	7,319,252	7,412,931	7,498,312	7,584,024
Number of women newly diagnosed with breast cancer	11,762	11,936	12,100	12,254	12,396	12,540
Number of men newly diagnosed with breast cancer	90	91	92	94	95	96
Number of people newly diagnosed with breast cancer	11,852	12,027	12,192	12,347	12,491	12,635
Early-Stage (Stage 1 or Stage 2) Invasive Breast Cancer						
Stage 1 breast cancer	5,084	5,160	5,231	5,297	5,359	5,421
Stage 2 breast cancer	4,539	4,606	4,670	4,729	4,784	4,839
Eligible people with LN−[a]	5,075	5,150	5,221	5,287	5,349	5,411
Eligible people with LN+[b]	1,180	1,198	1,214	1,230	1,244	1,258

Abbreviations: GEP, gene expression profiling; LN, lymph node.
Numbers may appear inexact due to rounding.
[a]65% of people with stage 1 breast cancer, and 39% of people with stage 2 breast cancer.
[b]26% of people with stage 2 breast cancer.

Current Intervention Mix

Currently, GEP tests are funded in Ontario through the Ministry's out-of-country program. Cancer Care Ontario[c] recommendations support the use of EndoPredict, Oncotype DX, and Prosigna for adjuvant chemotherapy decisions for people with ER+, LN−, HER2− breast cancer; and also suggest that Oncotype DX and Prosigna may be used in people with LN+ breast cancer.[29] The alternative to GEP testing is to use clinical expertise and/or decision tools (e.g., the PREDICT tool[163]), which use clinical and molecular characteristics to determine the potential benefit of adjuvant chemotherapy.

Oncotype DX is the most commonly used GEP test and is performed by the manufacturer in the United States. In the 2017/2018 fiscal year, there were 2,030 requests for Oncotype DX and 47 requests for EndoPredict. There were no requests for MammaPrint or Prosigna). Based on

[c] Cancer Care Ontario is now Ontario Health (Cancer Care Ontario).

this and on a 2013 study by Levine et al,[84] we estimated that the uptake of GEP tests through the out-of-country program is about 40% of eligible people and has been stable for several years. Therefore, in our current scenario, we assumed that 40% of eligible people receive Oncotype DX through the out-of-country program and that the remaining 60% do not get tested (Table 27).

Uptake of the New Intervention and Future Intervention Mix

With public funding of GEP tests conducted in Ontario, we expect that GEP tests will become more accessible to oncologists and people with breast cancer and, therefore, uptake will increase compared with the current scenario (Table 27). In the first year of our analysis, we assumed that 80% of eligible people would receive a GEP test in Ontario. In subsequent years, we assumed the uptake either would increase by 5% annually (in the reference case) or remain constant at 80% (in the scenario analysis).

Table 27: Target Population and Uptake of GEP Tests

		Year				
Scenario	**Test**	**2019**	**2020**	**2021**	**2022**	**2023**
Target population: ER+, LN−, HER2−		5,150	5,221	5,287	5,349	5,411
Reference Case Analysis						
Current scenario (out-of-country program, 40% uptake, no increase)	Tested	2,060	2,088	2,115	2,140	2,164
	Untested	3,090	3,133	3,172	3,209	3,246
New scenario, reference case analysis (high uptake 80%, 5% increase per year)	Tested	4,120	4,438	4,759	5,081	5,411
	Untested	1,030	783	529	267	0
Scenario Analysis						
New scenario, high uptake (80%), no increase in uptake	Tested	4,120	4,177	4,230	4,279	4,329
	Untested	1,030	1,044	1,057	1,070	1,082
New scenario, moderate uptake (60%), 5% increase per year	Tested	3,090	3,394	3,701	4,012	4,329
	Untested	2,060	1,827	1,586	1,337	1,082
New scenario, moderate uptake (60%), no increase in uptake	Tested	3,090	3,133	3,172	3,209	3,246
	Untested	2,060	2,088	2,115	2,140	2,164

Abbreviations: ER, estrogen receptor; GEP, gene expression profiling; HER2, human epidermal growth factor receptor 2; LN, lymph node.

In our reference case, we projected that 4,210 people would receive a test in Ontario in year 1, increasing to 5,411 in year 5. We assumed that 80% of people were tested with Oncotype DX, 10% with EndoPredict, and 5% with each of MammaPrint and Prosigna. The number of each GEP test projected for the next five years is summarized in Table 28.

Table 28: Number of Tests and Uptake of Each Type of Test

Scenario	Test	Year 1	Year 2	Year 3	Year 4	Year 5
Current scenario (out-of-country program), no increase in uptake	EndoPredict	0	0	0	0	0
	MammaPrint	0	0	0	0	0
	Oncotype DX	2,060	2,088	2,155	2,140	2,164
	Prosigna	0	0	0	0	0
	Tested (Total)	2,060	2,088	2,155	2,140	2,164
	Untested	3,090	3,133	3,172	3,209	3,246
New scenario in the reference case analysis (80% uptake, 5% increase in uptake per year)	EndoPredict	412	444	476	508	541
	MammaPrint	206	222	238	254	271
	Oncotype DX	3,296	3,550	3,807	4,065	4,329
	Prosigna	206	222	238	254	271
	Tested (Total)	4,120	4,438	4,759	5,081	5,411
	Untested	1,030	783	529	267	0

Resource Use and Costs

We included direct health care costs only, including health technology-associated (GEP test-related) and disease-associated (downstream breast cancer management–related) resources and costs. The inputs for the resource use and costs included in our budget impact analysis were derived from our primary economic evaluation (see Table A24, Appendix 11).

In our current scenario, eligible people may still receive testing through the out-of-country program. The difference in costs between the new scenario and the current scenario arises from the volume of testing, the type of test used, the price difference between testing in Ontario and through the out-of-country program (we assumed that the Ontario cost of Oncotype DX testing would be 90% of the out-of-country cost), and downstream costs.

We used our economic model to estimate the per person 5-year undiscounted cost for people who receive GEP testing and who do not receive GEP testing. We multiplied these costs by the number of people that we expected would receive tests under our scenarios (Table 27).

All costs were reported in 2018 Canadian Dollars.

Internal Validation

Formal internal validation was conducted by the secondary health economist. This included checking for errors and accuracy of parameter inputs and equations in the budget impact analysis.

Analysis

Reference Case Analysis

We calculated the budget required to publicly fund GEP tests in people with early-stage, ER+, LN−, HER2− breast cancer in Ontario by calculating the budget impact as the cost difference between our new scenario (public funding for GEP tests in Ontario, 80% uptake rate of GEP tests in the first year, with a 5% increase annually) and the current scenario (funding of GEP tests through the out-of-country program, 35% stable uptake in the next 5 years). Total costs were presented along with cost breakdowns (i.e., GEP tests, adjuvant chemotherapy, follow-up treatments).

Subgroups and Scenario Analyses

We conducted several scenario analyses that estimated the budget impact given:

- No funding for GEP as the current scenario
- Various assumptions on uptake
- Funding for people with LN− and LN+ breast cancer
- Various market shares of each GEP test
- Various prices for each GEP test
- Funding GEP tests for selected risk groups only
- Funding only one of the available four tests (e.g., 100% of tests are Oncotype DX)

Results

Reference Case

Table 29 presents the results of our reference case budget impact analysis. In our new scenario, 80% of eligible people receive GEP tests in Ontario in year 1. This increases to 100% by year 5. Compared to the current scenario, in which 40% of people get tests through the out-of-country program (with 60% receiving no test), the new scenario led to new costs of $1.29 million in year 1, increasing to $2.22 million in year 5, for a total cost increase of $8.13 million.

Table 29: Budget Impact Analysis Results: Reference Case

	Budget Impact, $ Million[a]					
Scenario	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**
Current Scenario: Publicly Funding GEP Tests Through the Out-of-Country program						
Test cost	11.14	11.30	11.44	11.57	11.71	57.16
Adjuvant chemotherapy cost	32.59	33.04	33.45	33.84	34.24	167.16
Incurred prior to distant recurrence	92.96	113.33	118.42	123.06	127.34	575.11
Incurred following distant recurrence	0.37	1.06	1.91	2.86	3.88	10.09
Incurred over last 3 months of life	0.12	0.51	1.10	1.81	2.60	6.15
Total cost	137.19	159.24	166.32	173.15	179.76	815.66
New Scenario: Publicly Funding GEP Tests Conducted in Ontario						
Test cost	18.57	20.01	21.45	22.91	24.39	107.33
Adjuvant chemotherapy cost	24.08	23.23	22.34	21.39	20.42	111.46
Incurred prior to distant recurrence	95.33	115.59	120.97	125.91	130.50	588.31
Incurred following distant recurrence	0.38	1.09	1.97	2.95	4.00	10.38
Incurred over last 3 months of life	0.12	0.52	1.13	1.86	2.68	6.32
Total cost	138.48	160.45	167.86	175.02	181.99	823.79
Budget Impact						
Test cost	7.43	8.71	10.01	11.33	12.68	50.16
Adjuvant chemotherapy cost	−8.51	−9.80	−11.12	−12.45	−13.82	−55.69
Incurred prior to distant recurrence	2.36	2.26	2.56	2.86	3.16	13.20
Incurred following distant recurrence	0.006	0.03	0.05	0.09	0.12	0.29
Incurred over last 3 months of life	0.001	0.01	0.03	0.05	0.08	0.17
Total cost	1.29	1.21	1.53	1.87	2.22	8.13

Numbers may appear inexact due to rounding.
[a]In 2018 Canadian dollars.

Sensitivity Analysis

Tables 30 and 31, and Table A25 (Appendix 11) summarize the results from our sensitivity analyses. Table 30 summarizes the different GEP market share scenarios of funding GEP tests in Ontario. When the uptake level of new scenario remains stable at the current level of 40%, publicly funding GEP tests conducted in Ontario would always be cost saving compared to Oncotype DX through out-of-country program, regardless of market share of GEP tests. If the uptake remains the same (40%) after public funding GEP tests in Ontario, then between $1.25 million and $1.30 million could be saved annually as we assume more people would choose a GEP test in Ontario that is less expensive than Oncotype DX (either in Ontario or out of country) with our assumption of the market share of four GEP tests.

Table 30: Budget Impact Sensitivity Analysis Results, Uptake of 40% After Funding, No Increase

Scenario	Budget Impact, $ Million[a]					
	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Same market share as reference case (80% Oncotype DX, 10% EndoPredict, 5% MammaPrint, and 5% Prosigna)	−1.30	−1.26	−1.26	−1.26	−1.25	−6.33
Oncotype DX only, in Ontario vs. out-of-country	−1.11	−1.13	−1.14	−1.16	−1.17	−5.72
EndoPredict only in Ontario vs. Oncotype DX out-of-country	−2.07	−1.82	−1.74	−1.64	−1.53	−8.80
MammaPrint only in Ontario vs. Oncotype DX out-of-country	−1.10	−0.95	−0.96	−0.97	−0.98	−4.95
Prosigna only in Ontario vs. Oncotype DX out-of-country	−2.82	−2.57	−2.48	−2.37	−2.25	−12.49

[a]In 2018 Canadian dollars.

Table 31: Budget Impact Sensitivity Analysis Results: Other Scenarios

Scenario	Budget Impact, $ Million[a]					
	Year 1	Year 2	Year 3	Year 4	Year 5	Total
New scenario vs. no funding for gene expression profiling tests	5.17	4.61	4.99	5.37	5.77	25.91
Current uptake level, Oncotype DX only through out-of-country vs. no funding for GEP tests	3.88	3.40	3.45	3.50	3.55	17.78
High uptake (80%), no increase in uptake	1.29	0.88	0.93	0.99	1.04	5.13
Moderate uptake (60%), 5% increase per year	−0.001	0.14	0.44	0.75	1.08	2.40
Moderate uptake (60%), no increase in uptake	−0.001	−0.19	−0.17	−0.14	−0.10	−0.60
Low uptake (40%), 5% increase per year	−1.30	−0.93	−0.66	−0.37	−0.07	−3.33
Both people with LN– and LN+ breast cancer	2.34	2.12	2.54	2.97	3.42	13.40
Oncotype DX market price[b]	3.07	3.13	3.59	4.07	4.57	18.43

Abbreviation: GEP, gene expression profiling; LN, lymph node.
[a]In 2018 Canadian dollars.
[b]If the budget impact is estimated according to the market price of Oncotype DX, rather than 90% of market price (assumed Oncotype DX price in Ontario).

Table 31 summarizes other sensitivity analyses results. In general, the budget impact remained relatively low (below $7 million per year).

If the new scenario was compared with no funding for any GEP testing (whether through the out-of-country program or in Ontario), the budget impact would be $5.17 million in year 1, increasing to $5.77 million in year 5. Another scenario, comparing maintaining the current uptake (40%) for Oncotype DX through the out-of-country program for next 5 years with the scenario of no funding for any GEP testing, yielded a budget impact of $3.88 million in year 1, increasing to $3.55 million in year 5.

Compared to publicly funding through the out-of-country program, funding both people with LN− and with LN+ breast cancer for GEP tests in Ontario would lead to a budget impact of $2.34 million in year 1, increasing to $3.42 million in year 5. If we assumed Oncotype DX testing conducted in Ontario would cost the same as the Oncotype DX market price through the out-of-country program, the budget impact of publicly funding GEP testing in Ontario versus through the out-of-country program would be $3.07 million in year 1, increasing to $4.57 million in year 5 (assuming an increase in uptake when GEP testing becomes available in Ontario).

Several scenarios reduced the budget impact and showed that publicly funding GEP testing conducted in Ontario could lead to cost-savings (see Table A24, Appendix 11). This included scenarios related to lower uptake rates and prices. If all test prices are reduced by 20%, there could be a cost savings of between $2.42 million and $2.65 million.

The budget impact was robust to other scenarios, including funding only one GEP test, or increasing the market share of Prosigna (the least expensive test) to 45% over 5 years while decreasing that of Oncotype DX to 40%.

Discussion

Eligible people in Ontario have the option of receiving a GEP test through the out-of-country program. We compared a new scenario of publicly funded GEP testing in Ontario to the current scenario, and our analysis suggested that publicly funding GEP tests in Ontario would increase costs by $1.29 million to $2.22 million annually. We conducted several sensitivity and scenario analyses and our results were generally robust, remaining lower than $7 million annually. One reason for our relatively low budget impact results is our assumption that the new scenario would be less expensive per test than the current scenario and that the projected cost increases are driven primarily by a greater uptake rate among eligible people. There are 2,000 tests annually in the current scenario (2017/2018 fiscal year). Even aggressive assumptions in the uptake rate (e.g., an 80% uptake rate in year 1, increasing to 100% in year 5), show a relatively low budget impact. It is unclear how moving testing to Ontario will impact uptake; therefore, we also modelled scenarios with little to no increased uptake. In these scenarios, we found GEP testing may be cost-saving compared with funding through the out-of-country program. We also evaluated the budget impact of funding GEP testing in Ontario versus no GEP testing at all and estimated the budget impact to be approximately $5 million per year.

Variations in test pricing can be another source of cost savings. We expect that GEP tests conducted in Ontario will be less expensive than tests conducted out of country. Although we used a reduced test price in our analysis, further reductions in price may be possible. Another factor contributing to the low budget impact arises from our assumption that, in the new scenario, more people would receive less expensive tests (i.e., Oncotype DX is the predominant and most expensive test in the current scenario).

Strengths and Limitations

There were several strengths to our budget impact analysis. First, our budget impact relied on our primary economic evaluation, which allowed us to consider both costs related to the tests and to downstream clinical outcomes. Additionally, our analysis was based on the most recent Ontario cost.[126]

Our analysis also had limitations. We based our uptake rates and market share estimates on current clinical practice, expert inputs, and assumptions, and thus were unable to determine the actual uptake of each GEP test. Uptake rates and market share of GEP tests may be influenced by many factors, including clinical evidence, patient preference, test prices, and implementation considerations. For example, patients may have a strong preference to receive chemotherapy to lower their recurrence risk such that GEP testing would not change clinical practice. There may be no benefit to providing GEP tests to these people. Another limitation is that our estimate of the proportion of eligible people with early-stage breast cancer was based on assumptions. Last, we based our market-price estimates on data provided by manufacturers because the market prices of tests in Ontario are not available. Actual prices will depend in part on how testing is implemented in Ontario. Nevertheless, the budget impact was robust to the variation in assumptions and price changes.

Conclusions

Publicly funding GEP testing conducted in Ontario is estimated to cost an additional $1.29 million to $2.22 million per year compared with funding GEP tests through the out-of-country program. Lower uptake, lower prices, and/or increased use of less expensive tests could further lower the budget impact.

PREFERENCES AND VALUES EVIDENCE

Objective

The objective of this analysis was to explore the underlying values, needs, and priorities of those who have lived experience of early-stage breast cancer, as well the preferences and perceptions of both patients and providers of gene expression profiling (GEP) tests.

Background

Gene expression profiling tests can be used as a decision-making tool to help decide if people who have been diagnosed with early-stage breast cancer should receive adjuvant chemotherapy. Those who are at low risk likely do not benefit from adjuvant chemotherapy.

In our analysis, we reviewed the quantitative literature for patient and physician preferences for GEP testing for breast cancer (Quantitative Evidence) and, in addition, we interviewed people and family members who have been diagnosed with and treated for early-stage invasive breast cancer, whether they received a GEP test or not (Direct Patient Engagement). We also considered the results from a review by the Canadian Agency for Drugs and Technologies in Health (CADTH) of the published qualitative evidence.[198]

Quantitative Evidence

Research Questions

- What is the relative preference of patients and providers for gene expression profiling (GEP) tests compared with non-genetic prognostic tests or no testing?
- What is the relative importance of key attributes of GEP tests, and what trade-offs between attributes are patients and providers willing to make?
- How do GEP tests impact patients' and providers' decisional conflict, psychological well-being, and quality of life?
- How satisfied are patients and providers with GEP tests?
- What are patients' and providers' knowledge and understanding of GEP tests and their use?

Methods

We conducted an evaluation of patient and health care providers' preferences for GEP testing as a literature survey using methods different from those of the clinical and economic systematic reviews. The objective was to describe and understand patients' and providers' values and preferences regarding GEP testing for early-stage invasive breast cancer. Results are summarized narratively in text and in tables.

Literature Search

We performed a targeted literature search of preferences and values on December 17, 2018, for quantitative studies published from inception until the search date. We used the Ovid interface to search MEDLINE only.

The search strategy was based on the economic literature search strategy, with a methodological search filter by Selva et al[199] applied, which limited the retrieval of studies to quantitative evidence of preferences and values. We further modified the search filter to include additional key terms relevant to psychological and emotional outcomes, specific types of health care providers, and patient or provider satisfaction. The final search strategy was peer reviewed using the PRESS Checklist.[41]

We created database auto-alerts in MEDLINE and monitored them for the duration of the assessment period. See Appendix 4 for our literature search strategies, including all search terms.

Eligibility Criteria

Studies

Inclusion Criteria

- English-language full-text publications
- Studies published from database inception until December 17, 2018
- Randomized controlled trials, cohort studies, cross-sectional studies that examined:
    - Patients' or providers' preferences for adjuvant chemotherapy treatment decision-making for early-stage invasive breast cancer, and

- Utility measures: direct techniques (standard gamble, time trade-off, rating scales) or conjoint analysis (discrete choice experiment, contingent valuation and willingness-to-pay, probability trade-off), or
- Non-utility quantitative measures: direct choice techniques, decision aids, surveys, questionnaires

Exclusion Criteria

- Animal and in vitro studies
- Nonsystematic reviews, narrative reviews, abstracts, editorials, letters, case reports, commentaries, and qualitative studies

Participants

Inclusion Criteria

- People with any type of breast cancer of any age (any breast cancer stage, tumour receptor status, lymph node status, number of lymph nodes involved, any menopausal status)
- Health care providers who use a GEP test or consult patients on the use or results of a GEP test

Exclusion Criteria

- People who do not have breast cancer (e.g., patients' family members, general public, etc.)

Interventions

Inclusion Criteria

- EndoPredict, MammaPrint, Oncotype DX, and Prosigna (PAM50)
- Unspecified general GEP test or hypothetical GEP test
- Comparator: standard test or other included GEP test (head-to-head comparison)

Exclusion Criteria

- Any GEP test not included above

Outcome Measures

- Preferences for GEP test and test characteristics and trade-offs
- Decisional conflict
- Psychological effects (e.g., anxiety, worry)
- Quality of life
- Satisfaction
- Knowledge and understanding of GEP tests and their use

Literature Screening

A single reviewer conducted an initial screening of titles and abstracts using Covidence,[42] and then obtained the full text 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.

Data Extraction

We extracted relevant data on study characteristics using a data form to collect information about the following:

- Source (e.g., citation information, contact details, study type)
- Methods (e.g., study design, study duration, participant recruitment)
- Outcomes (e.g., outcomes measured, outcome definition and source of information, unit of measurement, upper and lower limits [for scales], time points at which the outcomes were assessed)

Statistical Analysis

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.

Critical Appraisal of Evidence

We did not critically appraise the included studies. The purpose of our literature survey is to gain a broad overview of the quantitative evidence of preferences of patients and health care providers.

Results

Literature Search

The database search of the quantitative evidence of preferences and values yielded 370 citations published between inception and December 17, 2018. We identified 31 nonrandomized studies that met our inclusion criteria. We identified three additional studies[200-202] from reference lists and another study from auto-alerts[49], for a total of 35 studies. Figure 8 presents the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram for the literature search for quantitative evidence of preferences and values.

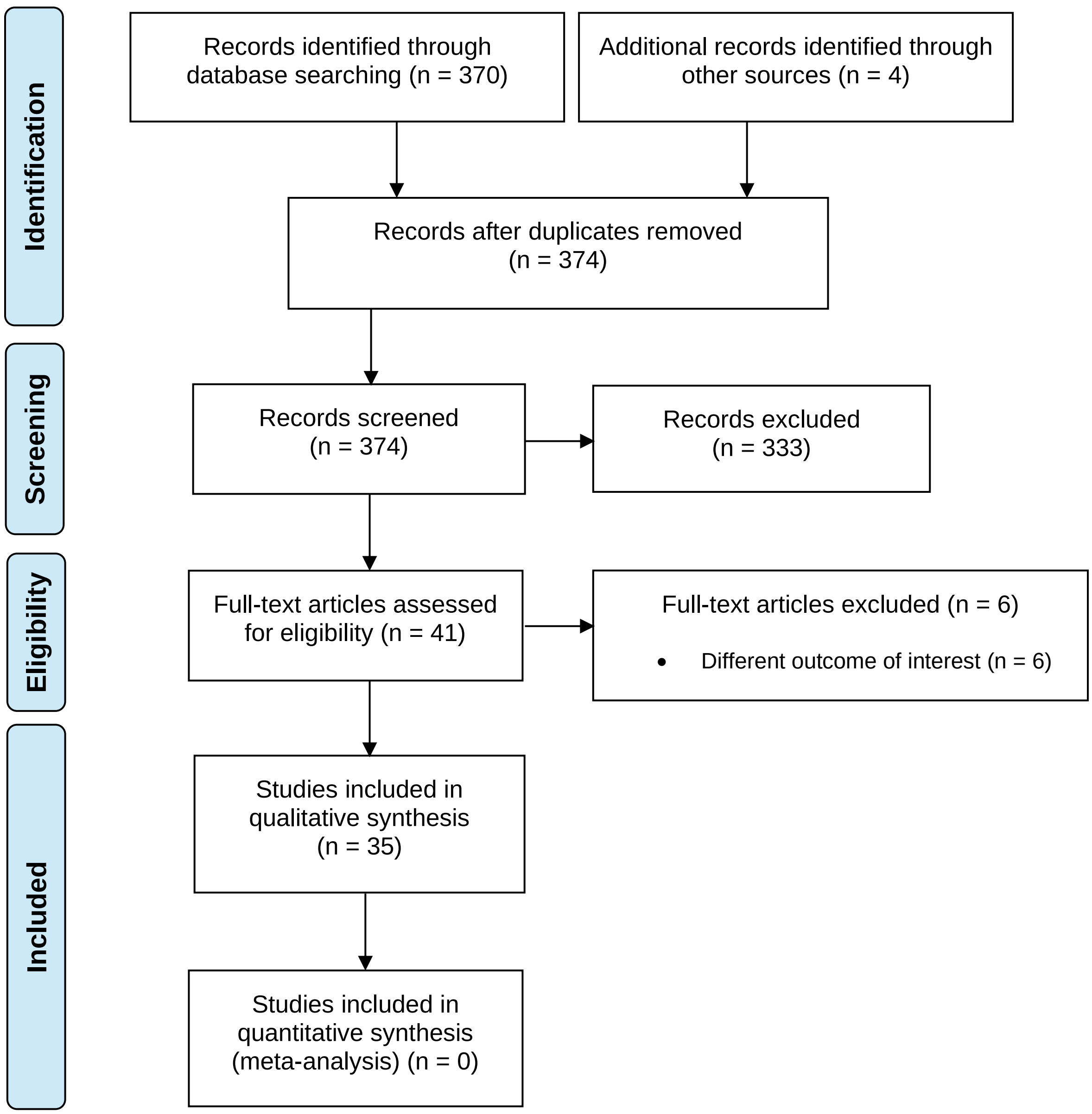


Figure 8: PRISMA Flow Diagram—Quantitative Evidence of Preferences and Values Search Strategy

Source: Adapted from Moher et al, 2009.[50]

Characteristics of Included Studies

Table 32 shows the characteristics of the included studies. Almost all studies evaluated the Oncotype DX test, three studies evaluated Prosigna,[88,89,203] three studies evaluated MammaPrint,[49,204,205] and one study evaluated EndoPredict.[105] Most of the studies focused on the quantitative evidence of preferences of people diagnosed with breast cancer or people who had previously been treated for breast cancer. The providers included in the studies were either oncologists, breast cancer surgeons, or radiation oncologists.

We found one systematic review that examined the experiences and attitudes toward risk of recurrence testing in women with breast cancer.[206] We did not include this systematic review in our analysis because it included studies within the general population (i.e., people who did not have breast cancer). It also included other non-genetic prognostic tests such as Adjuvant! Online. We reviewed the included studies within the systematic review and included the relevant studies that met our inclusion criteria. Similarly, we also found a systematic review by Yanes et al[207] that evaluated the psychosocial and behavioural outcomes of genomic testing in people diagnosed with cancer.[207] This systematic review was excluded because it evaluated all genomic tests for breast cancer (e.g., genome- or exome-wide sequencing, single nucleotide polymorphism tests), but we also searched the reference list for relevant studies.

Preferences for Gene Expression Profiling Tests

Table 32 summarizes the preferences of patients and physicians for GEP tests and the perceived value of the tests. In general, both patients and physicians valued the test because of the additional information the results provided and they reported that test results helped inform chemotherapy decision-making. Patients thought that GEP test results gave them a better understanding of a treatment option's chance of success.

Most patients also preferred an active or shared role in treatment decision-making when using GEP tests. People who took a GEP test reported that the risk of Recurrence Score provided important information when deciding whether to receive chemotherapy. Patients were generally more confident after GEP testing and gave more weight to GEP test results compared with non-genetic prognostic test results.[201,208]

Table 32: Preferences for GEP Tests

Author, Year	N	Measurement Method	Results
Hypothetical GEP Test			
Brewer et al, 2009[201]	165 patients	Questionnaire	• Both GEP and non-genetic tests elicited greater interest in chemotherapy when test results indicated high risk (89% vs. 26%, and 87% vs. 22%, respectively, $P < .001$) • Chemotherapy preferences were more strongly affected by recurrence risk information from GEP tests than non-genetic tests
DeFrank et al, 2013[208]	77 patients	Questionnaire	• High recurrence risk scores increased patients' perception of risk and preference for chemotherapy ($P < .001$) • Perceived risk mediated the effect of test results on chemotherapy preferences • When test results conflicted, patients gave more weight to GEP tests than to non-genetic tests
DeFrank et al, 2013[209]	132 patients	Questionnaire	• People who received the GEP test were more likely to be unsure about receiving chemotherapy than people who did not receive the GEP test ($P < .05$), suggesting that people who receive GEP tests are those who benefit most from the information provided • People who received the GEP test were less concerned their cancer would recur than people who did not receive the test • Most people who described their decision-making style as active received the test (75%); few people who described their style as passive received the test (12%; $P < .01$)
O'Neill et al, 2007[200]	139 patients	Questionnaire	• Participants rated potential benefits of testing higher than potential concerns ($P < .001$) • People willing to pay an average of $997 USD (95% CI $840–$1,155) for out-of-pocket testing; those who had heard of testing before the study were willing to pay more • Most participants preferred to be involved in treatment decision-making pre- and post-test; a majority preferred shared or active decision-making related to test results • 87% of participants stated they "definitely" wanted to know their test results, and 84% wanted to include the results in their treatment decision-making process • Participants trusted an intermediate test result the least • Participants who perceived more benefits, had chemotherapy, or had sufficient finances were more trusting of test results ($P = 0.2$, 0.2, and 0.1, respectively)
Panattoni et al, 2019[210]	833 patients	Questionnaire	• Confidence among those who reported high or low GEP scores was not significantly different from those who did not have a GEP test • Compared with people who reported no test, people who reported an intermediate score were less likely to report post hoc confidence (adjusted OR 0.34, 95% CI 0.20–0.58), as were people with an unknown Recurrence Score (adjusted OR 0.09, 95% CI 0.05–0.18) • People who reported unknown test receipt were also less likely to report post hoc confidence (adjusted OR 0.37, 95% CI 0.24–0.57)

Author, Year	N	Measurement Method	Results
Oncotype DX			
Evans et al, 2016[110]	193 patients	Questionnaire	• Perceived pros and cons of chemotherapy increased pre-test vs. post-test ($P < .001$) • Perceived risk of breast cancer recurrence decreased over time ($P = .004$)
Friese et al, 2017[211]	1,527 patients	SEER registry	• Among 420 people who reported low risk RS results, 65.0% indicated that RS shifted their opinion against chemotherapy; 73.1% of those who reported high scores reported that their RS result shifted their opinion toward chemotherapy
Gligorov et al, 2015[212]	94 physicians	Questionnaire	• 80% physicians (95% CI 70%–87%) agreed or strongly agreed that test results provided additional information
Kurian et al, 2018[213]	304 oncologists	Questionnaire	• When presented with information that the patient initially wanted chemotherapy, oncologists were more likely to recommend chemotherapy and order a GEP test before making a decision ($P < .001$) • When asked how their recommendation would change if the test predicted a high risk of distant recurrence (RS = 34), almost all oncologists would recommend chemotherapy; less likely to recommend chemotherapy for low RS (RS = 16) • For the less favourable prognosis scenario, virtually all oncologists would recommend chemotherapy, and few would order the test before making a decision • When presented with information that the patient initially wanted to avoid chemotherapy, oncologists were somewhat less likely to recommend chemotherapy ($P < .001$)
Lillie et al, 2007[214]	163 patients	Questionnaire	• Patients wanted further information for recurrence risk, specifically regarding treatment (i.e., how results affect treatment that physician recommends and the different treatments available); desire for additional information was related to recurrence risk • Patients wanted the least information about test development
Lipkus et al, 2011[215]	64 patients	Questionnaire	• Patients' most desired information about GEP tests (90%): what GEP tests are, accuracy in predicting recurrence risk, how results are used to guide treatment decision-making, what additional information providers use to guide treatment decisions • Patients' least desired information (71%–78%): stories of how patients used test results to inform treatment decisions, references to scientific studies, exercises to help clarify what is important to patients in treatment decision-making
Lo et al, 2010[85]	93 patients 17 medical oncologists	Questionnaire	• 94% of physicians stated that test provided additional information in treatment decision-making process • 88% of physicians believed results from test influenced their treatment recommendations • 83% of patients stated test influenced their treatment decision-making

Author, Year	N	Measurement Method	Results
Murciano-Gorof et al, 2018[216]	732 oncologists and surgeons	Questionnaire	• 10.9% of physicians believed that the test was too difficult to arrange • 54.4% of physicians believed that the test was too expensive • 94.3% of physicians believed that the test would help in the management of people with breast cancer • Medical oncologists ordered the test more frequently than surgeons (OR 3.37, *P* < .001) • Physicians were more likely to order the test if they believed testing would be covered by patients' insurance (OR 7.33, *P* < .001)
Ngoi et al, 2013[217]	200 patients 67 physicians	Questionnaire	• Patients regarded proven medical benefit, affordability, and accuracy as important criteria influencing test decisions • Patients' reasons for reversing testing decisions included ambiguity in management of intermediate test results • Patients' reasons for maintaining testing decisions: influence on management decisions and to facilitate better understanding of their condition • Fear of cancer recurrence was an important factor in patients' interest in testing • Physicians' most common reason to recommend testing was belief that results would influence management decisions • Physicians' most common reason to recommend against testing: lack of utility in influencing treatment decisions, cost, ambiguity in management of intermediate RS results
Ozmen et al, 2016[87]	165 patients NR physicians	Questionnaire	• Pre-test vs. post-test for physicians: 34.1% vs. 88% of physicians "strongly believed" that test result would contribute to final treatment decision • Pre-test vs. post-test for patients: 41.2% vs. 85% of patients "strongly believed" that tests would provide additional information
Patil et al, 2015[218]	119 oncologists (medical or surgical)	Questionnaire	• 54.62% of physicians used a test only sometimes and only when they felt that the test was necessary for the particular patient • 88.23% of physicians thought the test's risk classification was "somewhat useful" or "very useful" for treatment decision-making • 69.93% of physicians thought that ordering the test was easy • As perceived ease of use increased, the associated perceived usefulness also increased (*P* = .003) • The insurance status of that patient was negatively associated with physicians' use of the test
Richman et al, 2011[219]	78 patients	Questionnaire	• 95% of people agreed that test gave them a better understanding of the success of treatment options • 76% thought that the test was useful because it could predict if there was high Risk of Recurrence • 77% thought results could be trusted; 71% thought results were accurate

Author, Year	N	Measurement Method	Results
Torres et al, 2018[103]	71 patients	Questionnaire	• Patients' confidence increased in 54% of cases, stayed the same in 32%, and decreased in 14% • 74% of patients agreed or strongly agreed that test results made clear what choice of treatment was best for them (vs. 38% at pre-test)
Tzeng et al, 2010[193]	77 patients	Questionnaire	• 95% of patients thought test results gave them better understanding of the chance of success of treatment options • Most believed test results were accurate and found the test useful to determine with certainty whether their cancer had a high chance of recurrence • People with higher perceived benefits and concerns were less concerned about testing (*P* = .001)
Unspecified GEP Test			
Seror et al, 2013[220]	43 patients	Questionnaire	• Main reason for agreeing to undergo testing was to access the most appropriate treatment (67.4% of patients)

Abbreviations: CI, confidence interval; GEP, gene expression profiling; NR, not reported; OR, odds ratio; RS, Recurrence Score; SEER, Surveillance, Epidemiology, and End Results.

Satisfaction

Patients were generally satisfied with GEP tests and would take the test again (if needed) or recommend its use to others. At 12 months after diagnosis, patients still felt satisfied with their decision to undergo testing and still believed that test results influenced their treatment decision.[85] Similarly, many physicians would also use GEP tests again in the future.[85] Table 33 summarizes the results for patient satisfaction.

Table 33: Results for Test Satisfaction

Author, Year	N	Measurement Method	Results
Hypothetical GEP Test			
O'Neill et al, 2007[200]	139 patients	Questionnaire	• Participants anticipated the most potential regret for an intermediate test result ($P < .001$); those who expressed more concerns about testing anticipated greater regret ($P < .02$)
MammaPrint			
Retel et al, 2009[205]	77 patients	Questionnaire	• Satisfaction of receiving test per risk group was 76% • Overall satisfaction from diagnosis to time of interview (approximately 2 months after surgery) was 82%
Retel et al, 2013[204]	347 patients	Questionnaire	• 97% of patients were satisfied with their experience from diagnosis to time of questionnaire • 94% of patients expressed overall satisfaction with the test information received • 28% were unsatisfied with the waiting time for results • 9% were dissatisfied with the way physicians conveyed results
Oncotype DX			
Friese et al, 2017[211]	1,527 patients	SEER registry	• High satisfaction with decision-making with RS and receipt of chemotherapy; scores did not differ substantively according to whether patients did or did not receive chemotherapy • Among people who received RS, 63.9% reported that the test was "very" or "extremely" helpful
Lo et al, 2010[85]	93 patients 17 medical oncologists	Questionnaire	• 97%–100% of physicians would order or use the test again • 95% patents said they were glad they took the test • At 12 months, 92.5% of patients continued to feel satisfied that they had used the test, and 80.6% continued to believe the results influenced their treatment decision • At 12 months, 95.5% of patients were satisfied with their adjuvant treatment decision • Patients who were not satisfied noted a negative impact on their quality of life, treatment side effects, and a negative impact on their self-image
Richman et al, 2011[219]	78 patients	Questionnaire	• Most people reported being satisfied with the test • 96% said they would have the test again if needed • 95% would recommend the test to others
Tzeng et al, 2010[193]	77 patients	Questionnaire	• 96% would have the test again if needed • 95% would recommend the test to others

Abbreviations: GEP, gene expression profiling; RS, Recurrence Score; SEER, Surveillance, Epidemiology, and End Results.

Decisional Conflict

Gene expression profiling tests generally decreased the uncertainty among patients about which option to choose (decisional conflict). Many studies noted a significant decrease in either the total or subscale scores of the Decisional Conflict Scale post versus pre-test. The Decisional

Conflict Scale consists of 16 items in five response categories. The scale measures personal perceptions of uncertainty when choosing among options, of modifiable factors contributing to uncertainty (e.g., feeling uninformed, unclear about personal values, unsupported decision-making), and of effective decision-making (e.g., feeling the choice is informed, value-based). Table 34 summarizes the results for decisional conflict.

Table 34: Results for Decisional Conflict

Author, Year	N	Measurement Method	Results
EndoPredict			
Fallowfield et al, 2018[105]	136 patients	Decisional Conflict Scale	• Significant decrease in total pre-test vs. post-test score: 17.74 (SD 13.59) vs. 14.59 (SD 14.26; *P* < .022) • No difference in pre- vs. post-test scores for patients who had treatment changes (upgraded or downgraded therapy decisions) • For patients with unchanged treatment decisions, significant difference in pre- vs. post-test score: 16.90 (SD 12.77) vs. 12.11 (SD 11.85; *P* = .001)
MammaPrint			
Wuerstlein et al, 2019[49]	430 patients	Decisional Conflict Scale	• Overall and subscale scores improved after test (*P* < .001) • Test risk result and post-test chemotherapy recommendations were strongly associated with post-test decisional conflict • Among the high-risk test group, discordant results between initial chemotherapy and test result did not affect decisional conflict • Among the low-risk test group, the concordant group had lower decisional conflict
Oncotype DX			
Davidson et al, 2013[137]	147 patients	Decisional Conflict Scale	• After post-RS discussion with the provider, the total score and all subscores of the Decision Conflict Scale significantly decreased (*P* < .001)
Eiermann et al, 2013[83]	325 patients	Decisional Conflict Scale	• Mean pre-test vs. post-test score: 1.72 vs. 1.61 (*P* = .028) • Significant decrease in decisional conflict post-test for all patients
Holt et al, 2013[143]	40 patients	Decisional Conflict Scale	• Mean pre-test vs. post-test score: 14.8 (95% CI 10.9–18.7) vs. 10.7 (95% CI 6.9–14.4; *P* = .03) • Significant decrease post-test vs. pre-test for informed and uncertainty subscores; no changes for values clarity, support, or effective decision subscores
Kuchel et al, 2016[111]	132 patients	Decisional Conflict Scale	• Mean pre-test vs. post-test score: 22.1 vs. 12.7 (*P* < .0001) • Significant decrease post-test vs. pre-test for informed, clarity, and effective decision subscores; no change for support subscore
Levine et al, 2016[84]	956 patients	Decisional Conflict Scale	• Mean total pre-test vs. post-test score: 34 vs. 19 (*P* < .0001) • Significant decrease in all mean subscores (*P* < .0001) • Significant difference in mean scores between patients who chose and did not choose chemotherapy

Author, Year	N	Measurement Method	Results
Lo et al, 2010[85]	93 patients	Decisional Conflict Scale	• Mean pre-test vs. post-test results: 1.99 (SD 0.62) vs. 1.69 (SD 0.50; *P* < .001)
Sulayman et al, 2012[221]	81 patients	Decisional Conflict Scale	• Moderate decisional conflict (mean 1.70, SD 0.50) • 30% of people reported problematic levels of decisional conflict (scores ≥ 2)
Yamauchi et al, 2014[222]	116 patients	Decisional Conflict Scale	• Total scores and subscale scores all decreased post-test (*P* = .014 for informed subscale; *P* < .001 for all others) • Mean total score improved by 26% post-test
Prosigna			
Hequet et al, 2017[88]	158 patients	Decisional Conflict Scale	• Significant decrease post-test vs. pre-test for total score and informed, values clarity, uncertainty subscores; no change in support and effective decision subscores
Martin et al, 2015[203]	183 patients	Decisional Conflict Scale	• No change in mean pre-test vs. post-test score: 16.90 (SD 12.5) vs. 16.95 (SD 12.7; *P* = .957)
Wuerstlein et al, 2016[89]	198 patients	Decision Conflict Scale	• Total mean score pre-test vs. post-test: 17.0 vs. 12.8 (*P* < .001) • Patients had greater knowledge about breast cancer status and treatment options (values clarity subscore) and higher engagement in informed decision-making (informed subscore) post-test compared with no test (*P* < .01) • Less uncertainty (uncertainty subscore) and more effective decision-making (effective decision subscore) post-test (*P* < .05)
Any GEP Test			
Seror et al, 2013[220]	43 patients	Questionnaire	• 42.4% and 30.3% of patients expressed decisional conflict on informed and values clarity subscales • 9.1% of patients expressed decisional conflict on uncertainty subscale • Education levels of at least secondary school were associated with both higher global decisional conflict scores (median 2.1 vs. 1.95, *P* = .04) and higher values clarity subscale scores (median 2.1 vs. 1.9, *P* = .03). • High or very high risk perception of chemotherapy-related side effects was significantly associated with a higher support subscale score (median 2.0 vs. 1.7, *P* = .02) • No significant relationship between patients' decisional conflict scores and their understanding of their genomic test results (*P* > .12 in all cases)

Abbreviations: CI, confidence interval; GEP, gene expression profiling; RS, Recurrence Score; SD, standard deviation.

Quality of Life

All studies used the Functional Assessment of Cancer Therapy (FACT) instrument to measure health-related quality of life among patients. There is a general cancer version of the test (FACT-G), as well as a breast cancer–specific version (FACT-B). Generally, studies did not find a significant change in quality of life pre- versus post-test, although some studies did find changes within certain subscale scores (mainly emotional and physical well-being subscales[88,89]). Table 35 summarizes the results for quality of life.

Table 35: Results for Quality of Life

Author, Year	N	Measurement Method	Results
MammaPrint			
Retel et al, 2013[204]	347 patients	FACT-B	• Older age was associated with better health-related quality of life (*P* > .001) • Risk perception was associated with lower health-related quality of life (*P* < .001) • Adjusting for demographic and process factors, only the clinically high/no MammaPrint group reported lower health-related quality of life compared with the reference group (clinically low/MammaPrint low)
Oncotype DX			
Lo et al, 2010[85]	93 patients	FACT-B	• Mean pre-test vs. post-test scores: 112.2 (SD 17.4) vs. 114.3 (SD 18.6; *P* = .55) • Scores also stable between pre-test vs. post-test at 12 months (*P* = .49)
Sulayman et al, 2012[221]	81 patients	FACT-B	• Decision style (active/shared or passive) and RS category were associated with quality of life (*P* < .04) • Among people who preferred a passive role, those with intermediate results reported poorer quality of life (compared with low or high RS)
Prosigna			
Hequet et al, 2017[88]	151 patients	FACT-G, version 4	• Significant difference in emotional well-being subscore post-test vs. pre-test (*P* < .001); no difference in physical, social/family, or functional well-being subscores
Martin et al, 2015[203]	183 patients	FACT-G, version 4	• No change in mean pre-test vs. post-test score: 79.19 (SD 15.6) vs. 79.57 (SD 14.6; *P* = .713)
Wuerstlein et al, 2016[89]	198 patients	FACT-G, version 4	• Test results improved emotional and functional well-being subscores for people categorized as high ROR (*P* < .05 and < .01, respectively)

Abbreviation: FACT-B, Functional Assessment of Cancer Therapy–Breast; FACT-G, Functional Assessment of Cancer Therapy–General; ROR, Risk of Recurrence; RS, Recurrence Score; SD, standard deviation.

Psychological Effects

Almost all studies used the State–Trait Anxiety Inventory to evaluate anxiety, which was the primary psychological outcome. The instrument is based on a four-point Likert scale and consists of 40 questions. The studies reported on anxiety in two broad categories: state-anxiety and trait-anxiety. State-anxiety refers to short-term anxiety related to a specific event. Trait-anxiety refers to the predisposition of a person to react with anxiety in stressful situations. In general, changes were found for state-anxiety but not for trait-anxiety.

People who chose to downgrade their chemotherapy recommendations had lower anxiety scores, whereas those who chose to upgrade their chemotherapy recommendations had higher anxiety scores. Anxiety was also typically impacted by the GEP test score, with high-risk test scores causing more anxiety. Table 36 summarizes the psychological effects of GEP testing.

Table 36: Results for Psychological Effects

Author, Year	N	Measurement Method	Results
Hypothetical GEP Test			
O'Neill et al, 2007[200]	139 patients	Questionnaire	• Participants' worry increased as function of GEP test result (*P* < .001)
EndoPredict			
Fallowfield et al, 2018[105]	149 patients	State–Trait Anxiety Inventory	• Anxiety scores were stable in patients with unchanged decisions for endocrine therapy alone or chemoendocrine therapy • Patients who had downgraded therapy had significantly lower anxiety scores (*P* = .045) • Patients who had upgraded therapy had significantly higher anxiety scores (*P* = .001) • No significant difference in scores pre- vs. post-test for people who had high or low anxiety
MammaPrint			
Retel et al, 2009[205]	74 patients	Questionnaire, Cancer Worries scale	• People with discordant clinical low/MammaPrint high and clinical high/no MammaPrint (due to failure process) had the highest negative affect scores • 43% of patients with clinical low/poor signature and 29% clinical high/no MammaPrint often worried about cancer recurrence
Retel et al, 2013[204]	347 patients	10 items adapted from Lynch's distress scale 7-item version of Lerman's Cancer Worry Scale	• Clinically low/MammaPrint low group had lowest distress, similar to the clinically high/MammaPrint low group (*P* = .18) • Higher distress was associated with unavailable test results, discordant groups, and clinically high/MammaPrint high group • Higher levels of worry were seen in people with lower satisfaction (*P* < .001) and higher perceived risk (*P* < .001)
Wuerstlein et al, 2019[49]	430 patients	State–Trait Anxiety Inventory	• Scores significantly improved pre- vs. post-test (*P* < .001); slight increase among high-risk people, and decrease among low-risk people • Trait-anxiety remained virtually unchanged among all patients • Test risk category and post-test chemotherapy recommendation were strongly associated with post-test decisional conflict (*P* < .001) • Among high-risk test group, discordant results between initial chemotherapy and test result did not impact anxiety • Among low-risk test group, concordant group had lower anxiety
Oncotype DX			
Evans et al, 2016[110]	193 patients	Questionnaire	• Cancer-related distress did not increase pre-test vs. post-test (*P* = .09)

Author, Year	N	Measurement Method	Results
Lo et al, 2010[85]	93 patients	State–Trait Anxiety Inventory	• Mean scores decreased significantly over time (*P* = .007) • Trait-anxiety did not change significantly (*P* = .27) • State-anxiety was positively correlated with decisional conflict both pre- and post-test (*P* = .001 and < .001, respectively)
Sulayman et al, 2012[221]	81 patients	15-item Impact of Event Scale or distress, 4-point Likert scale for cancer worry	• Moderate cancer-related distress (mean 19.10, SD 17.50) • 38.7% of people reported high levels of distress • Moderate cancer worry (mean 1.70, SD .80) • Among people who preferred an active/shared role in care, RS was unrelated to distress; among people who had an intermediate RS, people who preferred passive role had higher distress (*P* < .008)
Tzeng et al, 2010[193]	77 patients	4 items adapted from Cancer Worry Scale	• 26% of people agreed or strongly agreed that the test result made them worried and anxious • Greater distress was associated with higher RS • Stronger feelings of distress was related to getting chemotherapy, not getting radiation, and more frequent worrying of recurrence
Prosigna			
Hequet et al, 2017[88]	200 patients	State–Trait Anxiety Inventory	• State-anxiety was significantly decreased post-test vs. pre-test (*P* = .02); no change in trait-anxiety (*P* = .115) • Test was most helpful in decreasing anxiety for people with low-risk ROR
Martin et al, 2015[203]	180 patients	State–Trait Anxiety Inventory	• State-anxiety pre-test vs. post-test: 42.61 (SD 12.5) vs. 39.79 (SD 13.3; *P* = .003) • No significant change in trait-anxiety pre- vs. post-test (*P* = .858) • Significant association between changes in state- and trait-anxiety and ROR category
Wuerstlein et al, 2016[89]	198 patients	State–Trait Anxiety Inventory	• Significant association between changes in score and ROR risk status (*P* < .01) • Knowledge of test results decreased anxiety in people with low ROR

Abbreviation: GEP, gene expression profiling; RS, Recurrence Score; ROR, Risk of Recurrence; SD, standard deviation.

Knowledge and Understanding

Patients were often misinformed regarding the prognostic ability of GEP tests and did not understand that recurrence risk is conditional on further treatment. Knowledge about recurrence risk was generally low among patients. However, most patients knew that the test can help people avoid chemotherapy and that chemotherapy is most beneficial for people with high-risk test scores. In addition, people who actively sought information were more knowledgeable about GEP testing than those who did not.[215] Physicians included in the studies showed a higher level of understanding of recurrence risk, and more than 90% were aware of GEP testing or did not have difficulty interpreting test results.[218] Table 37 summarizes the results on patients' and providers' knowledge and understanding of GEP testing.

Table 37: Results for Knowledge and Understanding of GEP Testing

Author, Year	N	Measurement Method	Results
Hypothetical GEP Test			

Brewer et al, 2009[202]	163 patients	Rapid Estimate of Adult Literacy in Medicine	• Average health literacy score: 63.6 (range 30–66)
			  ○ High literacy (≥ 63): 125 people
			  ○ Low literacy (< 63): 38 people
			• People with lower health literacy gave higher mean estimates of recurrence risk for a hypothetical group of people with early-stage breast cancer than people with higher health literacy (52% vs. 30%, $P < .001$)
			• People with lower health literacy gave more variable estimates of recurrence risk
			• When making chemotherapy decisions using risks presented in verbal formats, decisions by people with lower health literacy were less sensitive to the difference between low and high recurrence risk
			• People with lower health literacy expressed lower ease of understanding than people with higher health literacy ($P = .002$)

MammaPrint			
Retel et al, 2009[205]	77 patients	Questionnaire	• 87% of patients scored incorrect on questions about the predictive accuracy of the test; 66% scored incorrect on questions about the consequences of the test
Retel et al, 2013[204]	347 patients	Questionnaire	• 6% of patients had heard of the test before diagnosis • Knowledge about GEP testing was relatively high among participants (mean correct answers was 75%) • 43% of patients did not know if the result of the genomic profile was always correct; 53% did not know if the recurrence risk was 50% within the next 10 years if the test result was high • People with relatives who had undergone chemotherapy answered more questions correctly ($P = .006$) • Knowledge about recurrence risk was low (mean knowledge score = 67%) • People with higher numeracy scores, higher health literacy, more education, less comorbidity had higher knowledge scores • People diagnosed in the 12 months before the questionnaire, had active decision-making roles, or had fewer reported concerns about testing, had higher knowledge scores
Oncotype DX			

Lillie et al, 2007[214]	163 patients	Rapid Estimate of Adult Literacy in Medicine	• Average health literacy score: 63.6 (range 30–66)
			  ○ High literacy (≥ 63): 125 people
			  ○ Low literacy (< 63): 38 people
			• 58% of people with higher health literacy reported a preference for active decision-making; 41% preference among people with lower health literacy

Author, Year	N	Measurement Method	Results
Lipkus et al, 2011[215]	64 patients	Questionnaire	• People less likely to know that test predicts Risk of Recurrence conditional on further treatment • <50% acknowledged that the test is not always correct • 87% knew that the test was to help people avoid unneeded chemotherapy; 92% knew that test served as a decision-making aid • 76% recognized that chemotherapy was most beneficial for people with high recurrence risk • 75% knew that the test results were based on breast tumour genes • 76% knew that the sample for the test is from breast tissue; 82% knew that the sample was taken during surgery • Patient knowledge was negatively associated with increasing age (*P* < .0002) and positively associated with educational level (*P* < .0001) • Top 4 sources of information for patients were internet (48%), health care provider (31%), pamphlets/brochures (22%), and books (12%) • Patients who sought information were significantly more knowledgeable about GEP testing than those who did not seek information
Lo et al, 2010[85]	93 patients 17 medical oncologists	Questionnaire	• 87% patients stated they understood how the test worked, and 89% felt that the results were easy to understand
Ngoi et al, 2013[217]	200 patients 67 physicians	Questionnaire	• 40% of patients indicated previous awareness of testing • 91% of physicians were aware of testing
Patil et al, 2015[218]	119 oncologists (medical or surgical)	Questionnaire	• 92.92% of physicians did not have difficulty interpreting test results
Tzeng et al, 2010[193]	77 patients	Questionnaire	• 11% of patients had heard of the test before diagnosis
Any GEP Test			
Seror et al, 2013[220]	43 patients	Questionnaire	• 62.7% of patients misunderstood the test results • Good understanding of test results was not significantly related to perceived risk of chemotherapy-related side effects (*P* = .45)

Abbreviation: GEP, gene expression profiling.

Discussion

To our knowledge, our literature survey is the first to summarize the evidence on the quantitative evidence of preferences and values of patients and providers for GEP testing. Other systematic reviews have been published on genomic testing[207] and recurrence testing[206] in breast cancer, but they were broader in scope and not as extensive in their evaluated outcomes.

Patients and physicians reported valuing the added information that GEP testing provides in the treatment decision-making process. Patients reported being satisfied with GEP testing, and test results generally reduced decisional conflict and psychological outcomes. Patients' risk category based on GEP test results also impacted outcomes.

The quantitative literature on patient and provider preferences for GEP testing was heterogeneous in terms of patient population and study methods. Unlike our clinical evidence review on the effectiveness of GEP testing, we included both prospective and retrospective studies in our quantitative evidence of preferences literature survey. While some of the included studies were embedded within prospective studies, we also included cross-sectional and retrospective studies, which may be more prone to methodological limitations. As part of our literature survey methodology, we did not critically appraise the literature, so we do not know the impact of potential risk of bias on study results. However, despite the clinical and methodological heterogeneity, we still found the same general trends within the results, which suggests the evidence is generalizable.

The majority of the evidence focused on the Oncotype DX test; however, the results from studies on EndoPredict, MammaPrint, and Prosigna were similar. We did not find any comparative studies that evaluated all four GEP tests, and so we cannot determine whether patients or physicians preferred one GEP test over another. However, compared with non-genetic prognostic tools, patients seem to prefer GEP testing to no GEP testing because they feel the information is more individualized than with other decision-making tools.

Qualitative Evidence—Direct Patient Engagement

Ontario Health (Quality) collaborated with the Canadian Agency for Drugs and Technologies in Health (CADTH) to conduct this health technology assessment. CADTH conducted a review of qualitative literature on patient perspectives.[198]

Background

Exploring patient preferences and values provides a unique source of information about people's experiences of a health condition and the health technologies or interventions used to manage or treat that health condition. It includes the impact of the condition and its treatment on the patient, the person with the health condition, their family and other caregivers, and the person's personal environment. Engagement also provides insights into how a health condition is managed by the province's health system.

Information shared from lived experience can also identify gaps or limitations in published research (e.g., outcomes important to those with lived experience that are not reflected in the literature).[223-225] Additionally, lived experience can provide information and perspectives on the ethical and social values implications of health technologies or interventions.

Because the needs, preferences, priorities, and values of those with lived experience in Ontario are often inadequately explored in published literature, we may speak directly with people who live with a given health condition, including those with experience with the technology or intervention we are exploring.

Methods

Partnership Plan

The engagement plan for this health technology assessment focused on consultation to examine the experiences of people who have been diagnosed with early-stage breast cancer and those of their families and other caregivers. We engaged people via telephone interviews and follow-up was done through email.

We used a qualitative interview, as this method allowed us to explore the meaning of central themes in the experiences of people who have been diagnosed with early-stage breast cancer, as well as those of their families and caregivers.[226] The sensitive nature of exploring people's experiences of a health condition and their quality of life are other factors that supported our choice of an interview methodology.

Participant Outreach

We used an approach called purposive sampling,[227-229] which involves actively reaching out to people with direct experience of the health condition and health technology or intervention being reviewed. We reached out by email to individuals who have experience with early-stage breast cancer and the GEP test, various clinical experts, health teams in hospitals that provide care for patients with early-stage breast cancer, and organizations and support groups.

Inclusion Criteria

Patients and their family members who have been actively managing their condition after being diagnosed with early-stage breast cancer.

Exclusion Criteria

We did not set specific exclusion criteria.

Participants

For this project, we interviewed six people who had been diagnosed with breast cancer and one family member, all of whom were living in Ontario. Participants were from different socio-economic backgrounds. Participants shared their experiences and perceptions through phone and email. Of the six people interviewed who had been diagnosed with breast cancer, three had received GEP tests, while the other three had experience with pathology testing.

Approach

At the beginning of the interview, we explained the role of our organization, the purpose of the health technology assessment, the risks of participation, and how participants' personal health information would be protected. We gave this information to participants in a printed letter of information (Appendix 12). We obtained participants' verbal consent before starting the interview. With participants' consent, we audio-recorded and then transcribed the interviews.

Interviews lasted approximately 30 minutes. Interviews were loosely structured and consisted of a series of open-ended questions. Questions were based on a list developed by the Health Technology Assessment International Interest Group on Patient and Citizen Involvement in Health Technology Assessment.[230] Questions focused on the impact of early-stage breast cancer on patients' and families' quality of life, and their perceptions of the benefits or limitations of receiving a GEP test as a tool to inform treatment decision-making. See Appendix 13 for our interview guide.

Data Extraction and Analysis

We used a modified version of a grounded-theory methodology to analyze interview transcripts. The grounded-theory approach allowed us to organize and compare information on experiences across participants. This method consists of a repetitive process of obtaining, documenting, and analyzing responses while simultaneously collecting, analyzing, and comparing information.[231,232] We used the qualitative data analysis software program NVivo[233] to identify and interpret patterns in interview data. The patterns we identified then allowed us to highlight the impact of early-stage breast cancer and GEP testing on the people, family members, and caregivers we interviewed.

Results

Impacts of Early-Stage Breast Cancer on Patients and Caregivers

Physical Impacts

Before being diagnosed with early-stage breast cancer, people described the physical symptoms—lumps in or around their breasts—which led them to go to the doctor and get the necessary tests done. This was a period of uncertainty for the patient as they were not yet aware of the cause of the lumps:

> *I found a lump on my, well I didn't find a lump, I found something on my breast, which I originally thought was some sort of muscle. And I wasn't concerned, but it lasted quite a while. So, I thought when I see my doctor, I will discuss it with him.*

People we interviewed also highlighted the uncertainty they felt due to delays in diagnosis. Delays may be due to an inconclusive mammogram, which then requires the patient to go through additional steps. It may take weeks to receive a final diagnosis of early-stage breast cancer:

> *They were looking on the right breast [in the ultrasound], but they were looking in the wrong area … Then they [looked] where the lump was. What they were concerned with was not what I found. So, the mammogram did not find what I found.*

> *My experience with early-stage breast cancer was a delay in diagnosis, which led to a later stage of, finally, discovery. I was a stage 3C. So I'm not happy about that. And I've been working very hard with [cancer organizations] on suspicion to decision [the time from finding a lump to deciding on a treatment plan], studying up a better pathway, standardising procedures and amalgamating so that everything is done as a one stop shop.*

Work–Life Impacts

Some people described the impact that the diagnosis of early-stage breast cancer had on their careers. The time it took to go through different tests, biopsy, and surgery, made it challenging for some people to return to work for a long period time, which caused additional stress:

> *Because I'm self-employed, I don't have any insurance for sick leave, that kind of thing. So, basically, if I don't work, money doesn't come in. I would say that the major stress I [went] through at that time [was] finances. … Because I don't have any kind of support. I ended up having to go for the [retirement savings plan] to support myself to go through this. I mean, I could not go back to work. I was having the operation and chemotherapy.*

Emotional Impacts

Patients and family members were able to highlight the significant impact that early-stage breast cancer had on their quality of life. Both patients and family members reported on the emotional burden of being diagnosed with breast cancer, especially where they did not have a family history of breast cancer. Upon diagnosis, some patients described how hard it was for them and their family members to accept the news and manage its impact on their day-to-day lives:

> *The experience in dealing with my family was almost as difficult as the diagnosis. Number one, my husband. His mother died of breast cancer.*

> *In terms of quality of life, it did not have an immediate impact until a few months after [my mother's] tumorectomy, when they realized that she actually had a more extensive tumour size than they had originally anticipated from the tumorectomy analysis. So that's when it had a higher impact on her overall mental state. Also, my mom had to make the choice of whether she was going to have a second tumorectomy or would opt to have a mastectomy. So that led to more physical quality of life changes.*

> *Actually it didn't change my day-to-day life much, except of course when I had the radiation. That was several days in a row… my whole life centred around the cancer.*

Overview of Currently Available Treatments

In Ontario, patients who did not receive a GEP test had to go through other types of testing and determine their treatment based on the results. This included regular pathology testing, which would also consist of biopsy. These patients would be provided with a pathology report that would outline the size of the tumor, the nuclear grade, and whether the tumor is ER+ or ER−:

> *To determine my treatment, the biggest factor [was] the biopsy report … that's what they used to set up my treatment plan. That and the size of the tumour, all of the things in the pathology report: nuclear grade, tumour size, and ER positive or negative.*

> *In her case, it was based on whether or not her tumour sample had certain tumour markers, which determined whether or not she would be receptive to radiotherapy and chemotherapy, so she went through regular pathology testing.*

Intervention Under Study: Gene Expression Profiling Test

Of the seven patients interviewed, three received a GEP test—Oncotype DX, which is the most commonly performed GEP test in Ontario. The other four patients expressed an interest in receiving one.

Process of Receiving a Gene Expression Profiling Test

After going through an initial round of testing, patients reported that they took it upon themselves to research GEP tests. With the information they had gathered, they would ask their provider (oncologist or surgeon) to apply for a test. Their eligibility to receive the test is based on the results of their pathology tests. The process requires that the patient or their provider fill out documents that are then sent to a lab outside of Canada along with the pathology test samples:

> *[My care provider] made an appointment with an oncologist. In the meantime, I did some reading and I learned about the [GEP test and thought that] I would be a good candidate. So when I went in to see the [doctor], I was ready to discuss that … I would like this test. And he said, that's good, because I also agree, you should have that test.*

> *It had a bunch of forms to fill out, mainly for the pathologist … then the surgeon signed and it was then sent [outside of Canada]. I got an email in which my score [was provided].*

One patient also highlighted that the result from the lumpectomy can help establish patient eligibility to receive a GEP test:

> *It was the result of the lumpectomy that determined my eligibility for [the GEP test] and that was the only test.*

Interpretation of Results and Decision-Making

After the test is processed, the patient is presented with a score. Depending on a person's age and lymph node status, Oncotype DX test results fall within either two (low or high) or three (low, intermediate, or high) risk categories. The score is used to inform decision-making around whether or not to proceed with adjuvant chemotherapy. The three patients who did have a test received a low score, which helped them to decide to forgo chemotherapy:

> *I had the test and I can't remember the number I came out with—13 or something—which meant that I was in a relatively low-risk [category]. The chemo would not likely be a good choice, so I did not have chemo. I did have radiation.*

One patient reported that, since it was hard to get an appointment with the oncologist, they had the option to meet a patient navigator—a service provided by the hospital she was attending. The navigator helped the patient receive the information faster and to be well informed before meeting the oncologist to go over the patient's risk:

> *I was really anxious waiting for it, so I didn't want to wait for the appointment with the oncologist to hear, it was hard to schedule an appointment with the oncologist so I had the patient navigator get the information for me and tell me the information and then when I met with the oncologist he went over my risk ... she was able to facilitate appointments for me and answer questions that I could not figure out myself. Not everybody has this option.*

Patients also pointed out that their providers valued their decision and opinion on whether to continue with chemotherapy or not based on the results of the GEP test:

> *Once I saw the likelihood that chemo would help, you know, it was very low. So that was the end of that … And my oncologist certainly agreed. I didn't have to fight him over it or anything, it was just the right decision for me. And I'm not sorry about making that decision. I might feel different if I get cancer again, but, at the moment, no.*

Perceived Benefits to Receiving Gene Expression Profiling Test

Better Decision-Making

Patients stated that the main benefit of receiving the GEP test is that it provides them with information on whether to go ahead with chemotherapy. A GEP test also helps patients learn about their diagnosis and the most appropriate treatment for them:

> *If I had had any information that would indicate I was at a low risk and did not require chemotherapy, I absolutely would have followed that pathway.*

> *I'm biased given my training background. Given my background, I would say it would be very useful if most, if not all, patients … could have access to gene profiling. It can help improve diagnosis, it can help refine prognosis, it can help better orient treatment decisions. It can help in risk stratification.*

> *I had confidence in it and because of that I was able to make a decision. With [the GEP test] saying I didn't need chemo, I trusted the [GEP test].*

> *Prognosis is very important, are you about to go back to work and be productive again? Or wait to die?*

Positive Emotional Impact

Another patient stated that the main benefit of receiving a GEP test is to reduce the uncertainty and anxiety that they had prior to the test, as the diagnosis of early-stage breast cancer already had a significant impact on their emotional well being:

> *The benefit was huge because it removed uncertainty and gave me some peace of mind. I had grade 2 cancer … so, you know, aggressive. But you don't know if you should have chemo or not, the test looks at your own genes and gives you information that can allow you to avoid going through chemo. Otherwise I was just going to rely on my oncologist's opinion. If there was no [GEP test], it's either chemo yes or chemo no. I'm feeling very, very lucky that I did not have to go through chemo.*

Interpretation of Gene Expression Profiling Test Results

While some patients found it easy to understand the results and took it upon themselves to research the information to learn more, others had a hard time understanding their results because there was little descriptive information to help them understand the numbers. Participants reported that discussions with their provider helped them understand the results and make a decision:

> *It doesn't tell you very much. I call it the kindergarten sheet. It would be helpful to have something else that would give you a little bit more information about what that all means.*

> *I think a lot of times people hear, well there's 90% chance of this, or a 2% chance of that, and they don't realize that that's just chance. So, they really don't explain statistics to people and somewhere along the line, we need to know that.*

> *I thought it was relatively easy to understand, but I was not quite sure … because of course it's all in terms of percentages—the odds [for] and the odds against, whether it's going to help or not. And that's why it's not completely clear.*

One patient stated that the surgeon did not provide her with much information to help make the decision and she felt confused as to what decision she would like to make:

> *I was told by the surgeon [that] this is what you will do, not this is what you might do or this is what's suggested. It was, 'this is what will happen.'*

Eligibility Criteria for Gene Expression Profiling Test Results

Gene expression profiling tests are typically used in people with early-stage breast cancer (stages 1 and 2). Some patients reported that they were not eligible for a GEP test because they had stage 3 breast cancer (a more advanced stage of cancer). One patient whose cancer did not qualify her for a GEP test said that if a GEP test were beneficial for her condition, it would have helped her decide what pathway of care to follow:

> *If it were applicable to me … if I had any information that would indicate I was at a low risk and did not require chemotherapy, I absolutely would have followed that pathway.*

Perceived Barriers to Receiving Gene Expression Profiling Test

The two main barriers that were identified by patients and caregivers were cost when required to pay for the test out of pocket and the limited information available to people not fluent in English.

Perceived Financial Barrier

Two of the three people who received a GEP test before it was publicly funded in Ontario through the out-of-country program paid for the test out of pocket. Most patients and caregivers reported that gene expression is an expensive test and that paying for it might be a challenge for those with limited income:

> *OHIP doesn't pay for it … I was fortunate enough that I could pay for the test … it was not a big deal for me … It was not something I had to worry about, you know, was I going to be able to manage groceries after that. So, I was okay, but I'm sure there are many people that, you know, that was a, that's a barrier.*
>
> *I had to pay for the test myself … $4,000 US … of course the cost is a barrier.*

Perceived Language Barrier

One family member pointed out that the patient is French speaking and does not have access to staff that can communicate in French. It can be a challenge for the patient to understand the information they are receiving and to learn about the types of treatment they may receive. The patient had to rely on a family member for interpretation and to communicate any other information to the medical staff:

> *We're French Ontarians, so there's a strong language barrier. It's very challenging to have the full spectrum of your staff be French speaking. So I would say the nurse that was looking after my mom was relatively bilingual, but a lot of the specialists—the pathologist, the radiotherapist, the oncologist … only spoke English ... at a very high level. In this context [my mother] was really relying on my interpretation [of what the specialists told me].*

Discussion

Evidence provided through the quantitative literature survey and direct patient engagement highlighted that patients valued the use of GEP testing to provide guidance when making treatment decisions around adjuvant chemotherapy. Patients preferred to be part of the decision-making process and the test acted as a trusted tool to help that decision. Furthermore, patients also highlighted that the results helped reduce the uncertainty and anxiety that they experienced before they received the test results. In contrast, patients who did not receive a GEP test based their treatment decisions on clinical factors and the results of their pathology tests. These patients reported feeling that a GEP test could also help them to make decisions about chemotherapy.

Our findings are limited in both our direct patient engagement and in the quantitative evidence review by the fact that a majority of the evidence is on Oncotype DX. All patients interviewed who had received a GEP test, had the Oncotype DX test.

To complement the quantitative evidence survey and our direct patient engagement, CADTH published a qualitative evidence review that evaluated patients' and providers' expectations, understanding, communication, experiences, perceptions, and decisions surrounding GEP tests.[198] In general, the key findings aligned with our results. The CADTH review also found that GEP tests are a valued chemotherapy treatment decision-making tool for both patients and providers, with many people heavily relying on the test results. The review by CADTH found that patients expected GEP tests to provide valid, individualized results to determine the appropriate course of treatment, and while low- or high-risk test results met these expectations, intermediate results did not. Our quantitative evidence survey also found that some patients thought intermediate test results were the least useful and led to more anxiety compared with low- or high-risk results. Similarly, both reviews found that not all patients understood the nature of testing and the possibility of inaccurate risk classification.

The CADTH review also explored provider communication around GEP testing. They noted that knowing patients' eligibility and preferences for chemotherapy was seen by many providers as the key to deciding whether or not to order GEP testing. In addition, the CADTH review found that communicating GEP test results requires longer patient consultation time (particularly for intermediate test results) and the timing of ordering GEP tests is important to ensure timely availability of test results for treatment decision-making.

Overall, both patients and providers value GEP testing for patients' treatment decision-making process. Our findings highlight the importance of GEP testing to the goal of enabling patients to be more involved with their treatment decision-making and to be better informed about their health condition.

Conclusions

People interviewed who have been diagnosed with early-stage breast cancer discussed the emotional and physical impact it had on their quality of life. They were able to highlight the process of receiving a GEP test, which enabled them to receive information about their condition and to guide them in making decisions around whether to accept or forgo chemotherapy. Our patient engagement and the quantitative evidence survey of patient and provider preferences and values found that patients were eager to be part of the decision-making process. Whether they received a test or not, patients felt that GEP testing is a valuable tool to help them become more educated about their condition and treatment options. It could relieve some of the uncertainty and anxiety they may experience. However, all patients we interviewed received Oncotype DX. They did not have any experience with other GEP tests, and therefore the evidence was mostly based on this test. Overall, patients receiving Oncotype DX reported that it was a positive experience. They would recommend that it be used in the patient's pathway of care.

CONCLUSIONS OF THE HEALTH TECHNOLOGY ASSESSMENT

In the LN− patient population, GEP tests are likely prognostic for freedom from distant recurrence (GRADE: Moderate) and may be prognostic for disease-free and overall survival (GRADE: Low). In the LN+ patient population, GEP tests may be prognostic for freedom from distant recurrence (GRADE: Low). They may also be prognostic for disease-free and overall survival (GRADE: Very Low), but we are very uncertain. Some GEP tests may predict chemotherapy benefit in the LN− population (GRADE: Low). They may also predict chemotherapy benefit in the LN+ population (GRADE: Very Low), but we are very uncertain about this. Gene expression profiling tests may lead to changes in treatment recommendations (GRADE: Low). The GEP tests may also increase physician confidence in treatment recommendations (GRADE: Very Low), but we are very uncertain.

Gene expression profiling tests were likely cost-effective compared with usual care in people with ER+, LN−, HER2− breast cancer. All GEP tests except Oncotype DX were likely cost-effective in people with LN+ breast cancer. All GEP tests were likely cost-effective in people who are premenopausal. We are uncertain about the cost-effectiveness of GEP tests compared with each other. Publicly funding GEP tests conducted in Ontario is estimated to cost an additional $1.29 million to $2.22 million per year compared with funding GEP tests conducted through Ontario’s out-of-country program.

The quantitative evidence on patient and provider preferences and values and our interviews with patients and family members highlighted patients’ willingness to be part of the decision-making process. Whether they received a test or not, patients felt that a GEP test is a valuable tool to help them become more educated about their condition and treatment options. It could relieve some of the decisional uncertainty and anxiety they experience. Overall, patients receiving a GEP test reported that it was a positive experience. They would recommend it as a test to be used in the patient’s pathway of care.

ABBREVIATIONS

AOL	Adjuvant! Online
CADTH	Canadian Agency for Drugs and Technologies in Health
CCO	Cancer Care Ontario[d]
CI	Confidence interval
EPclin	EndoPredict clinical score
ER	Estrogen receptor
GEP	Gene expression profiling
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
ICER	Incremental cost-effectiveness ratio
IHC	Immunohistochemistry
LN	Lymph node
NICE	National Institute for Health and Care Excellence
NPI	Nottingham Prognostic Index
OR	Odds ratio
PR	Progesterone receptor
QALY	Quality-adjusted life-year
RCT	Randomized controlled trial
ROR	Prosigna Risk of Recurrence
RS	Oncotype DX Recurrence Score

[d] Cancer Care Ontario is now Ontario Health (Cancer Care Ontario).

GLOSSARY

Adverse event	An adverse event is any unexpected problem that happens during or as a result of treatment, regardless of the cause or severity.
Budget impact analysis	A budget impact analysis estimates the financial impact of adopting a new health care intervention on the current budget (i.e., the affordability of the new intervention). It is based on predictions of how changes in the intervention mix will impact the level of health care spending for a specific population. Budget impact analyses are typically conducted for a short-term period (e.g., 5 years). The budget impact, sometimes referred to as the net budget impact, is the estimated cost difference between the current scenario (i.e., the anticipated amount of spending for a specific population without using the new intervention) and the new scenario (i.e., the anticipated amount of spending for a specific population following the introduction of the new intervention).
Cost-effective	A health care intervention is considered cost-effective when it provides additional benefits, compared with relevant alternatives, at an additional cost that is acceptable to a decision-maker based on the maximum willingness-to-pay value.
Cost-effectiveness acceptability curve	In economic evaluations, a cost-effectiveness acceptability curve is a graphical representation of the results of a probabilistic sensitivity analysis. It illustrates the probability of health care interventions being cost-effective over a range of willingness-to-pay values. Willingness-to-pay values are plotted on the horizontal axis of the graph, and the probability of the intervention of interest and its comparator(s) being cost-effective at corresponding willingness-to-pay values is plotted on the vertical axis.
Cost-effectiveness acceptability frontier	In economic evaluations, a cost-effectiveness acceptability frontier is a graph summarizing the probability of a number of health care interventions being cost-effective over a range of willingness-to-pay values. Like cost-effectiveness acceptability curves, cost-effectiveness acceptability frontiers plot willingness-to-pay values on the horizontal axis and the probability of the interventions being cost-effective at particular willingness-to-pay values on the vertical axis.
Cost-effectiveness analysis	Used broadly, “cost-effectiveness analysis” may refer to an economic evaluation used to compare the benefits of two or more health care interventions with their costs. It may encompass several types of analysis (e.g., cost-effectiveness analysis, cost–utility analysis). Used more specifically, “cost-effectiveness analysis” may refer to a type of economic evaluation in which the main outcome measure is the incremental cost per natural unit of health (e.g., life-year, symptom-free day) gained.

Cost–utility analysis	A cost–utility analysis is a type of economic evaluation used to compare the benefits of two or more health care interventions with their costs. The benefits are measured using quality-adjusted life-years, which capture both the quality and quantity of life. In a cost–utility analysis, the main outcome measure is the incremental cost per quality-adjusted life-year gained.
Decision tree	A decision tree is a type of economic model used to assess the costs and benefits of two or more alternative health care interventions. Each intervention may be associated with different outcomes, which are represented by distinct branches in the tree. Each outcome may have a different probability of occurring and may lead to different costs and benefits.
Deterministic sensitivity analysis	Deterministic sensitivity analysis is an approach used to explore uncertainty in the results of an economic evaluation by varying parameter values to observe the potential impact on the cost-effectiveness of the health care intervention of interest. One-way sensitivity analysis accounts for uncertainty in parameter values one at a time, whereas multiway sensitivity analysis accounts for uncertainty in a combination of parameter values simultaneously.
Distant recurrence	Recurrence of the cancer after surgery in areas of the body away from the breast area where it originated.
Dominant	A health care intervention is considered dominant when it is more effective and less costly than its comparator(s).
Dominated	A health care intervention is considered dominated when it is less effective and more costly than its comparator(s).
EuroQol–Five Dimensions (EQ-5D)	The EQ-5D is a generic health-related quality-of-life classification system widely used in clinical studies. In economic evaluations, it is used as an indirect method of obtaining health state preferences (i.e., utility values). The EQ-5D questionnaire consists of five questions relating to different domains of quality of life: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. For each domain, there are three response options: no problems, some problems, or severe problems. A newer instrument, the EQ-5D-5L, includes five response options for each domain. A scoring table is used to convert EQ-5D scores to utility values.
Health-related quality of life	Health-related quality of life is a measure of the impact of a health care intervention on a person's health. It includes the dimensions of physiology, function, social life, cognition, emotions, sleep and rest, energy and vitality, health perception, and general life satisfaction.
Health state	A health state is a particular status of health (e.g., sick, well, dead). A health state is associated with some amount of benefit and may be associated with specific costs. Benefit is captured through individual or societal preferences for the time spent in each health state and is expressed in quality-adjusted weights called utility values. In a Markov model, a finite number of mutually exclusive health states are used to represent discrete states of health.

Incremental cost	The incremental cost is the additional cost, typically per person, of a health care intervention versus a comparator.
Incremental cost-effectiveness ratio (ICER)	The incremental cost-effectiveness ratio (ICER) is a summary measure that indicates, for a given health care intervention, how much more a health care consumer must pay to get an additional unit of benefit relative to an alternative intervention. It is obtained by dividing the incremental cost by the incremental effectiveness. Incremental cost-effectiveness ratios are typically presented as the cost per life-year gained or the cost per quality-adjusted life-year gained.
Markov model	A Markov model is a type of decision-analytic model used in economic evaluations to estimate the costs and health outcomes (e.g., quality-adjusted life-years gained) associated with using a particular health care intervention. Markov models are useful for clinical problems that involve events of interest that may recur over time (e.g., stroke). A Markov model consists of mutually exclusive, exhaustive health states. Patients remain in a given health state for a certain period of time before moving to another health state based on transition probabilities. The health states and events modelled may be associated with specific costs and health outcomes.
Ministry of Health perspective	The perspective adopted in economic evaluations determines the types of costs and health benefits to include. Ontario Health (Quality) develops health technology assessment reports from the perspective of the Ontario Ministry of Health. This perspective includes all costs and health benefits attributable to the Ministry of Health, such as treatment costs (e.g., drugs, administration, monitoring, hospital stays) and costs associated with managing adverse events caused by treatments. This perspective does not include out-of-pocket costs incurred by patients related to obtaining care (e.g., transportation) or loss of productivity (e.g., absenteeism).
One-way sensitivity analysis	A one-way sensitivity analysis is used to explore uncertainty in the results of an economic evaluation. It is done by varying one model input (i.e., a parameter) at a time between its minimum and maximum values to observe the potential impact on the cost-effectiveness of the health care intervention of interest.
Predictive ability	The degree to which GEP tests can identify people who will benefit most from chemotherapy.
Prognostic ability	The degree to which GEP tests can accurately predict the risk of an outcome and discriminate people with different outcomes.
Quality-adjusted life-year (QALY)	The quality-adjusted life-year (QALY) is a generic health outcome measure commonly used in cost–utility analyses to reflect the quantity and quality of life-years lived. The life-years lived are adjusted for quality of life using individual or societal preferences (i.e., utility values) for being in a particular health state. One year of perfect health is represented by one quality-adjusted life-year.

Reference case	The reference case is a preferred set of methods and principles that provide the guidelines for economic evaluations. Its purpose is to standardize the approach of conducting and reporting economic evaluations, so that results can be compared across studies.
Scenario analysis	A scenario analysis is used to explore uncertainty in the results of an economic evaluation. It is done by observing the potential impact of different scenarios on the cost-effectiveness of a health care intervention. Scenario analyses include varying structural assumptions from the reference case.
Sensitivity analysis	Every economic evaluation contains some degree of uncertainty, and results can vary depending on the values taken by key parameters and the assumptions made. Sensitivity analysis allows these factors to be varied and shows the impact of these variations on the results of the evaluation. There are various types of sensitivity analysis, including deterministic, probabilistic, and scenario.
Time horizon	In economic evaluations, the time horizon is the time frame over which costs and benefits are examined and calculated. The relevant time horizon is chosen based on the nature of the disease and health care intervention being assessed, as well as the purpose of the analysis. For instance, a lifetime horizon would be chosen to capture the long-term health and cost consequences over a patient's lifetime.
Tornado diagram	In economic evaluations, a tornado diagram is used to determine which model parameters have the greatest influence on results. Tornado diagrams present the results of multiple one-way sensitivity analyses in a single graph.
Utility	A utility is a value that represents a person's preference for various health states. Typically, utility values are anchored at 0 (death) and 1 (perfect health). In some scoring systems, a negative utility value indicates a state of health valued as being worse than death. Utility values can be aggregated over time to derive quality-adjusted life-years, a common outcome measure in economic evaluations.
Willingness-to-pay value	A willingness-to-pay value is the monetary value a health care consumer is willing to pay for added health benefits. When conducting a cost–utility analysis, the willingness-to-pay value represents the cost a consumer is willing to pay for an additional quality-adjusted life-year. If the incremental cost-effectiveness ratio is less than the willingness-to-pay value, the health care intervention of interest is considered cost-effective. If the incremental cost-effectiveness ratio is more than the willingness-to-pay value, the intervention is considered not to be cost-effective.

APPENDICES

Appendix 1: Public Funding in Canada

Table A1: Public Funding Criteria for GEP Tests for Breast Cancer in Canada

Test	Province/Territory	Public Funding Eligibility Criteria
EndoPredict	Ontario	• Newly diagnosed with early invasive breast cancer, ER+, HER2−, LN− or N1mi, postmenopausal, fit to receive chemotherapy (receiving or intend to receive tamoxifen or an aromatase inhibitor [anastrazole, letrozole, exemestane]) • Tumour is ≤ 1 cm, tumour grade is 2 or 3 OR LN micrometastases are present • Results will be used to guide decisions about withholding adjuvant systemic chemotherapy
MammaPrint	None	Not applicable
Oncotype DX	Ontario	• Newly diagnosed with early invasive breast cancer, ER+, HER2−, LN− or N1mi, pre- or postmenopausal, fit to receive chemotherapy (receiving or intend to receive tamoxifen or an aromatase inhibitor [anastrazole, letrozole, exemestane]) • Tumour is ≤ 1 cm, tumour grade is 2 or 3 OR LN micrometastases are present • Results will be used to guide decisions about offering or withholding adjuvant systemic chemotherapy or withholding adjuvant systemic chemotherapy for N1mi disease
	Alberta[234]	• Grade 2 or 3 invasive breast cancer, early-stage resected LN− (including N0i+) or N1mi, patient is medically fit to receive adjuvant breast cancer chemotherapy • Exclusion criteria: people unwilling to consider or medically unfit to receive adjuvant chemotherapy, LN+ or HER2+ breast cancer, metastatic breast cancer, grade 1 invasive breast cancer

	British Columbia[235]	• ≤ 80 years old, fit to receive chemotherapy, ER/PR+, HER2−
		• For LN− population (pN0 or pN0i+):
		  ○ Any grade 1–2 and ≤ 40 years old
		  ○ Any grade 2 and pT1b or larger
		  ○ Any grade 3
		• For LN+ population (pN1mi only, 0.3–2mm micrometastases in 1 LN): Any grade

	Manitoba[a]	Publicly funded for specific people but no published criteria available
	Newfoundland and Labrador[a]	Publicly funded for specific people but no published criteria available
	New Brunswick[a]	Publicly funded for specific people but no published criteria available
	Nova Scotia[236]	• Newly diagnosed early-stage LN− breast cancer (stage 1 or 2 pN0 or pN0i+), ER/PR+, HER2−, plan to receive adjuvant endocrine therapy • Medical oncologist recommends considering adjuvant chemotherapy based on high risk features • Pros and cons of Oncotype DX testing were discussed and the person agrees to the test being ordered and will accept the results as informative in regard to the benefit, or lack thereof, of adjuvant chemotherapy
	Prince Edward Island[a]	Publicly funded for specific people but no published criteria available

	Quebec[37]	• Newly diagnosed with invasive breast cancer, ER+, HER2−, LN−, or micrometastasis only (N1mi) with 1 of the following characteristics:
		  ○ Stage pT1b AND histological grade 2 and weak hormone receptor expression, histological grade 2 and young age, nuclear or histological grade 3, or high proliferation index
		  ○ Stage pT1c AND histological grade 1 and weak hormone receptor expression, histological grade 1 and young age, histological grade 1 and high proliferation index, or histological grade 2 or 3
		  ○ Stage pT2 AND histological grade 1 or 2, histological grade 3 and PR ≥ 20%
		• Ineligibility criteria: postmenopausal people with classical lobular carcinoma and no unfavourable factors, people with adenoid cystic or tubular carcinoma, people > 80 years old, people who will not receive adjuvant hormone therapy (tamoxifen or aromatase inhibitors)

	Saskatchewan[a]	Publicly funded for specific people but no published criteria available
Prosigna	Ontario	• Newly diagnosed with early invasive breast cancer, ER+, HER2−, LN− or N1mi, postmenopausal, fit to receive chemotherapy (receiving or intend to receive tamoxifen or an aromatase inhibitor [anastrazole, letrozole, exemestane]) • Tumour is ≤ 1 cm, tumour grade is 2 or 3 OR LN micrometastases are present • Results will be used to guide decisions about withholding adjuvant systemic chemotherapy or withholding adjuvant systemic chemotherapy for N1mi disease
	Alberta[234]	• Grade 2 or 3 invasive breast cancer, early-stage resected LN− (including N0i+) or N1mi, patient is medically fit to receive adjuvant breast cancer chemotherapy • Exclusion criteria: people unwilling to consider or medically unfit to receive adjuvant chemotherapy, LN+ or HER2+ breast cancer, metastatic breast cancer, grade 1 invasive breast cancer

	British Columbia[235]	• ≤ 80 years old, fit to receive chemotherapy, ER/PR-positive, HER2−
		• For LN− population (pN0 or PN0i+):
		  ○ Any grade 1–2 and ≤ 40 years old
		  ○ Any grade 2 and pT1b or larger
		  ○ Any grade 3
		• For LN+ population (pN1mi only, 0.3–2mm micrometastases in 1 LN): Any grade

Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; i, isolated tumour cells; LN, lymph node; mi, micrometastasis; N, node; p, pathologic stage; PR, progesterone receptor; T, tumour.
[a]Based on information received from Genomic Health (March 1, 2019).

Appendix 2: Summary of Guideline Recommendations on Gene Expression Profiling Tests

Table A2: Guideline Recommendations on GEP Tests for Breast Cancer

<table>
<tr><th>Author, Year</th><th>Recommendation Excerpt (Verbatim)</th></tr>
<tr><td colspan="2">Canadian Guidelines[a]</td></tr>
<tr><td>Alberta Health Services, 2018[234]</td><td>Risk categories for LN− breast cancer
• Adverse prognostic factors:
○ Age < 35 years
○ HER2 overexpression (HER2+)
○ Presence of lymphovascular invasion
○ Grade 3
○ HR− disease
○ Genomic test score (e.g., Prosigna, Oncotype DX): higher risk
• Lower risk:
○ ≤ 2 cm, grade 1, with no other adverse prognostic factors
○ < 0.5 cm with any other feature
○ Genomic test score (e.g., Prosigna, Oncotype DX): lower risk
• Intermediate risk:
○ All other combinations of factors that do not fit into either the low- or high-risk criteria
○ Genomic test score (e.g., Prosigna, Oncotype DX): intermediate risk
• Higher risk:
○ > 1 cm with any 2 or more adverse prognostic factors
○ > 2 cm with any 1 or more adverse prognostic factors
○ > 3 cm ± adverse prognostic factors
○ Special considerations for HER2+ breast cancer
○ Genomic test score (e.g., Prosigna, Oncotype DX): higher risk
Genomic testing for systemic therapy decision-making
• Inclusion criteria:
○ Patient is medically fit to receive adjuvant breast cancer chemotherapy, and
○ Has early stage resected LN− (including N01+) or N1mi, and
○ Either grade 2 or grade 3 invasive breast cancer
• Exclusion criteria:
○ Patients unwilling to consider or are medically unfit to receive adjuvant breast cancer chemotherapy
○ LN+ breast cancer</td></tr>
</table>

Author, Year	Recommendation Excerpt (Verbatim)
	○ Metastatic breast cancer ○ HER2+ breast cancer ○ Grade 1 invasive breast cancer (e.g., Prosigna, Oncotype DX)
Cancer Care Ontario, 2016[30] (Chang et al, 2017[29])	• Clinicians may offer multigene profile assay testing to potential chemotherapy candidates with invasive breast carcinoma that is ER+, HER2−. (Recommendation type: evidence-based; evidence quality: level IB; strength of recommendation: moderate) • In patients with node-negative, ER+, HER2− disease, clinicians may use a low-risk result from the Oncotype DX, Prosigna, or EndoPredict assay to support a decision to withhold chemotherapy. (Recommendation type: evidence-based; evidence quality: IB; strength of recommendation: moderate) • In patients with node−, ER+, HER2− disease, clinicians may use a high-risk result from Oncotype DX to support a decision to offer chemotherapy. A high-risk Oncotype DX result in this subpopulation has been associated with both poor prognosis without chemotherapy and a predicted benefit from chemotherapy. (Recommendation type: evidence-based; evidence quality: IB–II; strength of recommendation: weak) • In some patients with ER+, HER2− tumours and with 1–3 involved nodes (N1a disease), clinicians may withhold chemotherapy based on a low-risk Oncotype DX or Prosigna score if the decision is supported by other clinical, pathology, or patient-related factors. (Recommendation type: consensus-based; evidence quality: level II; strength of recommendation: weak) • In patients with ER+ disease, the evidence is insufficient to recommend the use of multigene profiling assays to inform clinical decision-making for late Risk of Recurrence. A high-risk score using Prosigna or EndoPredict prognosticates for late recurrence; however, evidence that those tests predict a benefit for the use of extended adjuvant endocrine treatment beyond 5 years is lacking. (Recommendation type: consensus-based; evidence quality: lack of evidence; strength of recommendation: weak)
Cancer Care Ontario, 2014;[237] Eisen et al, 2015[238]	• The following risk stratification tools can be used in determining the utility of certain systemic therapies in patients with early-stage breast cancer: ○ Oncotype DX score (for HR+; N0, N1mic, or isolated tumour cell; and HER2− ○ Adjuvant! Online (http://www.adjuvantonline.com). • When considering LN− tumours greater than 5 mm in size, these features should be considered high-risk (with the patients therefore considered candidates for chemotherapy): ○ Grade 3 ○ Triple-negative (ER−, PR−, and HER2−) ○ Positive for lymphovascular invasion ○ Oncotype DX Recurrence Score associated with an estimated 15% or greater risk of distant relapse at 10 years ○ HER2+. • Adjuvant chemotherapy might not be required in patients with HER2−, strongly ER+, PR+ breast cancer with any of these additional characteristics: ○ LN+ with micrometastasis (< 2 mm) only, or ○ Tumour < 5 mm in size, or ○ Oncotype DX Recurrence Score with an estimated distant relapse risk of less than 15% at 10 years.

Author, Year	Recommendation Excerpt (Verbatim)
International Guidelines	
National Comprehensive Cancer Network (NCCN), 2019[33]	21-gene assay (Oncotype DX) for node-negative: • Prognostic: yes • Predictive: yes • NCCN category of preference: Preferred (interventions that are based on superior efficacy, safety, and evidence; and, when appropriate, affordability) • NCCN category of evidence and consensus: 1 (based upon high-level evidence, there is uniform NCCN consensus that the ntervention is appropriate) 21-gene assay (Oncotype DX) for node-positive: • Prognostic: yes • Predictive: Not available; awaiting results of RxPONDER study • NCCN category of preference: Other (other interventions that may be somewhat less efficacious, more toxic, or based on less mature data; or significantly less affordable for similar outcomes) • NCCN category of evidence and consensus: 2A (based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate) 70-gene assay (MammaPrint) for node-negative and 1–3 positive nodes: • Prognostic: yes • Predictive: not determined • NCCN category of preference: Other (other interventions that may be somewhat less efficacious, more toxic, or based on less mature data; or significantly less affordable for similar outcomes) • NCCN category of evidence and consensus: 1 (based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate) 50-gene assay (PAM 50) for node-negative and 1–3 positive nodes: • Prognostic: yes • Predictive: not determined • NCCN category of preference: Other (other interventions that may be somewhat less efficacious, more toxic, or based on less mature data; or significantly less affordable for similar outcomes) • NCCN category of evidence and consensus: 2A (based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate)

Author, Year	Recommendation Excerpt (Verbatim)
	12-gene assay (EndoPredict) for node-negative and 1–3 positive nodes: • Prognostic: yes • Predictive: not determined • NCCN category of preference: Other (other interventions that may be somewhat less efficacious, more toxic, or based on less mature data; or significantly less affordable for similar outcomes) • NCCN category of evidence and consensus: 2A (based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate) Note: Multigene assays provide prognostic and therapy-predictive information that complements T, N, M and biomarker information. Use of these assays is not required for staging. The 21-gene assay (Oncotype DX) is preferred by the NCCN Breast Cancer Panel for node-negative breast cancer. Other prognostic multigene assays can provide additional prognostic information in patients with 1–3 positive lymph nodes.

National Institute for Health and Care Excellence (NICE), 2018[36]	• EndoPredict (EPclin score), Oncotype DX Breast Recurrence Score and Prosigna are recommended as options for guiding adjuvant chemotherapy decisions for people with oestrogen receptor (ER)-positive, human epidermal growth factor receptor 2 (HER2)-negative and lymph node (LN)-negative (including micrometastatic disease) early breast cancer, only if:
	  ○ they have an intermediate risk of distant recurrence using a validated tool such as PREDICT or the Nottingham Prognostic Index
	  ○ information provided by the test would help them choose, with their clinician, whether or not to have adjuvant chemotherapy taking into account their preference
	  ○ the companies provide the tests to the NHS with the discounts agreed in the access proposals and
	  ○ clinicians and companies make timely, complete and linkable record-level test data available to the National Cancer Registration and Analysis Service as described in the data collection arrangements agreed with NICE.
	• MammaPrint is not recommended for guiding adjuvant chemotherapy decisions for people with ER-positive, HER2-negative and LN-negative early breast cancer because it is not cost effective.
	• IHC4+C is not recommended for guiding adjuvant chemotherapy decisions for people with ER-positive, HER2-negative and LN-negative early breast cancer because the analytical validity of the test is uncertain.

Spanish Society of Medical Oncology and the Spanish Society of Pathology, 2018[239]	• In node-negative ER-positive breast cancer patients, one of several available genetic prognostic platforms (MammaPrint, Oncotype DX, Prosigna or EndoPredict) may be used in order to establish a prognostic category and to discuss with the patient whether adjuvant treatment may be limited to hormonal therapy. Oncotype DX • 5-year recurrence risk prognosis: good evidence to support a recommendation for use in patients with a low Recurrence Score (evidence from ≥ 1 properly randomized, controlled trial); moderate evidence to support a recommendation for use in patients with other Recurrence Scores (evidence from ≥ 1 properly randomized, controlled trial) • 10-year recurrence risk prognosis: moderate evidence to support a recommendation for use (evidence from ≥ 1 properly randomized, controlled trial)

Author, Year	Recommendation Excerpt (Verbatim)
	• Chemotherapy benefit prediction: good evidence to support a recommendation for use in patients with a low Recurrence Score (evidence from ≥ 1 properly randomized, controlled trial); moderate evidence to support a recommendation for use in patients with other Recurrence Scores (evidence from ≥ 1 properly randomized, controlled trial) Prosigna • 5-year recurrence risk prognosis: moderate evidence to support a recommendation for use (evidence from ≥ 1 properly randomized, controlled trial) • 10-year recurrence risk prognosis: moderate evidence to support a recommendation for use (evidence from ≥ 1 properly randomized, controlled trial) • Chemotherapy benefit prediction: not available MammaPrint • 5-year recurrence risk prognosis: moderate evidence to support a recommendation for use (evidence from ≥ 1 properly randomized, controlled trial) • 10-year recurrence risk prognosis: not available • Chemotherapy benefit prediction: not available EndoPredict • 5-year recurrence risk prognosis: moderate evidence to support a recommendation for use (evidence from ≥ 1 properly randomized, controlled trial) • 10-year recurrence risk prognosis: moderate evidence to support a recommendation for use (evidence from ≥ 1 properly randomized, controlled trial) • Chemotherapy benefit prediction: not available
American Society of Clinical Oncology, 2019[240]	• If a patient has ER/PR+, HER2− (node-negative) breast cancer, the clinician may use the 21-gene RS (Oncotype DX; Genomic Health, Redwood, CA) to guide decisions for adjuvant systemic chemotherapy. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: strong) • For patients older than 50 years whose tumors have Oncotype DX RSs < 26 and for patients age 50 years or younger whose tumors have Oncotype DX RSs < 16, there is little to no benefit from chemotherapy. Clinicians may offer endocrine therapy alone. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: strong) • For patients 50 years of age or younger with Oncotype DX RSs of 16 to 25, clinicians may offer chemoendocrine therapy. (Recommendation type: evidence based; evidence quality: intermediate; strength of recommendation: moderate) • Patients with Oncotype DX RSs > 30 should be considered candidates for chemoendocrine therapy. (Recommendation type: evidence based; evidence quality: high; strength of recommendation: strong) • Based on expert panel consensus, oncologists may offer chemoendocrine therapy to patients with Oncotype DX scores of 26–30. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: moderate)

Author, Year	Recommendation Excerpt (Verbatim)
	• If a patient has ER/PR+, HER2− (node-positive) breast cancer, the clinician should not use the 21-gene RS (Oncotype DX; Genomic Health) to guide decisions for adjuvant systemic chemotherapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use the 21-gene RS (Oncotype DX; Genomic Health) to guide decisions for adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has ER/PR+, HER2− (node-negative) breast cancer, the clinician may use the 12-gene risk score (EndoPredict; Sividon Diagnostics, Köln, Germany) to guide decisions for adjuvant systemic chemotherapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has ER/PR+, HER2− (node-positive) breast cancer, the clinician should not use the 12-gene risk score (EndoPredict; Sividon Diagnostics) to guide decisions for adjuvant systemic chemotherapy. (Recommendation type: evidence-based; evidence quality: insufficient; strength of recommendation: moderate) • If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use 12-gene risk score (EndoPredict; Sividon Diagnostics) to guide decisions for adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has ER/PR+, HER2−, node-negative, breast cancer, the MammaPrint (Agendia, Irvine, CA) assay may be used in those with high clinical risk per MINDACT categorization to inform decisions on withholding adjuvant systemic chemotherapy due to its ability to identify a good-prognosis population with potentially limited chemotherapy benefit. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: strong) • If a patient has ER/PR+, HER2−, node-negative, breast cancer, the MammaPrint (Agendia) assay should not be used in those with low clinical risk per MINDACT categorization to inform decisions on withholding adjuvant systemic chemotherapy, because women in the low clinical risk category had excellent outcomes and did not appear to benefit from chemotherapy even with a genomic high-risk cancer. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: strong) • If a patient has ER/PR+, HER2−, node-positive, breast cancer, the MammaPrint (Agendia) assay may be used in patients with one to three positive nodes and at high clinical risk per MINDACT categorization to inform decisions on withholding adjuvant systemic chemotherapy due to its ability to identify a good prognosis population with potentially limited chemotherapy benefit. However, such patients should be informed that a benefit of chemotherapy cannot be excluded, particularly in patients with greater than one involved lymph node. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: moderate) • If a patient has ER/PR+, HER2−, node-positive, breast cancer, the MammaPrint (Agendia) assay should not be used in patients with one to three positive nodes and at low clinical risk per MINDACT categorization to inform decisions on withholding adjuvant systemic chemotherapy. There are insufficient data on the clinical utility of MammaPrint in this specific patient population. (Recommendation type: informal consensus; evidence quality: low; strength of recommendation: moderate) • If a patient has HER2+ breast cancer, the clinician should not use the MammaPrint (Agendia) assay to guide decisions on adjuvant systemic therapy. Additional studies are required to address the role of MammaPrint in patients with this tumour subtype who are also receiving HER2-targeted therapy. (Recommendation type: informal consensus; evidence quality: low; strength of recommendation: moderate) • If a patient has ER/PR− and HER2-negative breast cancer (triple-negative), the clinician should not use the MammaPrint (Agendia) assay to guide decisions on adjuvant systemic chemotherapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong)

Author, Year	Recommendation Excerpt (Verbatim)
	• If a patient has ER/PR+, HER2− (node-negative) breast cancer, the clinician may use the PAM50 ROR score (Prosigna Breast Cancer Prognostic Gene Signature Assay; NanoString Technologies, Seattle, WA) in conjunction with other clinicopathologic variables to guide decisions on adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: strong) • If a patient has ER/PR+, HER2− (node-positive) breast cancer, the clinician should not use the PAM50 ROR score (Prosigna Breast Cancer Prognostic Gene Signature Assay; NanoString Technologies) to guide decisions on adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has HER2+ breast cancer, the clinician should not use the PAM50 ROR score (Prosigna Breast Cancer Prognostic Gene Signature Assay; NanoString Technologies) to guide decisions on adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has triple-negative breast cancer, the clinician should not use the PAM50 ROR score (Prosigna Breast Cancer Prognostic Gene Signature Assay; NanoString Technologies) to guide decisions for adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has ER/PR+, HER2−, node-negative breast cancer, the clinician may use the Breast Cancer Index (bioTheranostics, San Diego, CA) to guide decisions on adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has ER/PR+, HER2−, node-positive breast cancer, the clinician should not use the Breast Cancer Index (bioTheranostics) to guide decisions on adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use the Breast Cancer Index (bioTheranostics) to guide decisions on adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has ER/PR+, HER2− (node-positive or node-negative) breast cancer, the clinician should not use the five-protein assay Mammostrat (GE Healthcare, Aliso Viejo, CA) to guide decisions on adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use the five-protein assay Mammostrat (GE Healthcare) to guide decisions on adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has ER/PR+, HER2− (node-positive or node-negative) breast cancer, the clinician should not use IHC-4 to guide decisions on adjuvant systemic chemotherapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use IHC-4 to guide decisions on adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has ER/PR+, HER2− (node-negative) breast cancer, the clinician may use the uPA and PAI-1 to guide decisions on adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: high; strength of recommendation: weak)

Author, Year	Recommendation Excerpt (Verbatim)
	• If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use the uPA and PAI-1 to guide decisions on adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: weak) • The clinician should not use circulating tumour cells to guide decisions for adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: strong) • If a patient has ER/PR+, HER2− (node-positive or node-negative) breast cancer, the clinician should not use tumour-infiltrating lymphocytes to guide decisions for adjuvant systemic therapy. (Recommendation type: informal consensus; evidence quality: insufficient; strength of recommendation: strong) • If a patient has HER2+ breast cancer or triple-negative breast cancer, the clinician should not use tumour-infiltrating lymphocytes to guide decisions on adjuvant systemic therapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: strong) • Ki67 labeling index by immunohistochemistry should not be used to guide the choice of adjuvant chemotherapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate) • If a patient has ER/PR+, HER2− (node-negative) breast cancer and has had 5 years of endocrine therapy without evidence of recurrence, the clinician should not use multiparameter gene expression or protein assays (Oncotype DX, EndoPredict, PAM50, Breast Cancer Index, or IHC-4) to guide decisions on extended endocrine therapy. (Recommendation type: evidence-based; evidence quality: intermediate; strength of recommendation: moderate)
St. Gallen International Consensus Guidelines, 2019[34]	The Panel strongly endorsed the value of genomic assays for determining whether to recommend chemotherapy in T1/T2 N0 tumors, T3 N0 tumors, and TxN1 (1–3 positive LN).
European Group on Tumor Markers, 2017[241]	Oncotype DX recommendation: • Oncotype DX RS may provide added value to established factors for determining prognosis and aiding decision-making with respect to administration of adjuvant chemotherapy in newly diagnosed breast cancer patients with LN− invasive disease that is ER+ but HER2− (Level of evidence: IB; strength of recommendation: A). In addition, Oncotype DX may be considered for identifying HER2−, ER+ patients with 1–3 involved lymph nodes for treatment with adjuvant chemotherapy (Level of evidence: IB; strength of recommendation: A). • Before performing the test, any biopsy cavity in the cancer specimen should be removed by manual dissection. Oncotype DX recommendation for further research: • Two of the most important questions relating to the use of Oncotype DX are currently being addressed in prospective randomized trials, i.e. whether LN− ER+ patients with intermediate RS benefit from adding adjuvant chemotherapy to endocrine therapy (TAILORx trial) and whether LN+ (1–3 nodes positive), ER+ patients with low to intermediate RS benefit from adjuvant chemotherapy (RxPONDER trial). In the RxPONDER trial, women with 1–3 positive LNs who have HR+ but HER2− disease with RS ≤ 25 are randomized to receive endocrine therapy alone or endocrine therapy plus chemotherapy. • Establish if Oncotype DX can predict response to specific forms of adjuvant chemotherapy. MammaPrint recommendation:

Author, Year	Recommendation Excerpt (Verbatim)
	• MammaPrint may be used for determining prognosis and guiding decision-making with respect to the administration of adjuvant chemotherapy in patients with newly diagnosed invasive breast cancer that is LN− or LN+ (1–3 metastatic nodes). Patients at high-risk based on clinical and pathological criteria but at low risk based on MammaPrint may be the candidates for avoiding having to receive adjuvant chemotherapy (Level of evidence: IA; strength of recommendation: A). MammaPrint recommendation for further research: • Further validation after longer follow-up. • Investigate if MammaPrint can predict response to specific forms of systemic treatment. Prosigna recommendation: • In combination with established clinical and pathological factors, Prosigna may be used for predicting outcome and aiding adjuvant therapy decision-making in HR+, HER2− patients that are either LN− or LN+ (1–3 metastatic nodes) (Level of evidence: IB; strength of recommendation: A). Prosigna recommendation for further research: • Validation in a prospective randomized trial. This is currently ongoing as part of the OPTIMA trial (ISRCTN42400492). • Establish if Prosigna can predict benefit from adjuvant chemotherapy. • Further validation for predicting late recurrences following adjuvant endocrine therapy. • Further validation in premenopausal patients. EndoPredict recommendation: • In combination with established clinical and pathological factors, EndoPredict may be used for predicting outcome and aiding adjuvant therapy decision-making in HR+, HER2− patients that are either LN− or LN+ (1–3 metastatic nodes) (Level of evidence: IB; strength of recommendation: A). EndoPredict: recommendation for further research • Validation in a prospective randomized trial. This is currently ongoing as part of the UNIRAD trial (NCT01805271). • Establish if EndoPredict can predict benefit from adjuvant chemotherapy. • Further validation for predicting late recurrences following adjuvant endocrine therapy. • Further validation in premenopausal patients. Breast Cancer Index recommendation: • In combination with established clinical and pathological factors, BCI may be used for predicting outcome and aiding adjuvant therapy decision-making in LN−, HR+ and HER2− patients (Level of evidence: IB; strength of recommendation: A).

Author, Year	Recommendation Excerpt (Verbatim)
	Breast Cancer Index: recommendation for further research • Validation in patients with LN+ disease. • Validation in a prospective randomized trial. • Further validation for predicting late recurrences following adjuvant endocrine therapy.
European Society for Medical Oncology, 2019[35]	• Validated gene expression profiles may be used to gain additional prognostic and/or predictive information to complement pathology assessment and help in adjuvant chemotherapy decision making (level of evidence I: evidence from at least one large randomised controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity; grade of recommendation A: strong evidence for efficacy with a substantial clinical benefit, strongly recommended) • The choice of treatment strategy should be based on the tumour burden/location (size and location of primary tumour, number of lesions, extent of lymph node involvement) and biology (pathology, including biomarkers and gene expression), as well as the age, menopausal status, general health status and preferences of the patient (level of evidence V: studies without control group, case reports, expert opinions; grade of recommendation A: strong evidence for efficacy with a substantial clinical benefit, strongly recommended) • In cases of uncertainty regarding indications for adjuvant chemotherapy (after consideration of all clinical and pathological factors), expression of uPA-PAI1 [level of evidence I: evidence from at least one large randomised, controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity; grade of recommendation A: strong evidence for efficacy with a substantial clinical benefit, strongly recommended] or gene expression assays, such as MammaPrint [level of evidence I: evidence from at least one large randomised controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity; grade of recommendation A: strong evidence for efficacy with a substantial clinical benefit, strongly recommended], Oncotype DX [level of evidence I: evidence from at least one large randomised, controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity; grade of recommendation A: strong evidence for efficacy with a substantial clinical benefit, strongly recommended], Prosigna, EndoPredict or Breast Cancer Index, can be used • Expression of uPA-PAI1 or multigene panels, such as MammaPrint, Oncotype DX, EndoPredict, Prosigna, or Breast Cancer Index, may be used in conjunction with all clinicopathological factors to guide systemic treatment decisions in patients where these decisions are challenging, such as luminal B-like/HER2-negative and node-negative/nodes 1–3-positive breast cancer [level of evidence I: evidence from at least one large randomised controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity; grade of recommendation A: strong evidence for efficacy with a substantial clinical benefit, strongly recommended]

Abbreviations: EPclin, EndoPredict clinical score; ER, estrogen receptor; GEP, gene expression profiling; HER2, human growth factor receptor 2; HR, hormone receptor; IHC-4, immunohistochemistry 4; IHC4+C: immunohistochemistry 4 plus clinical treatment score; LN, lymph node; M, metastasis; mi, micrometastasis; mic, micrometastasis; N, node; NCCN, National Comprehensive Cancer Network; NHS, National Health Service; PAI-1, plasminogen activator inhibitor type 1; PR, progesterone receptor; ROR, Risk of Recurrence; RS, Recurrence Score; T, tumour; uPA, urokinase plasminogen activator.
[a]Guidelines are reproduced verbatim from the listed sources.

Table A3: Study Categories and Levels of Evidence Based on the Tumour Marker Utility Grading System[a]

Descriptor	Definition
Study Categories	
A	Randomized controlled trial designed with tumour biomarker or biomarker assay as the intervention
B	Randomized controlled trial designed to address a treatment intervention that is not the tumour biomarker or biomarker assay; study prospectively enrols and follows patients and collects tumour samples, and then uses archived tumour tissue retrospectively to evaluate the tumour biomarker or biomarker assay
C	Prospective observational registry study that prospectively enrols patients in a registry and collects, processes, and archives tumour specimens, but that uses standard-of-care treatment and follow-up; archived tumour tissue is used retrospectively to evaluate the tumour biomarker or biomarker assay
D	Retrospective study
Levels of Evidence	
IA	1 category A study
IB	At least 2 category B studies with consistent results
II	1 category B study, *or* multiple category B studies with inconsistent results, *or* at least 2 category C studies with consistent results
III	1 category C study *or* multiple category C studies with inconsistent results
IV	Any number of category D studies (level IV evidence is not sufficient for determining clinical utility)

[a]Used by Cancer Care Ontario guidelines[29,30]

Table A4: Elements of Tumour Marker Studies That Constitute Levels of Evidence Determination[a]

Category Element	A Prospective	B Prospective Using Archived Samples	C Prospective/Observational	D Retrospective/Observational
Clinical trial	PCT designed to address tumour marker	Prospective trial not designed to address tumour marker, but design accommodates tumour marker utility Accommodation of predictive marker requires PRCT	Prospective observational registry; treatment and follow-up not indicated	No prospective aspect to study
Patients and patient data	Prospectively enrolled, treated, and followed in PCT	Prospectively enrolled, treated, and followed in clinical trial and, especially if a predictive utility is considered, a PRCT addressing the treatment of interest	Prospectively enrolled in registry, but treatment and follow-up standard of care	No prospective stipulation of treatment or follow-up; patient data collected by retrospective chart review
Specimen collection, processing, and archival	Specimens collected, processed, and assayed for specific marker in real time	Specimens collected, processed, and archived prospectively using generic SOPs; assayed after trial completion	Specimens collected, processed, and archived prospectively using generic SOPs; assayed after trial completion	Specimens collected, processed and archived with no prospective SOPs
Statistical design and analysis	Study powered to address tumour marker question	Study powered to address therapeutic question and underpowered to address tumor marker question Focused analysis plan for marker question developed before doing assays	Study not prospectively powered at all; retrospective study design confounded by selection of specimens for study Focused analysis plan for marker question developed before doing assays	Study not prospectively powered at all. Retrospective study design confounded by selection of specimens for study No focused analysis plan for marker question developed before doing assays
Validation	Result unlikely to be play of chance Although preferred, validation not required	Result more likely to be play of chance that A but less likely than C Requires one or more validation studies	Result very likely to be play of chance Requires subsequent validation studies	Result very likely to be play of chance Requires subsequent validation

Abbreviations: PCT, prospective controlled trial; PRCT, prospective randomized controlled trial; SOP, standard operating practices.
[a]Used by the European Group on Tumor Markers 2017 Guidelines[241]

Table A5: Revised Determination of Levels of Evidence Using Elements of Tumour Marker Studies[a]

Level of Evidence	Category	Validation Studies Available
I	A	None required
I	B	One or more with consistent results
II	B	None or inconsistent results
II	C	2 or more with consistent results
III	C	None or 1 with consistent results or inconsistent results
IV–V	D	Not applicable (studies will never be satisfactory for determination of medical utility)

[a]Used by the European Group on Tumor Markers 2017 Guidelines[241]

Appendix 3: Recent Health Technology Assessments and Systematic Reviews on GEP Tests for Breast Cancer

Table A6: Recent Health Technology Assessments and Systematic Reviews on Gene Expression Profiling Tests for Breast Cancer

Author, Year	Search Period	Databases	Test(s)	No. Included Studies	Primary Conclusions
Broder et al, 2013[242]	2004 to 2012 Conference abstracts: 2010 to 2012s	PubMed ASCO, Breast Cancer Symposium, and SABCS conference abstracts	BreastOncPx, MammaPrint, Mammostrat, Molecular Grade Index, Oncotype DX, EndoPredict, PAM50 (Prosigna), NuvoSelect, IHC4	20 studies	Published evidence of clinical utility for only Oncotype DX and MammaPrint Substantial evidence that Oncotype DX changes treatment decisions in about one-third of patients and reduces chemotherapy use by more than 20% Evidence from 3 studies that MammaPrint changes treatment decision-making, but not overall reduction in chemotherapy
Carlson et al, 2013[243]	MEDLINE, Embase and ASCO: Jun 2005 to Mar 2012 SABCS: Jan 2008 and Mar 2012	MEDLINE, Embase, ASCO and SABCS conference abstracts	Oncotype DX	23 studies	Low risk people more likely to follow suggested treatment than high-risk people, suggesting tendency toward less aggressive treatment Lack of studies on impact of Oncotype DX vs. standard approaches
Tiwana et al, 2013[168] University of Calgary	Dec 2012	MEDLINE, Embase, CINAHL, Cochrane library, Web of Science, BIOSIS	Oncotype DX, IHC4	14 studies	Limited, low-quality evidence supporting the clinical utility of Oncotype DX to predict benefit from chemotherapy Based on observational studies of low to medium quality, Oncotype DX results lead to changes in adjuvant chemotherapy decisions Oncologists and pathologists in Alberta have mixed opinions about the analytic utility of Oncotype DX, especially for patients in the intermediate risk group, and a lack of consensus about the communication and usability of the results obtained from IHC4 testing For patients, genetic testing may present complex information that may be hard to understand

Author, Year	Search Period	Databases	Test(s)	No. Included Studies	Primary Conclusions
Meleth et al, 2014[244] Agency for Healthcare Research and Quality (AHRQ)	Inception to Nov 2013	PubMed, Cochrane, Embase Manually searched for unpublished articles using test developers' websites, ClincialTrials.gov, FDA website, Health Services Research Projects in Progress, EU Clinical Trials Register; requested information from College of American Pathologists and relevant companies	MammaPrint, Oncotype DX	112 studies	Modest evidence supporting prognostic value (clinical validity) beyond traditional prognostic factors for MammaPrint and Oncotype DX Moderate strength of evidence that Oncotype DX leads to changes in treatment decisions resulting in less chemotherapy use No studies that directly assessed the impact of test use on downstream health outcomes to establish clinical utility; no evidence that using test was related to improved outcomes for patients
CADTH, 2014[245]	Jan 1, 2008 to Dec 18, 2013	PubMed, Cochrane library, CRD, HTA agencies, focused internet search	Oncotype DX	4 HTAs/SRs, 4 recent primary studies, 6 guidelines	Some benefit of Oncotype DX for prognosis and treatment planning Extent of benefit is unclear; differences in patient clinical outcomes as a result of treatment decision-making remain unknown Unclear recommended action for people at intermediate risk
CADTH, 2014[181]	Jan 1, 2008 to Dec 18, 2013	PubMed, Cochrane library, CRD, HTA agencies, focused internet search	Oncotype DX	1 HTA, 1 guideline	Clinical effectiveness is uncertain; only 3 retrospective studies were identified Guidelines were scarce and largely uninformative
Augustovski et al, 2015[246]	Search performed in Jul 2014	MEDLINE, Embase	Oncotype DX	15 studies	Oncotype DX predicts recurrence and treatment response Main benefits: better tailoring of treatment based on patient risk, changing decisions in 30% of women, sparing chemotherapy in low risk patients and increasing use of chemotherapy in high-risk patients Decisional impact could be higher in real-life settings or in select patient populations with greater uncertainty regarding initial treatment decisions Further research to clarify how large decision impact is in real-life settings, how it translates to patient relevant outcomes using high level of evidence research in all risk groups and also intermediate risk with higher uncertainty, and comparative effectiveness with existing risk-stratification tools

Author, Year	Search Period	Databases	Test(s)	No. Included Studies	Primary Conclusions
San Miguel et al, 2015[247] Belgian Health Care Knowledge Centre (KCE)	Search performed on Jun 27, 2014	MEDLINE, Embase, Cochrane library	EndoPredict, MammaPrint, Oncotype DX, Prosigna (PAM50) Others: MapQuant DX, BCI, BluePrint, Randox Breast Cancer Array, Mammostrat, NPI+, IHC4, uPA/PAI-1	13 reviews	Oncotype DX: more robust evidence than other tests, moderate to high quality of evidence supporting prognostic ability of test (clinical validity), no prospective studies on impact of test on long-term outcomes such as overall survival, 4 studies indicated that test leads to changes in decision-making 2 studies on predictive benefit of test were identified, 1 for LN+ patients; first evidence relating to improvements in quality of life and reduction in patient anxiety as result of using test has been reported but based on a small number of patients; further evidence is required
Marrone et al, 2015[248]	2007 to Dec 2013	PubMed, Embase, Cochrane library, CINAHL	Oncotype DX, MammaPrint	5 SRs	No direct evidence of clinical utility for Oncotype DX or MammaPrint Indirect evidence showed Oncotype DX can predict treatment effects of adjuvant chemotherapy, on evidence found for MammaPrint Both tests influenced treatment changes Cost-effectiveness of Oncotype DX varied depending on comparator; uncertainty in cost-effectiveness for MammaPrint
Institute for Quality and Efficiency in Health Care (IQWiG), 2016[249]	Nov 2015 Aug 2016 (for data on ongoing studies submitted by manufacturers)	MEDLINE, Embase, CENTRAL, CDSR, DARE, HTA database	EndoPredict, MammaPrint, Oncotype DX, PAM50 (Prosigna)	8 studies	No benefit or harm of biomarker-based strategy to support the decision for or against adjuvant chemotherapy
Issa et al, 2015[250]	Jan 2002 to Mar 2014	MEDLINE, Embase, PubMed	MammaPrint, Mammostrat, Oncotype DX, Rotterdam signature	25 studies	Meta-analysis: adjusted hazard ratio 3.538 (95% CI 1.513–8.469) Oncotype DX showed highest stability in estimation of likelihood of distant Risk of Recurrence RS led to 31.8% change in treatment recommendations
CADTH, 2017[251]	Jan 1, 2014, to Sep 12, 2017 Economic studies: since Jan 1, 2012	MEDLINE, PubMed, Cochrane library, CRD, HTA agencies, focused internet search	Oncotype DX, EndoPredict, MammaPrint, Prosigna, Mammostrat	1 clinical study, 1 cost-effectiveness study	Single comparative study found that Oncotype DX and EndoPredict were potentially useful assays in determining risk of distant recurrence Single cost-effectiveness study found that Mammostrat was more cost-effective than Oncotype DX from a United States third-party-payer perspective Overall findings limited by quantity and scope of each study found

Author, Year	Search Period	Databases	Test(s)	No. Included Studies	Primary Conclusions
Harnon et al, 2017[252] National Institute for Health and Care Excellence (NICE)	Inception to Feb 2017	MEDLINE, Embase, CDSR, DARE, CENTRAL, HTA database, NHS EED, Science Citation Index Expanded, Conference Proceedings Citation Index, WHO International Clinical Trials Registry Platform, ASCO, ESMO Contact with experts in the field, hand-searching of study reference lists, relevant studies from previous review by Ward et al 2013[172], evidence dossiers from manufacturers	Oncotype DX, EndoPredict, Prosigna, MammaPrint, IHC4	153 studies	All tests showed significant prognostic ability in unadjusted analyses in LN− and LN+ populations All tests provided additional prognostic information compared with common clinicopathological factors, with more variable results for the LN+ population Some evidence of chemotherapy benefit between risk groups for Oncotype DX Limited evidence of chemotherapy benefit for MammaPrint
Scope et al, 2017[253]	From Jan 2002 or Jan 2009, depending on test; update search from Jan 2013 to May 2016	MEDLINE, Embase, Cochrane library, Web of Science, BIOSIS Previews, Science Citation Index Expanded, Conference Proceedings Citation Index–Science	Oncotype DX, MammaPrint, IHC4, Mammostrat	40 studies	Some support for Oncotype DX in predicting chemotherapy benefit; test also impacts treatment decision-making Lower levels of evidence for IHC4, MammaPrint, and Mammostrat Study design limitations for all tests
Blok et al, 2018[117]	Searched for articles before Apr 2016	PubMed, Embase, Web of Science, Cochrane	EndoPredict, MammaPrint, Oncotype DX, Prosigna	149 studies	More evidence exists for Oncotype DX and MammaPrint, compared with EndoPredict and Prosigna Oncotype DX and MammaPrint are both useful prognostic tests that could reduce chemotherapy use with a generally favourable cost–benefit ratio, both have shown in prospective trials that low risk patients can safely forgo chemotherapy despite clinical risk factors Benefit of chemotherapy in high-risk patients has only been shown for Oncotype DX so far Need further prospective studies on all tests
EUnetHTA, 2018[254]	Jun 2014 to Apr 2017	MEDLINE, Embase, Cochrane library, ClinicalTrials.gov, EU Clinical Trials Register	MammaPrint	1 RCT	Clinical utility of MammaPrint not yet proven

Author, Year	Search Period	Databases	Test(s)	No. Included Studies	Primary Conclusions
King et al, 2018[255] Washington State Health Authority	Jan 2007 to Nov 2017 Dossier: Dec 2016	MEDLINE, CDSR, CENTRAL, AHRQ, Blue Cross/Blue Shield, NICE, Veterans Administration Evidence-based Synthesis Program, test manufacturer websites, dossier submitted to Washington State Agency Medicator Directors' Group, ClinicalTrials.gov, AHRQ National Guideline Clearinghouse, CMS, Aetna, Cigna, Regence websites for coverage policies	Oncotype DX, EndoPredict, Prosigna, MammaPrint, Mammostrat, BCI	22 reports 3 SRs, 17 primary studies, 2 economic studies	Moderate quality evidence to support use of MammaPrint and Oncotype DX for important outcomes related to early stage invasive breast cancer with any LN status Based on limited economic studies, tests are likely supported at currently conventional economic thresholds for use
Oregon Health Authority, 2018[256]	From 2012	MEDLINE, AHRQ, Blue Cross/Blue Shield, CADTH, Cochrane library, ICER, Medicaid Evidence-based Decisions Project, NICE, Tufts Cost-Effectiveness Analysis Registry, Veterans Administration Evidence-based Synthesis Program, Washington State Health Technology Assessment Program, national and international guidelines	Oncotype DX, EndoPredict, Prosigna, MammaPrint	2 SRs, 1 guideline, 1 RCT, 3 observational studies, 21 additional recent observational studies	The following breast cancer genome profile tests (1 test per primary breast cancer) are recommended for coverage in patients with early-stage breast cancer when the patient is willing to use the results of this testing in shared decision-making regarding adjuvant chemotherapy, and when the listed criteria are met (lymph nodes with micrometastases < 2 mm are considered node-negative): • Oncotype DX for ER+, HER2−, LN− (GRADE: strong recommendation) • Oncotype DX for ER+, HER2−, 1–3 positive nodes (GRADE: weak recommendation) • EndoPredict for ER+, HER2−, LN− (GRADE: weak recommendation) • Prosigna for ER+, HER2−, LN− (GRADE: weak recommendation) • MammaPrint for ER+, HER2−, LN−and only in those cases categorized as high clinical risk (GRADE: weak recommendation) • EndoPredict, Prosigjal and MammaPrint are not recommended for coverage in early-stage breast cancer with involved axillary lymph nodes (GRADE: weak recommendation)

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; ASCO, American Society of Clinical Oncology; BCI, Breast Cancer Index; CADTH, Canadian Agency for Drugs and Technologies in Health; CDSR, Cochrane Database of Systematic Reviews; CENTRAL, Cochrane Controlled Register of Trials; CI, confidence interval; CINAHL, Cumulative Index to Nursing and Allied Health Literature; CMS, Centers for Medicare & Medicaid Services; CRD, Centre for Reviews and Dissemination; DARE, Database of Abstracts of Reviews of Effects; ER, estrogen receptor; ESMO, European Society for Medical Oncology; EU, European Union; FDA, Food and Drug Administration; GRADE, Grading of Recommendations Assessment, Development and Evaluation; HER2, human epidermal growth factor receptor 2; HTA, health technology assessment; ICER, Institute for Clinical and Economic Review; IHC, immunohistochemistry; LN, lymph node; NHS EED, National Health Service Economic Evaluation Database; NICE, National Institute for Health and Care Excellence; NPI, Nottingham Prognostic Index; RS, Recurrence Score; SABCS, San Antonio Breast Cancer Symposium; uPA/PAI-1, urokinase plasminogen activator/plasminogen activator inhibitor type 1; WHO, World Health Organization.

Appendix 4: Literature Search Strategies

Clinical Evidence Search

Search date: November 28, 2018

Databases searched: Ovid MEDLINE, Embase

Database: Embase <1980 to 2018 Week 48>, Ovid MEDLINE(R) ALL <1946 to November 28, 2018>

Search strategy:

----

1 exp breast cancer/ (674538)
2 breast cancer.mp. (685132)
3 (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r$)).mp. (860481)
4 or/1-3 (868568)
5 (oncotype$ or 21 gene or recurrence score).mp. (4429)
6 (prosigna or PAM50).mp. (1055)
7 (mammaprint or 70 gene).mp. (2118)
8 endopredict.mp. (207)
9 or/5-8 (6868)
10 tailorx.mp. (124)
11 rxponder.mp. (39)
12 (swog adj (S1007 or "8814")).mp. (31)
13 (nsabp adj (b20 or b-20 or b 20)).mp. (26)
14 (nsabp adj (b14 or b-14 or b 14)).mp. (44)
15 transatac.mp. (60)
16 ((ma17 or ma 17 or ma-17 or ma12 or ma 12 or ma-12) adj (trial or study)).mp. (142)
17 (ABCSG-6 or ABCSG 6 or ABCSG-8 or ABCSG 8).mp. (45)
18 mindact.mp. (133)
19 (raster adj2 study).mp. (36)
20 (geicam 9906 or geicam-9906 or geicam9906).mp. (38)
21 (OPTIMA adj2 study).mp. (121)
22 or/10-21 (726)
23 exp randomized controlled trials as topic/ or exp clinical trials, phase III as topic/ or exp clinical trials, phase IV as topic/ (296071)
24 (randomized controlled trial or clinical trial, phase III or clinical trial, phase IV).pt. (475026)
25 random allocation/ or double blind method/ or single blind method/ (479988)
26 (randomi$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw. (495229)
27 or/23-26 (1392273)
28 (phase II or phase 2).tw. or exp clinical trial/ or exp clinical trial as topic/ (2462100)
29 (clinical trial or clinical trial, phase II or controlled clinical trial).pt. (554256)
30 (28 or 29) and random$.tw. (961725)
31 (clinic$ adj trial$1).tw. (764237)
32 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3 or dummy)).tw. (370875)
33 placebos/ (290691)
34 (placebo? or random allocation or randomly allocated or allocated randomly).tw. (528131)
35 (allocated adj2 random).tw. (1577)
36 Prospective study/ (972223)

37 Retrospective study/ (1428721)
38 Cohort study/ (623022)
39 or/30-38 (4483543)
40 27 or 39 (4968980)
41 (4 and 9 and 40) or 22 (1831)
42 (comment or letter or editorial or note or erratum or short survey or news or newspaper article or patient education handout or case report or historical article).pt. (5041955)
43 41 not 42 (1790)
44 exp animal/ not human/ (8532793)
45 43 not 44 (1785)
46 limit 45 to english language (1748)
47 limit 46 to yr="2018 -Current" (236)
48 47 use medall (41)
49 exp breast cancer/ (674538)
50 breast cancer.mp. (685132)
51 (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r$)).mp. (860481)
52 or/49-51 (868568)
53 (oncotype$ or 21 gene or recurrence score).mp. (4429)
54 (prosigna or PAM50).mp. (1055)
55 (mammaprint or 70 gene).mp. (2118)
56 endopredict.mp. (207)
57 or/53-56 (6868)
58 TAILORx.mp. (124)
59 rxponder.mp. (39)
60 (SWOG adj (S1007 or "8814")).mp. (31)
61 (nsabp adj (b20 or b-20)).mp. (26)
62 (nsabp adj (b14 or b-14)).mp. (44)
63 transatac.mp. (60)
64 ((ma17 or ma 17 or ma-17 or ma12 or ma 12 or ma-12) adj (trial or study)).mp. (142)
65 (ABCSG-6 or ABCSG 6 or ABCSG-8 or ABCSG 8).mp. (45)
66 mindact.mp. (133)
67 (raster adj2 study).mp. (36)
68 (geicam 9906 or geicam-9906 or geicam9906).mp. (38)
69 (OPTIMA adj2 study).mp. (121)
70 or/58-69 (726)
71 exp randomized controlled trial/ or exp phase 3 clinical trial/ or exp phase 4 clinical trial/ (1012719)
72 randomization/ or single blind procedure/ or double blind procedure/ (355618)
73 (randomi$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw. (495229)
74 or/71-73 (1484869)
75 (phase II or phase 2).tw. or exp clinical trial/ or exp prospective study/ or exp controlled clinical trial/ (2956073)
76 75 and random$.tw. (928951)
77 (clinic$ adj trial$1).tw. (764237)
78 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3 or dummy)).tw. (370875)
79 placebo/ (312955)
80 (placebo? or random allocation or randomly allocated or allocated randomly).tw. (528131)
81 (allocated adj2 random).tw. (1577)
82 Prospective study/ (972223)

83 Retrospective study/ (1428721)
84 Cohort study/ (623022)
85 or/77-84 (4061052)
86 74 or 76 or 85 (4900995)
87 (52 and 57 and 86) or 70 (1803)
88 (editorial or note or letter erratum or short survey).pt. or abstract report/ or letter/ or case study/ (5938142)
89 87 not 88 (1769)
90 animal/ not human/ (5399223)
91 89 not 90 (1768)
92 limit 91 to english language (1730)
93 limit 92 to yr="2018 -Current" (237)
94 93 use emez (196)
95 48 or 94 (237)
96 remove duplicates from 95 (207)

Economic Evidence Search

Search date: December 4, 2018

Databases searched: Ovid MEDLINE, Embase, Cochrane Database of Systematic Reviews, CRD Health Technology Assessment Database, Cochrane Central Register of Controlled Trials, and NHS Economic Evaluation Database

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <October 2018>, EBM Reviews - Cochrane Database of Systematic Reviews <2005 to November 30, 2018>, EBM Reviews - Health Technology Assessment <4th Quarter 2016>, EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>, Embase <1980 to 2018 Week 49>, Ovid MEDLINE(R) ALL <1946 to November 30, 2018>

Search strategy:

----

1 exp breast neoplasms/ (745819)
2 breast cancer.mp. (713151)
3 (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r$)).mp. (891290)
4 or/1-3 (900870)
5 ((genetic$ or gene$1 or genome$1 or genomic$) adj3 (profil$ or signature$ or assay$1)).ti. (41142)
6 ((genome$1 or genomic$) adj test$).ti. (672)
7 (oncotype$ or 21 gene or recurrence score).mp. (4652)
8 (prosigna$ or PAM50 or PAM 50 or 50 gene).mp. (1723)
9 (mammaprint$ or 70 gene).mp. (2231)
10 (endopredict$ or EPclin or 12 gene).mp. (3014)
11 or/5-10 (51052)
12 4 and 11 (7686)
13 economics/ (250016)
14 economics, medical/ or economics, pharmaceutical/ or exp economics, hospital/ or economics, nursing/ or economics, dental/ (800419)
15 economics.fs. (412239)

16 (econom* or price or prices or pricing or priced or discount* or expenditure* or budget* or pharmacoeconomic* or pharmaco-economic*).ti,ab,kf. (832818)
17 exp "costs and cost analysis"/ (561802)
18 (cost or costs or costing or costly).ti. (250319)
19 cost effective*.ti,ab,kf. (303427)
20 (cost* adj2 (util* or efficacy* or benefit* or minimi* or analy* or saving* or estimate* or allocation or control or sharing or instrument* or technolog*)).ab,kf. (199151)
21 models, economic/ (11977)
22 markov chains/ or monte carlo method/ (76492)
23 (decision adj1 (tree* or analy* or model*)).ti,ab,kf. (39190)
24 (markov or markow or monte carlo).ti,ab,kf. (121702)
25 quality-adjusted life years/ (37241)
26 (QOLY or QOLYs or HRQOL or HRQOLs or QALY or QALYs or QALE or QALEs).ti,ab,kf. (66648)
27 ((adjusted adj1 (quality or life)) or (willing* adj2 pay) or sensitivity analys*s).ti,ab,kf. (108586)
28 or/13-27 (2422214)
29 12 and 28 (676)
30 Case Reports/ or Comment.pt. or Editorial.pt. or Letter.pt. or Congresses.pt. (5048129)
31 29 not 30 (646)
32 exp Animals/ not Humans/ (15652250)
33 31 not 32 (346)
34 limit 33 to english language [Limit not valid in CDSR; records were retained] (334)
35 limit 34 to yr="2016 -Current" (114)
36 35 use medall,coch,cctr,clhta (60)
37 limit 12 to english language [Limit not valid in CDSR; records were retained] (7457)
38 limit 37 to yr="2016 -Current" (2545)
39 38 use cleed (0)
40 36 or 39 (60)
41 exp breast cancer/ (687477)
42 breast cancer.mp. (713151)
43 (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r$)).mp. (891290)
44 or/41-43 (899399)
45 ((genetic$ or gene$1 or genome$1 or genomic$) adj3 (profil$ or signature$ or assay$1)).ti. (41142)
46 ((genome$1 or genomic$) adj test$).ti. (672)
47 (oncotype$ or 21 gene or recurrence score).mp. (4652)
48 (prosigna$ or PAM50 or PAM 50 or 50 gene).mp. (1723)
49 (mammaprint$ or 70 gene).mp. (2231)
50 (endopredict$ or EPclin or 12 gene).mp. (3014)
51 or/45-50 (51052)
52 44 and 51 (7683)
53 Economics/ (250016)
54 Health Economics/ or Pharmacoeconomics/ or Drug Cost/ or Drug Formulary/ (124859)
55 Economic Aspect/ or exp Economic Evaluation/ (439295)
56 (econom* or price or prices or pricing or priced or discount* or expenditure* or budget* or pharmacoeconomic* or pharmaco-economic*).tw,kw. (857335)
57 exp "Cost"/ (561802)
58 (cost or costs or costing or costly).ti. (250319)
59 cost effective*.tw,kw. (314599)

60 (cost* adj2 (util* or efficac* or benefit* or minimi* or analy* or saving* or estimate* or allocation or control or sharing or instrument* or technolog*)).ab,kw. (207079)
61 Monte Carlo Method/ (61198)
62 (decision adj1 (tree* or analy* or model*)).tw,kw. (42894)
63 (markov or markow or monte carlo).tw,kw. (126694)
64 Quality-Adjusted Life Years/ (37241)
65 (QOLY or QOLYs or HRQOL or HRQOLs or QALY or QALYs or QALE or QALEs).tw,kw. (70455)
66 ((adjusted adj1 (quality or life)) or (willing* adj2 pay) or sensitivity analys*s).tw,kw. (128183)
67 or/53-66 (2068795)
68 52 and 67 (621)
69 Case Report/ or Comment/ or Editorial/ or Letter/ or conference abstract.pt. (10077342)
70 68 not 69 (398)
71 (exp animal/ or nonhuman/) not exp human/ (10095638)
72 70 not 71 (396)
73 limit 72 to english language [Limit not valid in CDSR; records were retained] (373)
74 limit 73 to yr="2016 -Current" (107)
75 74 use emez (53)
76 40 or 75 (113)
77 76 use medall (42)
78 76 use coch (0)
79 76 use cctr (16)
80 76 use clhta (2)
81 76 use cleed (0)
82 76 use emez (53)
83 remove duplicates from 76 (75)

Quantitative Preferences Evidence Search

Search date: December 17, 2018

Databases searched: Ovid MEDLINE

Search filter used: Quantitative preference evidence filter, modified from Selva et al[199]

Database: Ovid MEDLINE(R) ALL <1946 to December 14, 2018>

Search strategy:
----
1 exp breast neoplasms/ (270023)
2 breast cancer.mp. (244661)
3 (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r$)).mp. (349405)
4 or/1-3 (349495)
5 ((genetic$ or gene$1 or genome$1 or genomic$) adj3 (profil$ or signature$ or assay$1)).ti. (17007)
6 ((genomic$ or genome$1) adj test$).ti. (287)
7 (oncotype$ or 21 gene or recurrence score).mp. (1309)
8 (prosigna$ or PAM50 or PAM 50 or 50 gene).mp. (448)
9 (mammaprint$ or 70 gene).mp. (692)
10 (endopredict$ or EPclin or 12 gene).mp. (1114)

11 or/5-10 (20219)
12 4 and 11 (2274)
13 Attitude to Health/ (80892)
14 Health Knowledge, Attitudes, Practice/ (99788)
15 Patient Participation/ (23215)
16 exp Patient Satisfaction/ (81583)
17 Attitude of Health Personnel/ (113344)
18 *Professional-Patient Relations/ (10871)
19 *Physician-Patient Relations/ (33584)
20 Choice Behavior/ (30029)
21 (choice or choices or value* or valuation*).ti. (185574)
22 (preference* or expectation* or attitude* or acceptab* or knowledge or point of view).ti,ab. (1071636)
23 ((patient*1 or user*1 or men or women or personal or provider* or practitioner* or professional*1 or (health* adj2 worker*) or clinician* or physician* or doctor* or surgeon* or oncologist* or pathologist*) adj2 (participation or perspective* or perception* or misperception* or perceiv* or view* or understand* or misunderstand* or satisf* or dissatisf* or value*1)).ti,ab. (163512)
24 health perception*.ti,ab. (2469)
25 Stress, Psychological/ (110905)
26 (psycholog* or psychosocial or psycho social or emotion* or anxiet* or anxious* or worry* or worries or confus* or distress* or reassur* or re assur*).ti,ab. (696216)
27 *Decision Making/ (37744)
28 (patient*1 or user*1 or men or women or personal or provider* or practitioner* or professional*1 or (health* adj2 worker*) or clinician* or physician* or doctor* or surgeon* or oncologist* or pathologist*).ti. (2268199)
29 27 and 28 (6984)
30 (decision* and mak*).ti. (25445)
31 (decision mak* or decisions mak*).ti,ab. (120417)
32 30 or 31 (121853)
33 (patient*1 or user*1 or men or women or personal or provider* or practitioner* or professional*1 or (health* adj2 worker*) or clinician* or physician* or doctor* or surgeon* or oncologist* or pathologist*).ti,ab. (7464314)
34 32 and 33 (76216)
35 (discrete choice* or decision board* or decision analy* or decision-support or decision tool* or decision aid* or latent class* or decision* conflict* or decision* regret*).ti,ab. (28888)
36 Decision Support Techniques/ (18212)
37 (health and utilit*).ti. (1287)
38 (gamble* or prospect theory or health utilit* or utility value* or utility score* or utility estimate* or health state or feeling thermometer* or best-worst scaling or time trade-off or TTO or probability trade-off).ti,ab. (11645)
39 (preference based or preference score* or preference elicitation or multiattribute or multi attribute).ti,ab. (2433)
40 or/13-26,29,34-39 (2312948)
41 12 and 40 (382)
42 limit 41 to english language (370)

Grey Literature Search

Performed: December 4–6, 2018
Websites searched:
HTA Database Canadian Repository, Alberta Health Technologies Decision Process reviews, BC Health Technology Assessments, Canadian Agency for Drugs and Technologies in Health (CADTH), Institut national d'excellence en santé et en services sociaux (INESSS), Institute of Health Economics (IHE), Laval University, McGill University Health Centre Health Technology Assessment Unit, National Institute for Health and Care Excellence (NICE), Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers, Australian Government Medical Services Advisory Committee, Queensland Health Technology Evaluation, Centers for Medicare & Medicaid Services Technology Assessments, Institute for Clinical and Economic Review, Healthcare Improvement Scotland, Ireland Health Information and Quality Authority Health Technology Assessments, Washington State Health Care Authority Health Technology Reviews, ClinicalTrials.gov, PROSPERO, EUnetHTA, Epistemonikos, Tuft's Cost-Effectiveness Analysis Registry

Keywords used: Oncotype, OncotypeDX, Prosigna, Endopredict, Mammaprint, gene profiling, genetic profiling, expression profiling

Results (included in PRISMA): 7

Ongoing clinical trials: 21 (ClinicalTrials.gov)

Ongoing HTAs: 4 (PROSPERO/Australian Government Medical Services Advisory Committee)

Additional results from grey literature update: May 2–3, 2019

Results (included in PRISMA): 0

Ongoing clinical trials: 6 (ClinicalTrials.gov)

Ongoing HTAs: 1 (Alberta Health and Wellness)

Cancer Care Ontario Literature Search

Search strategy: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>

Search term (hits)

1. exp breast cancer/ (235614) Breast cancer terms
2. breast cancer.mp. (194555)
3. (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r)).mp. (290158)
4. or/1-3 (290182)
5. (oncotype or 21 gene or recurrence score).mp. (812)
6. (prosigna or PAM50).mp. (129) Profiling Assays
7. (mammaprint or 70 gene).mp. (550)

8. endopredict.mp. (25)
9. or/5-8 (1388)
10. tailorx.mp. (15) Terms for important studies
11. rxponder.mp. (7) known *a priori*
12. (swog adj (S1007 or "8814")).mp. (10)
13. (nsabp adj (b20 or b-20 or b 20)).mp. (8)
14. (nsabp adj (b14 or b-14 or b 14)).mp. (15)
15. transatac.mp. (13)
16. ((ma17 or ma 17 or ma-17 or ma12 or ma 12 or ma-12) adj (trial or study)).mp. (61)
17. (ABCSG-6 or ABCSG 6 or ABCSG-8 or ABCSG 8).mp. (17)
18. mindact.mp. (22)
19. (raster adj2 study).mp. (10)
20. (geicam 9906 or geicam-9906 or geicam9906).mp. (8)
21. (OPTIMA adj2 study).mp. (20)
22. or/10-21 (179)
23. exp randomized controlled trials as topic/ or exp clinical trials, as topic/ or exp clinical trials, phase IV as topic/ (107740)
24. (randomized controlled trial or clinical trial, phase III or clinical trial, phase IV).pt. (410812)
25. random allocation/ or double blind method/ or single blind method/ (229854)
26. (randomi$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw. (138935)
27. or/23-26 (664107)
28. (phase II or phase 2).tw. or exp clinical trial/ or exp clinical trial as topic/ (1077720)
29. (clinical trial or clinical trial, phase II or controlled clinical trial).pt. (537237)
30. (28 or 29) and random$.tw. (379415)
31. (clinic$ adj trial$1).tw. (245790)
32. ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3 or dummy)).tw. (140226)
33. placebos/ (33849)
34. (placebo? or random allocation or randomly allocated or allocated randomly).tw. (191968)
35. (allocated adj2 random).tw. (747)
36. Prospective study/ (401247)
37. Retrospective study/ (550579)
38. Cohort study/ (186361)
39. or/30-38 (1638866)
40. 27 or 39 (1852392)
41. (4 and 9 and 40) or 22 (323) (comment or letter or editorial or note or erratum or short survey or news or newspaper article or patient education handout or case report or historical article).pt. (1987212)
42. 41 not 42 (318)
43. exp animal/ not human/ (4096239)
44. 43 not 44 (317)
45. limit 45 to english (314)
46. limit 46 to yr="2002-2016" (309)

Search strategy: Embase <1996 to 2016 Week 7>

Search term (hits)

1. breast cancer/ (221856)
2. breast cancer.mp. (291233)
3. (breast adj2 (cancer$ or neoplasm$ or carcinoma$ or malignan$ or tumo?r)).mp. (334758)
4. or/1-3 (334758)
5. (oncotype or 21 gene or recurrence score).mp. (1747)
6. (prosigna or PAM50).mp. (317) Profiling Assays
7. (mammaprint or 70 gene).mp. (994)
8. endopredict.mp. (56)
9. or/5-8 (2756)
10. TAILORx.mp. (48)
11. rxponder.mp. (16)
12. (SWOG adj (S1007 or "8814")).mp. (16)
13. (nsabp adj (b20 or b-20)).mp. (16)
14. (nsabp adj (b14 or b-14)).mp. (16)
15. transatac.mp. (27)
16. ((ma17 or ma 17 or ma-17 or ma12 or ma 12 or ma-12) adj (trial or study)).mp. (76)
17. (ABCSG-6 or ABCSG 6 or ABCSG-8 or ABCSG 8).mp. (27)
18. mindact.mp. (64)
19. (raster adj2 study).mp. (17)
20. (geicam 9906 or geicam-9906 or geicam9906).mp. (21)
21. (OPTIMA adj2 study).mp. (69)
22. or/10-21 (354)
23. exp randomized controlled trial/ or exp phase 3 clinical trial/ or exp phase 4 clinical trial/ (347436)
24. randomization/ or single blind procedure/ or double blind procedure/ (174384)
25. (randomi$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw. (182769)
26. or/23-25 (517584)
27. (phase II or phase 2).tw. or exp clinical trial/ or exp prospective study/ or exp controlled clinical trial/ (1127414)
28. 27 and random$.tw. (347289)
29. (clinic$ adj trial$1).tw. (286181)
30. ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3 or dummy)).tw. (127920)
31. placebo/ (215324)
32. (placebo? or random allocation or randomly allocated or allocated randomly).tw. (192316)
33. (allocated adj2 random).tw. (303)
34. Prospective study/ (283378)
35. Retrospective study/ (391863)
36. Cohort study/ (180186)
37. or/29-36 (1322563)
38. 26 or 28 or 37 (1622632)
39. (4 and 9 and 38) or 22 (709)

40. (editorial or note or letter erratum or short survey).pt. or abstract report/ or letter/ or case study/ (1787525)
41. 39 not 40 (689)
42. animal/ not human/ (506080)
43. 41 not 42 (688)
44. limit 43 to english (671)
45. limit 44 to exclude medline journals (24)
46. limit 45 to yr="2002-2016" (24)

Appendix 5: Characteristics of Included Studies, Clinical Evidence Review

Table A7: Characteristics of Included Studies, Clinical Evidence Review

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
EndoPredict									
Dubsky et al, 2013[257]	Retrospective analysis of RCTs (ABCSG-6 and ABCSG-8)	Austria	1,702	Low: EPclin < 3.3 (EP < 5) High: EPclin ≥ 3.3 (EP ≥ 5)	Low: 1066 (63) High: 636 (37)	Median: 63.8 (range 41.5–80.7) Post: 100%	100% ER+ 79% PR+ 100% HER2−	68% LN− 32% LN+	ET
Ettl et al, 2017[104]	Prospective observational study	Germany	395	Low: EPclin < 3.3 (EP < 5) High: EPclin ≥ 3.3 (EP ≥ 5)	Low: 250 (63) High: 145 (37)	Median: 59 (range 29–88) Pre/Post: NR	100% HR+ 100% HER2−	77% LN− 23% LN+ (includes N1mic)	ET, ET + CT
Fallowfield et al, 2018[105]	Prospective observational study	United Kingdom	149	Low: NR High: NR	Low: 75 (50) High: 74 (50)	Mean: 56.4 (SD 10.9, range 26–77) Pre/Post: NR	100% ER+ 100% HER2−	66% LN− 34% LN+ (includes N1mic)	ET, ET + CT
Filipits et al, 2019[62]	Retrospective analysis of RCTs (ABCSG-6 and ABCSG-8)	Austria	1,702	Low: EPclin < 3.3 High: EPclin ≥ 3.3	Low: 1,066 (63) High: 636 (37)	Median: 63 (IQR 58–70) Post: 100%	100% ER+ 100% HER2−	69% LN− 31% LN+	ET
Mokbel et al, 2017[107] Mokbel et al, 2018[107]	Prospective observational study	United Kingdom	120	Low: EPclin < 3.3 High: EPclin ≥ 3.3	Low: 60 (50) High: 60 (50)	Median: 54 (range 31–77) Pre/Post: NR	100% ER+ 100% HER2−	74% LN− 26% LN+	ET, ET + CT
Sestak et al, 2019[102]	Retrospective analysis of RCTs (GEICAM/9906, GEICAM 2003/02, ABCSG-6, ABCSG-8, TransATAC)	Austria, Spain, United Kingdom	3,746	Low: EPclin < 3.3 High: EPclin ≥ 3.3	NR	Median: 61 (IQR 54–68) Pre: 15% Post: 85%	100% ER+ 100% HER2−	66% LN− 34% LN+	ET, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
MammaPrint									
Cardoso et al, 2016[51]	RCT (MINDACT)	9 European countries	6,693	Low: 0 to 1 High: −1 to 0	Low: 4,295 (64) High: 2,398 (36)	< 50: 33% ≥ 50: 67% Pre/Post: NR	88% HR+ 90% HER2−	79% LN− 21% LN+	ET, ET + CT
Cusumano et al, 2014[108]	Prospective observational study	Belgium, Italy, Netherlands	194	Low: 0 to 1 High: −1 to 0	Low: 85 (44) High: 109 (56)	Median: 56 (range 25–69) Pre/Post: NR	86% ER+ 88% HER2−	66% LN− 32% LN+ 1% unknown	ET, ET + CT
Drukker et al, 2013[63] Drukker et al, 2014[64]	Prospective observational study (RASTER)	Netherlands	427	Low: 0 to 1 High: −1 to 0	Low: 219 (51) High: 208 (49)	≤ 50: 68% >50: 32% Pre/Post: NR	80% ER+ 69% PR+ 84% HER2−	100% LN− (includes N1mic)	No ET, ET, ET + CT
Esserman et al, 2017[65]	Reanalysis of RCT (STO-3)	Sweden	652	Low: 0 to 1 High: −1 to 0	Low: 377 (58) High: 275 (42)	< 55: 11% ≥ 55: 89% Post: 100%	83% ER+ 58% PR+ 92% HER2−	100% LN−	No ET, ET
Kuijer et al, 2017[78]	Prospective observational study	Netherlands	660	Low: 0 to 1 High: −1 to 0	Low: 390 (59) High: 270 (41)	Median: 57 (SD 8.1) Pre/peri: 34% Post: 66% Unknown: 2%	100% ER+ 87% PR+ 97% HER2−	84% LN− 15% LN+ (includes N1mic) 1% unknown	ET, ET + CT
van de Vijver et al, 2002[66]	Retrospective analysis of Netherlands Cancer Institute database	Netherlands	295	Low: 0 to 1 High: −1 to 0	Low: 180 (61) High: 115 (39)	< 50: 83% ≥ 50: 17% Pre/Post: NR	77% ER+	51% LN− 49% LN+	ET, ET + CT
van't Veer et al, 2017[67]	Retrospective analysis of RCT (STO-3)	Sweden	538	Low: 0 to 1 High: −1 to 0	Low: 371 (69) High: 167 (31)	< 55: 10% ≥ 55: 90% Post: 100%	100% ER+ 68% PR+ 95% HER2−	100% LN−	No ET, ET
Wuerstlein et al, 2019[49]	Prospective observational study (WSG PRIMe)	Germany	430	Low: 0 to 1 High: −1 to 0	Low: 273 (63) High: 157 (36)	Median: 58 (range 33–88) Pre: 30% Post: 68% Unknown: 2%	99% ER+ 90% PR+ 99% HER2−	72% LN− 28% LN+	ET, CT, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Oncotype DX									
Albain et al, 2010[258]	Retrospective analysis of RCT (SWOG-8814)	United States	367	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 146 (40) Int: 103 (28) High: 118 (32)	Mean: 60 years (SD 7.5) Post: 100%	97% ER+ 80% PR+ 88% HER2−	38% LN− 62% LN+	ET, ET + CT
Albanell et al, 2012[77]	Prospective observational study (transGEICAM)	Spain	107	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 62 (58) Int: 35 (33) High: 10 (9)	< 50: 37% ≥ 50: 62% Pre/Post: NR	100% ER+ 84% PR+ 100% HER2−	100% LN−	ET, ET + CT
Albanell et al, 2016[79]	Retrospective analysis of 4 prospective observational studies	France, Germany, Spain, United Kingdom	565	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 312 (55) Int: 199 (35) High: 54 (10)	Mean: 56 (SD 10.1; range 25–85) Pre/Post: NR	100% ER+ 87% PR+ 100% HER2−	100% LN−	ET, ET + CT
Bargallo et al, 2015[80]	Prospective observational study	Mexico	96	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 46 (48) Int: 30 (31) High: 20 (21)	< 50: 30% ≥ 50: 70% Pre/Post: NR	100% ER+ 100% HER2−	65% LN− 35% LN+ (includes N1mic)	ET, ET + CT
Curtit et al, 2019[48]	Prospective observational study (PONDx)	France	866	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 474 (54) Int: 314 (36) High: 78 (9)	< 50: 27% > 70: 14% Pre: 23% Peri: 9% Post: 67%	100% ER+ 86% PR+ 99% HER2−	71% LN− 29% LN+ (includes N1mic)	ET, ET + CT
de Boer et al, 2013[81]	Prospective observational study	Australia	151	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 72 (48) Int: 59 (39) High: 20 (13)	Mean: 56.2 (NR)	90% HR+ 100% HER2−	67% LN− 33% LN+	ET, ET + CT
Dieci et al, 2018[82]	Prospective observational study (Breast DX Italy)	Italy	250	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 152 (61) Int: 81 (32) High: 17 (7)	Median: 55 (range 27–83) Pre: 41% Post: 59%	Median 90% ER+ (range 40%–100%) Median 80% PR+ (range: 0%–100%) 100% HER2−	50% LN− 50% LN+	ET, ET + CT
Dowsett et al, 2010[68]	Retrospective analysis of RCT (TransATAC)	United Kingdom	1,231	Low: ≤ 17 Int: 18–30 High: ≥ 31	NR	Mean: 64.3 (NR) Post: 100%	100% HR+	71% LN− 25% LN+ 4% unknown	ET
Eiermann et al, 2013[83]	Prospective observational study	Germany	366	Low: NR Int: NR High: NR	Low: 198 (54) Int: 139 (38) High: 29 (8)	Mean: 56 (NR) Pre/Post: NR	100% ER+ 89% PR+ 100% HER2−	67% LN− 33% LN+	ET, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Geyer et al, 2018[76]	Retrospective analysis of RCT (NSABP B-20)	United States	569	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 347 (61) Int: 125 (22) High: 97 (17)	Median: 51 (range 28–74) Pre/Post: NR	100% ER+ ≥ 86% PR+ 100% HER2−	100% LN−	ET, ET + CT
Gluz et al, 2016[94]	Retrospective analysis of RCT (WSG Plan B)	Germany	2,642	Low: ≤ 11 Int: 12–25 High: > 25	Low: 459 (17) Int: 1,544 (58) High: 550 (21) Unknown: 89 (3)	Median: 56 Pre/Post: NR	91% ER+ 75% PR+ 100% HER2−	59% LN− 41% LN+	Low: ET Int/high: ET + CT
Ibraheem et al, 2019[75]	Retrospective analysis of National Cancer Database	United States	73,185	Int: 11–30	Int: 73,185 (100)	Mean: 58 (SD 10.5) Pre/Post: NR	100% ER+ 92% PR+ 100% HER2−	82% LN− 17% LN+	ET, ET + CT
King et al, 2016[95]	Prospective observational study	United States	109	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 22 (20) Int: 29 (27) High: 50 (46) Unknown: 8 (7)	Median: 52 (range 21–79) Pre/Post: NR	66% HR+ and HER2− 18% HR+ and HER2+	16% LN− 71% LN+ 13% unknown	ET, CT, ET + CT
Kuchel et al, 2016[111]	Prospective observational study	United Kingdom	137	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 71 (52) Int: 58 (42) High: 8 (6)	Median: 55 (range 31–80) Pre/Post: NR	100% ER+ 100% HER2−	72% LN− 27% LN+ (includes N1mic) 1% unknown	ET, ET + CT
Leung et al, 2016[112]	Prospective observational study	Hong Kong	146	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 74 (51) Int: 51 (35) High: 21 (14)	< 50: 40% ≥ 50: 60% Pre: 47% Post: 53%	100% ER+ 49% HER2−	84% LN− 16% N1mic	ET, ET + CT
Levine et al, 2016[84]	Prospective study	Canada	979	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 565 (58) Int: 322 (33) High: 92 (9)	< 50: 21% ≥ 50: 79% Pre/Post: NR	100% ER+ 91% PR+ 99% HER2−	100% LN− (includes N1mic)	ET, ET + CT
Lo et al, 2010[85]	Prospective observational study	United States	89	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 38 (43) Int: 42 (47) High: 9 (10)	Mean: 55 (range 35–77) Pre/Post: NR	100% ER+ 93% HER2−	100% LN−	ET, ET + CT
Loncaster et al, 2017[86]	Retrospective analysis of prospective study	United Kingdom	201	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 86 (43) Int: 89 (44) High: 26 (13)	Mean: 55 (SD 10.0; range 24–77) Pre/Post: NR	100% ER+ 100% HER2−	68% LN− 32% LN+	ET, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Mamounas et al, 2010[69]	Retrospective analysis of RCTs (NSABP B-14 and NSABP B-20)	United States	1,674	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 862 (51) Int: 368 (22) High: 444 (27)	Age: NR Pre/Post: NR	100% ER+	100% LN−	No ET, ET, ET + CT
Mamounas et al, 2017[96]	Retrospective analysis of RCT (NSABP B-28)	United States	1,065	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 386 (36) Int: 364 (34) High: 315 (30)	< 50: 48% ≥ 50: 52% Pre/Post: NR	100% ER+	100% LN+	ET + CT
Martinez del Prado et al, 2018[113]	Prospective observational study	Spain	401	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 222 (55) Int: 153 (38) High: 26 (7)	< 50: 24% ≥ 50: 76% Pre: 35% Post: 64% Missing: 1%	100% ER+ 90% PR+ 100% HER2−	77% LN− 23% LN+ (includes N1mic)	ET, ET + CT
Nitz et al, 2017[55]	Prospective observational study (WSG Plan B)	Germany	2,642	Low: ≤ 11 Int: 12–25 High: ≥ 26	Low: 459 (17) Int: 1,544 (58) High: 550 (21) Unknown: 89 (3)	Median: 56 (range 25–77) Pre/Post: NR	90% ER+ 74% PR+ 100% HER2−	59% LN 41% LN+	Low: ET Int/high: ET + CT
Ozmen et al, 2016[87]	Prospective observational study	Turkey	165	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 93 (56) Int: 58 (35) High 14 (9)	Median: 49 (range 26–76) Pre/Post: NR	100% ER+ 67% PR+ 100% HER2−	93% LN− 7% LN+	ET, ET + CT
Paik et al, 2004[70]	Retrospective analysis of RCT (NSABP B-14)	United States	668	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 341 (51) Int: 147 (22) High 180 (27)	< 50: 29% ≥ 50: 71% Pre/Post: NR	100% ER+ 91% HER2−	100% LN−	ET
Paik et al, 2006[56]	Retrospective analysis of RCT (NSABP B-20)	United States	651	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 353 (54) Int: 134 (21) High 164 (25)	< 50: 45% ≥ 50: 55% Pre/Post: NR	100% ER+	100% LN−	ET, ET + CT
Penault-Llorca et al, 2018[97]	Retrospective analysis of RCT (PACS-01)	Belgium, France	530	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 209 (39) Int: 159 (30) High 162 (31)	< 50: 53% ≥ 50: 47% Pre/Post: NR	96% ER+ 67% PR+ 90% HER2−	100% LN+	ET + CT
Pestalozzi et al, 2017[114]	Prospective observational study (SAKK 26/10)	Switzerland	222	Low: < 18 Non-low: ≥ 18	Low: 154 (69) Non-low: 68 (31)	Median: 58 (range 32–82) Pre: 28% Peri: 4% Post: 68%	100% ER/PR ≥ 50% 100% HER2−	61% LN− 39% LN+	ET, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Petkov et al, 2016[71]	Retrospective analysis of SEER database	United States	44,825	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 24,454 (55) Int: 16,821 (38) High: 3,550 (8)	< 50: 27% ≥ 51: 73% Pre/Post: NR	100% HR+ 100% HER2−	90% LN− 10% LN+ (includes N1mic)	ET, ET + CT
Roberts et al, 2017[98]	Retrospective analysis of SEER database	United States	6,483	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 3,790 (59) Int: 2,263 (35) High: 430 (6)	< 50: 22% ≥ 51: 78% Pre/Post: NR	100% ER+ 92% PR+ 100% HER2−	100% LN+ (includes N1mic)	ET, ET + CT
Sparano et al, 2015[259]	Prospective observational study	United States	8,523	Low: ≤ 10 Int: 11–25	Low: 1,626 (19) Int: 6,897 (81)	≤ 50: 32% > 50: 68% Pre: 35% Post: 65%	99% ER+ 93% PR+ 100% HER2−	100% LN−	Low: ET
Sparano et al, 2018[26]	RCT (TAILORx)	United States	9,719	Low: ≤ 10 Int: 11–25 High: ≥ 26	Low: 1,619 (17) Int: 6,711 (69) High: 1,389 (14)	Median: 56 (range 23–75) Pre: 34% Post: 66%	99% ER+ 88% PR+ 100% HER2−	100% LN−	Low: ET Int: randomized to ET or ET + CT High: ET + CT
Stemmer et al, 2017[99]	Retrospective analysis of Clalit Health Services database	Israel	709	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 379 (53) Int: 258 (36) High: 72 (10)	Median: 62 (IQR 53–67) Pre/Post: NR	100% ER+ 100% HER2−	100% LN+ (includes N1mic)	ET, ET + CT
Stemmer et al, 2017[72]	Retrospective analysis of Clalit Health Services database	Israel	1,801	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 880 (49) Int: 733 (41) High: 188 (10)	Median: 60 (IQR 52–67) Pre/Post: NR	100% ER+ 100% HER2−	100% LN−	ET, ET + CT
Torres et al, 2018[103]	Prospective observational study	Canada	67	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 38 (57) Int: 23 (34) High: 6 (9)	Mean: 61 (range 37–84) Pre: 28% Post: 72%	100% ER+ 93% PR+ 100% HER2−	99% LN+ 1% unknown	ET, ET + CT
Voelker et al, 2018[115]	Prospective observational study	Germany	50	Low: ≤ 17 Int: 18–30 High: ≥ 31	Low: 31 (62) Int: 16 (32) High: 3 (6)	Mean: 53 ± 11 Pre/Post: NR	100% HR+ 100% HER2−	66% LN− 34% LN+	ET, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Prosigna									
Filipits et al, 2014[54]	Retrospective analysis of RCT (ABCSG-8)	Austria	1,246	*LN−* Low: ≤ 40 Int: 41–60 High: > 60 *LN+ (1–3 nodes)* Low: ≤ 15 Int: 16–40 High > 40	*LN−* Low: 448 (36) Int: 292 (23) High: 179 (14) *LN+ (1–3 nodes)* Low: 12 (1) Int: 124 (10) High: 191 (15)	Age: NR Post: 100%	100% HR+ 100% HER2−	89% LN− 11% LN+	ET
Fitzal et al, 2015[92]	Retrospective analysis of RCT (ABCSG-8)	Austria	1,324	Low: NR High: NR	Low: 641 (48) High: 683 (52)	Median: 63 (range 41–80) Post: 100%	100% ER+ 79% PR+ 100% HER2−	71% LN− 29% LN+	ET
Gnant et al, 2014[73]	Retrospective analysis of RCT (ABCSG-8)	Austria	1,478	*LN−* Low: ≤ 40 Int: 41–60 High: > 60 *LN+ (1–3 nodes)* Low: ≤ 15 Int: 16–40 High > 40	*LN−* Low: 487 (33) Int: 335 (23) High: 225 (15) *LN+ (1–3 nodes)* Low: 15 (1) Int: 143 (10) High: 273 (18)	Median: 63 years (range: 41–79) Post: 100%	99% ER+ 82% PR+ 95% HER2−	71% LN− 29% LN+	ET
Gnant et al, 2015[260]	Retrospective analysis of RCT (ABCSG-8 and TransATAC)	United Kingdom	2,197	Low: NR Int: NR High: NR	NR	Age: NR Post: 100%	100% HR+ 93% HER2−	75% LN− 25% LN+	ET
Hequet et al, 2017[88]	Prospective observational study	France	200	Low: ≤ 40 Int: 41–60 High: > 60	Low: 93 (46) Int: 67 (34) High: 40 (20)	Mean: 62 (99% ≥ 50) Post: 100%	100% ER+ 86% PR+ 100% HER2−	100% LN−	ET, ET + CT
Jensen et al, 2018[101]	Retrospective analysis of RCT (DBCG 77B)	Denmark	460	Low: 8–51 Int: 52–71 High: ≥ 72	Low: 155 (34) Int: 148 (32) High: 157 (34)	< 50: 72% ≥ 50%: 28% Pre: 100%	72% HR+	9% LN− 85% LN+ 4% unknown	No CT, CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Laenkholm et al, 2018[74]	Retrospective analysis of DBCG database	Denmark	2,558	*LN−* Low: ≤ 40 Int: 41–60 High: > 60 *LN+ (1 node)* Low: ≤ 35 Int: 36–55 High: > 55 *LN+ (2 nodes)* Low: ≤ 25 Int: 26–45 High: > 45 *LN+ (3 nodes)* Int: ≤ 25 High: > 25	Low: 720 (28) Int: 763 (30) High: 1,075 (42)	Median: 63 (range 50–89) Pre/Post: NR	100% ER+ 100% HER2−	46% LN− 54% LN+	ET
Liu et al, 2015[58]	Retrospective analysis of RCT (NCIC MA.21)	Canada, United States	1,094	Low: NR Int: NR High: NR	Low: 37 (3) Int: 196 (18) High: 861 (79)	Median: 47 (range 23–61) Pre: 69% Post: 31%	58% ER+ 71% HER2−	30% LN− 70% LN+	CT, ET + CT
Ohnstad et al, 2017[59]	Retrospective analysis of prospective observational study (Oslo1)	Norway	476	Low: ≤ 40 Int: 41–60 High: ≥ 61	Low: 180 (38) Int: 108 (23) High: 188 (39)	Median: 58 (range 28–93) Post: 100%	73% ER+ 89% HER2−	64% LN− 32% LN+	No ET, ET, or ET + CT
Sestak et al, 2015[60]	Retrospective analysis of RCTs (ABCSG-8 and TransATAC)	Austria, United Kingdom	2,137	Low: ≤ 26 Int: 27–68 High: ≥ 69	Low: 1,139 (53) Int: 693 (32) High: 305 (14)	≤ 65: 60% > 65: 40% Post: 100%	100% ER+	74% LN− 26% LN+	ET
Wuerstlein et al, 2016[89]	Prospective observational study (WSG BCIST)	Germany	198	Low: ≤ 40 Int: 41–60 High: > 60	Low: 85 (43) Int: 70 (35) High: 43 (22)	Median: 64 (range 40–81) Post: 100%	100% ER+ 87% PR+ 100% HER2−	100% LN−	ET, ET + CT

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Multiple Gene Expression Profiling Tests									
Buus et al, 2016[61]	Retrospective analysis of RCT (TransATAC)	United Kingdom	928	*EndoPredict* Low: EPclin < 3.3 (EP < 5) High: EPclin ≥ 3.3 (EP ≥ 5) *Oncotype DX* Low: < 18 Non-low: ≥ 18	*EndoPredict* Low: 546 (59) High: 382 (41) *Oncotype DX* Low: 573 (62) Non-low: 355 (38)	Mean: 64.7 (SD 8.3) Post: 100%	100% HR+ 100% HER2−	73% LN− 27% LN+	ET
Dowsett et al, 2013[57]	Retrospective analysis of RCT (TransATAC)	United Kingdom	739	*Oncotype DX* Low: NR Int: NR High: NR *Prosigna* Low: NR Int: NR High: NR	*Oncotype DX* Low: 434 (59) Int: 243 (33) High: 62 (8) *Prosigna* Low: 428 (59) Int: 192 (26) High: 119 (16)	Mean: 64 (SD 8.3) Post: 100%	100% ER+ 88% HER2−	100% LN−	ET
Martin et al, 2016[93]	Retrospective analysis of RCT (GEICAM)	Spain	536	*EndoPredict* Low: EPclin < 3.3 High: EPclin ≥ 3.3 *Prosigna* Low: < 18 Int: 18–65 High: > 65	*EndoPredict* Low: 69 (13) High: 467 (87) *Prosigna* Low: 99 (18) Int: 298 (56) High: 139 (26)	Median: 51 (range 23–76) Pre/Post: NR	100% ER+ 100% HER2−	100% LN+	ET + CT
Sestak et al, 2013[100]	Retrospective analysis of RCT (TransATAC)	United Kingdom	940	*Oncotype DX* Low: NR High: NR *Prosigna* Low: NR High: NR	*Oncotype DX* NR *Prosigna* NR	Age: NR Post: 100%	100% ER+	75% LN− 25% LN+	ET

Author, Year	Study Design	Country	Overall N	Risk Category Cut-Offs	Risk Categories, N (%)	Age, Years Menopausal Status (Pre/Peri/Post), %	Hormone Receptor	Lymph Node	Treatment
Sestak et al, 2018[52]	Retrospective analysis of RCT (TransATAC)	United Kingdom	774	*EndoPredict* Low: EPclin < 3.3 High: EPclin ≥ 3.3 *Oncotype DX* Low: ≤ 17 Int: 18–31 High: ≥ 32 *Prosigna* Low: ≤ 26 Int: 27–68 High: ≥ 69	*EndoPredict* Low: 472 (61) High: 302 (39) *Oncotype DX* Low: 479 (62) Int: 214 (28) High: 81 (11) *Prosigna* Low: 333 (43) Int: 236 (31) High: 205 (27)	Mean: 64 (SD 8.0) Post: 100%	100% ER+ 100% HER2−	76% LN− 24% LN+	ET
Tsai et al, 2018[109]	Prospective observational study (PROMIS)	United States	840	*MammaPrint* Low: 0 to 1 High: −1 to 0 *Oncotype DX* Int: 18–30	*MammaPrint* Low: 374 (45) High: 466 (55) *Oncotype DX* Int: 840 (100)	Mean: 59 (range 27–93) Pre/Peri: 23% Post: 72% Unknown: 5%	99.9% ER+ 86% PR+ 98% HER2−	71% LN− 27% LN+ (including N1mic) 1% unknown	ET, ET + CT

Abbreviations: ABCSG, Austrian Breast and Colorectal Cancer Study Group; BCIST, Breast Cancer Intrinsic Subtype; CT, chemotherapy; DBCG, Danish Breast Cancer Group; EP, EndoPredict; EPclin, EndoPredict clinical score; ER, estrogen receptor; ET, endocrine therapy; GEICAM, Spanish Breast Cancer Group; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; int, intermediate; IQR, interquartile range; LN, lymph node; MINDACT, Microarray in Node-Negative and 1 to 3 Positive Lymph Node Disease May Avoid Chemotherapy; N1mic, micrometastasis; NCIC, National Cancer Institute of Canada; NR, not reported; NSABP, National Surgical Adjuvant Breast and Bowel Project; PACS, Adjuvant Treatment of Breast Cancer; peri, peri-menopausal; PONDx, Prospective multicenter study of the ONcotype DX test; post, postmenopausal; PR, progesterone receptor; pre, premenopausal; PRIMe, PRospective study to measure the Impact of MammaPrint on adjuvant treatment in hormone receptor-positive HER2-negative breast cancer patients; PROMIS, Prospective Study of MammaPrint in Breast Cancer Patients With an Intermediate Recurrence Score; RASTER, Microarray Prognostics in Breast Cancer; RCT, randomized controlled trial; SAKK, Swiss Group for Clinical Cancer Research; SD, standard deviation; SEER, Surveillance, Epidemiology, and End Results; STO, Stockholm Tamoxifen; TAILORx, Trial Assigning IndividuaLized Options for Treatment (Rx); SWOG, Southwest Oncology Group; TransATAC, Translational Study of Anastrozole or Tamoxifen Alone or Combined; WSG, West German Study Group.

Note: Some percentage totals may appear inaccurate because of rounding and missing reported data.

Table A8: Quantitative Evidence of Preferences and Values of Patients and Providers—Characteristics of Included Studies

Author, Year	Country	N	Study Design/ Methods	Participants	Test
Brewer et al, 2009[202]	United States	165 patients	Health literacy assessment and questionnaire	Previously diagnosed with stage 1 or 2 primary breast cancer; completed surgery; had or had not received chemotherapy or currently receiving hormone therapy (tamoxifen); English-speaking Exclusion: cancer recurrence, life-threatening comorbid disease, second primary cancer diagnosis, metastasis, history of serious psychiatric illness, non-English-speaking	Unspecified GEP test, but roughly equivalent to Oncotype DX
Brewer et al, 2009[201]	United States	165 patients	6 hypothetical vignettes	Same as Brewer et al, 2009[202]	Oncotype DX
Davidson et al, 2013[137]	Canada	156 patients	Questionnaire	Stage 1 or 2 breast cancer, ER+, HER2−, $pN0_{IHC}$+, completed surgery	Oncotype DX
DeFrank et al, 2013[208]	United States	77 patients	6 hypothetical vignettes	Stage 1 or 2 breast cancer, ER/PR+, mixed LN status Exclusion: < 18 years old, non-English-speaking	Oncotype DX
DeFrank et al, 2013[209]	United States	132 patients	Questionnaire and medical chart review for all but 6 people	Stage 1 or 2 breast cancer, ER/PR+, mixed LN status Exclusion: < 18 years old, non-English-speaking, incarcerated, second primary cancer diagnosis or other life-threatening comorbid disease, history of serious psychiatric illness	Oncotype DX
Eiermann et al, 2013[83]	Germany	366 patients	Questionnaire	Operable early breast cancer; ER+, HER2−, tumour size ≥ 1 cm or < 1 cm if at least 1 unfavourable characteristic; mixed LN status; ≥ 18 years old; good performance status; no contraindication for systemic chemoendocrine therapy	Oncotype DX
Evans et al, 2016[110]	United States	193 patients	Interview and questionnaire	Recently diagnosed with breast cancer, stage 1 or 2, ER/PR+, received Oncotype DX Exclusion: prior cancer diagnosis, people who initiated chemotherapy or received RS prior to pre-test interview	Oncotype DX
Fallowfield et al, 2018[105]	United Kingdom	149 patients	Consultation with oncologist, questionnaire	Early-stage breast cancer, ER+, HER2−, equivocal indications for chemotherapy, able to read English	EndoPredict
Friese et al, 2017[211]	United States	1,527 patients	SEER registries of Los Angeles county and Georgia	Exclusion: stage 3 or 4 cancer, tumours > 5 cm, LN+ (> 3)	Oncotype DX
Gligorov et al, 2015[212]	France	1,000 patients 94 physicians	Questionnaire	Operable invasive early-stage breast cancer; ER+, HER2−, pN0 or pN1mi; potential candidate for systemic chemotherapy, good performance status; ≥ 18 years old	Oncotype DX
Hequet et al, 2017[88]	France	200 patients NR physicians	Questionnaire	Postmenopausal people; early-stage invasive breast cancer, T1–T2, LN−; no contraindication for adjuvant chemotherapy; ECOG score 0 or 1; ability to complete questionnaire without assistance	Prosigna

Author, Year	Country	N	Study Design/ Methods	Participants	Test
Holt et al, 2013[143]	United Kingdom	146 patients	Questionnaire	ER+, pN0, pN0i+, pN1mi Exclusion: < 18 years old, unable to comprehend details of trial, unable to complete documentation in English, previous history of breast cancer treatment	Oncotype DX
Kuchel et al, 2016[111]	United Kingdom	137 patients	Questionnaire	ER+, HER2−, LN− or micrometastases if ≤ 50 years old, LN+ (1–3 if > 50 years old), ECOG 0 or 1, fit for chemotherapy	Oncotype DX
Kurian et al, 2018[213]	United States	2,926 patients 304 oncologists	Questionnaire, identified through SEER	Stage 1 to 2, ER+, HER2−	Oncotype DX
Levine et al, 2016[84]	Canada	1,000 patients	Consultation with oncologist, questionnaire	ER+, LN− or micrometastases; had surgery; receiving or intend to receive chemotherapy; considered for chemotherapy Exclusion: inoperable, locally advanced, or metastatic breast cancer; previous neoadjuvant chemotherapy; HER2+	Oncotype DX
Lillie et al, 2007[214]	United States	163 patients	Health literacy assessment, interview	Previously had stage 1 or 2 primary breast cancer; had surgery; post-treatment people who did not receive neoadjuvant/adjuvant chemotherapy or who had completed it; currently receiving hormone therapy (tamoxifen) Exclusion: non-English-speaking, life-threatening comorbid disease, second primary cancer diagnosis, metastasis, history of serious psychiatric illness, no previous cancer recurrence	Oncotype DX
Lipkus et al, 2011[215]	United States	64 patients	Questionnaire	≥ 18 years old, Oncotype DX testing, can speak and read English, mailing address and working telephone number	Oncotype DX
Lo et al, 2010[85]	United States	93 patients 17 medical oncologists	Questionnaire	ER+, LN−, fit to receive chemotherapy	Oncotype DX
Martin et al, 2015[203]	Spain	200 patients	Questionnaire	Postmenopausal, ER+, HER2−, T1 or T2 tumours (< 5 cm), LN−, no metastasis, ECOG score 0 or 1, no contraindications for adjuvant chemotherapy	Prosigna
Murciano-Goroff et al, 2018[216]	United States	732 medical oncologists and surgeons	Questionnaire	From Florida, New Jersey, or New York; listed in 2010 AMA provider database or identified by breast cancer patients recruited from state cancer registries	Oncotype DX
Ngoi et al, 2013[217]	Singapore	200 patients 67 cancer physicians (medical oncologists, surgeons, radiation oncologists, others)	Questionnaire	Patients: previously had stage 0 to 3 breast cancer, ≥ 21 years old, attending tertiary cancer centre Cancer physicians: managed breast cancer patients from tertiary institutions, community-bated hospitals, or private practice	Oncotype DX
O'Neill et al, 2007[200]	United States	139 patients	Questionnaire	Same as Lillie et al, 2007[214]	Any GEP test
Ozmen et al, 2016[87]	Turkey	165 patients NR physicians	Questionnaire	pT1–3, pN0 or pN1mic, M0, ER+, HER2−	Oncotype DX

Author, Year	Country	N	Study Design/ Methods	Participants	Test
Panattoni et al, 2019[210]	United States	833 patients	Questionnaire	Previously had early-stage breast cancer, ER+, stage 1 or 2, LN−, HER2−, 25–74 years old at diagnosis	Any GEP test
Patil et al, 2015[218]	United States	119 oncologists (medical or surgical)	Questionnaire	From national physician panel treating breast cancer patients	Oncotype DX
Retel et al, 2009[205]	Netherlands	77 patients	Questionnaire	Early-stage invasive breast cancer, pT1–2, LN−, < 55 years old; received local therapy Exclusion: prior malignancies (except basal cell carcinomas and cervical dysplasia), bilateral breast cancer	MammaPrint
Retel et al, 2013[204]	Netherlands	347 patients	Questionnaire	Early-stage breast cancer, LN 0–3, able to read or write in English or Dutch	MammaPrint
Richman et al, 2011[219]	United States	78 patients	Questionnaire	Previously treated for early-stage breast cancer, had Oncotype DX test, stage 1 or 2, ER/PR+, LN− (1 person was LN+)	Oncotype DX
Seror et al, 2013[220]	France	43 patients	Questionnaire	LN+ treated with anthracycline-based chemotherapy without taxane	Any GEP test
Sulayman et al, 2012[221]	United States	81 patients	Questionnaire	Had Oncotype DX, completed chemotherapy, no recurrence of primary breast cancer or second cancer, no major comorbid disease or participated in clinical trials that would affect treatment decision	Oncotype DX
Torres et al, 2018[103]	Canada	71 patients	Questionnaire	Women > 18 years old, T1–3, LN+ (1–3), M0, ER+, HER2−, ECOG 0 or 1, candidates for chemotherapy where benefit of adding chemotherapy to hormonal therapy was unclear or not large enough to warrant risk of chemotherapy Exclusion: micrometastases only	Oncotype DX
Tzeng et al, 2010[193]	United States	77 patients	Questionnaire, medical chart review if patient consented	Previously treated for early-stage breast cancer, had Oncotype DX test, stage 1 or 2, LN− (1 person was LN+), ER/PR+ Exclusion: < 18 years old, non-English-speaking	Oncotype DX
Wuerstlein et al, 2016[89]	Germany	198 patients NR physicians	Questionnaire	Postmenopausal people, ER+, HER2−, LN−, pT1–2, M0	Prosigna
Wuerstlein et al, 2019[49]	Germany	430 patients NR physicians	Questionnaire	≥ 18 years old, pT1–3, pN0–1, ER/PR+, HER2− Exclusion: ≥ 4 LN involvement, multicentric or metastatic disease, prior malignancies within past 5 years	MammaPrint

Author, Year	Country	N	Study Design/ Methods	Participants	Test
Yamauchi et al, 2014[222]	Japan	124 patients 17 medical oncologists or surgeons	Questionnaire	Patients: operable breast cancer, ER+, HER2−, N0 for pre- or postmenopausal people, micrometastatic disease for postmenopausal people, LN 1–3 in postmenopausal people, ≥ 18 years old, adequate performance status, candidate for chemotherapy, answer written questions in Japanese Physicians: medical oncologist or surgeon making adjuvant treatment recommendations to breast cancer patients; at least 1 physician of participating centre had to have ordered Oncotype DX	Oncotype DX

Abbreviations: AMA, American Medical Association; ECOG, Eastern Cooperative Oncology Group; ER, estrogen receptor; GEP, gene expression profiling; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; LN, lymph node; M, metastasis; mi, micrometastasis; mic, micrometastasis; N, node; NR, not reported; p, histopathology; PR, progesterone receptor; RS, Recurrence Score; SEER, Surveillance, Epidemiology, and End Results; T, tumour.

Appendix 6: Critical Appraisal of Clinical Evidence

Table A9: Risk of Bias[a] Among Prognostic Studies (PROBAST)

	Risk of Bias				Applicability			Overall	
Author, Year	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
Albain et al, 2010[90]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Buus et al, 2016[61]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Dowsett et al, 2010[68]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Dowsett et al, 2013[57]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Drukker al, 2013[63]	High[c]	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Drukker et al, 2014[64]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Dubsky et al, 2013[91]	Low	Low	Low	High[b,d]	Low	Low	Low	Low	Low
Esserman et al, 2017[65]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Filipits et al, 2014[54]	High[c]	Low	Low	High[b]	Low	Low	Low	Low	Low
Filipits et al, 2019[62]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Fitzal et al, 2015[92]	High[c]	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Gluz et al, 2016[94]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Gnant et al, 2014[73]	Low	Low	Low	Low	Low	Low	Low	Low	Low
Gnant et al, 2015[260]	Low	Low	Low	Low	Low	Low	Low	Low	Low
Ibraheem et al, 2019[75]	High[c]	Low	Low	High[b]	Low	Low	Low	Low	Low
Laenkholm et al, 2018[74]	High[c]	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Liu et al, 2015[58]	Low	Low	Low	High[b,d]	Low	Low	Low	Unclear	Low
King et al, 2016[95]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Mamounas et al, 2010[69]	High[c]	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Mamounas et al, 2017[96]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Martin et al, 2016[93]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Mokbel et al, 2017[106]	High[c]	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Mokbel et al, 2018[107]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Nitz et al, 2017[55]	High[c]	Low	Low	Unclear	Low	Low	Low	Unclear	Low

Author, Year	Risk of Bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
Ohnstad et al, 2017[59]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Paik et al, 2004[70]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Paik et al, 2006[56]	High[c]	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Penault-Llorca et al, 2018[97]	Low	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Petkov et al, 2016[71]	High[c]	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Roberts et al, 2017[98]	High[c]	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Sestak et al, 2013[100]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Sestak et al, 2015[60]	High[c]	Low	Low	High[b]	Low	Low	Low	Unclear	Low
Sestak et al, 2018[52]	High[c]	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Sparano et al, 2015[259]	Low	Low	Low	Low	Low	Low	Low	Low	Low
Stemmer at al, 2017[72]	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Stemmer et al, 2017[99]	Low	Low	Low	Unclear	Low	Low	Low	Low	Low
Tsai et al, 2018[109]	Low	Low	Low	High[b,d]	Low	Low	Low	Low	Low
van de Vijver et al, 2002[66]	High[c]	Low	Low	High[b]	Low	Low	Low	Unclear	Low
van't Veer et al, 2017[67]	Low	Low	Low	High[b]	Low	Low	Low	Low	Low

Abbreviation: PROBAST, Prediction Model Risk of Bias Assessment Tool.
[a]Possible risk of bias levels: low, high, unclear.
[b]Some studies did not perform multivariate analyses to consider other potential confounding factors, or unclear how multivariables were chosen for the analysis.
[c]Unclear how patients were chosen and enrolled in study (e.g., if patient enrolment was consecutive).
[d]Selective reporting concerns where study did not report all preplanned or subgroup analyses.

Table A10: Risk of Bias[a] 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
Cardoso et al, 2016[51]	Unclear	Unclear[b]	Low	Low	Low	Low[c]
Sparano et al, 2018[26]	Unclear	Unclear[b]	Low	High[d]	Unclear	Low[c]

[a]Possible risk of bias levels: low, high, and unclear.
[b]Unclear method and process of randomization.
[c]Noninferiority randomized controlled trials because of ethical concerns related to withholding chemotherapy for distant recurrence.
[d]Selective reporting concerns where study did not report all preplanned or subgroup analyses.

Table A11: Risk of Bias[a] Among Nonrandomized Studies (RoBANS)

Author, Year	Selection of Participants	Confounding Variables	Measurement of Exposure	Blinding of Outcome Assessments	Incomplete Outcome Data	Selective Outcome Reporting
Albain et al, 2010[90]	High[b]	Low	Low	Unclear[c]	High[d]	Low
Albanell et al, 2012[77]	Low	High[e]	Low	Unclear[c]	Low	High[e]
Albanell et al, 2016[79]	Low	Low	Low	Unclear[c]	Low	Low
Bargallo et al, 2015[80]	Low	High[e]	Low	Unclear[c]	Low	High[e]
Curtit et al, 2019[48]	High[b]	High[e]	Low	Unclear[c]	High[d]	High[e]
Cusumano et al, 2014[108]	High[b]	Low	Low	Unclear[c]	Low	Low
de Boer et al, 2013[81]	Low	High[e]	Low	Unclear[c]	Low	High[e]
Dieci et al, 2018[82]	High[b]	High[e]	Low	Unclear[c]	High[d]	Low
Eiermann et al, 2013[83]	High[b]	High[e]	Low	Unclear[c]	Low	High[e]
Ettl et al, 2017[104]	High[b]	High[e]	Low	Unclear[c]	Low	High[e]
Fallowfield at al, 2017[105]	Low	Unclear	Low	Unclear[c]	Low	High[e]
Geyer et al, 2018[76]	High[b]	High[e]	Low	Unclear[c]	High[d]	Low
Hequet et al, 2017[88]	Low	High[e]	Low	Unclear[c]	Low	High[e]
Jensen et al, 2018[101]	High[b]	Unclear	Low	Unclear[c]	Low	Low
Kuchel et al, 2016[111]	Low	Unclear	Low	Unclear[c]	Low	High[e]
Kuijer et al, 2017[78]	High[b]	Low	Low	Unclear[c]	High[d]	High[e]
Levine et al, 2016[84]	High[b]	High[e]	Low	Low	Low	Low
Leung et al, 2016[112]	High[b]	Unclear	Low	Unclear[c]	High[d]	Low
Lo et al, 2010[85]	Low	High[e]	Low	Unclear[c]	High[d]	High[e]
Loncaster et al, 2017[86]	High[b]	Unclear	Low	Unclear[c]	Low	Low
Martinez del Prado et al, 2018[113]	High[b]	High[e]	Low	Low	High[d]	Unclear
Ozmen et al, 2016[87]	High[b]	High[e]	Low	Unclear[c]	Low	Low
Paik et al, 2006[56]	Low	Unclear	Low	Unclear[c]	High[d]	Low
Pestalozzi et al, 2014[114]	High[b]	High[e]	Low	Unclear[c]	Low	Low
Sestak et al, 2019[102]	Low	High[e]	Low	Unclear[c]	Low	Low
Torres et al, 2018[103]	High[b]	Unclear	Low	Unclear[c]	Low	Low

Author, Year	Selection of Participants	Confounding Variables	Measurement of Exposure	Blinding of Outcome Assessments	Incomplete Outcome Data	Selective Outcome Reporting
Voelker et al, 2018[115]	Low	High[e]	Low	Unclear[c]	Low	Low
Wuerstlein et al, 2016[89]	High[b]	High[e]	Low	Unclear[c]	Low	High[e]
Wuerstlein et al, 2019[49]	Low	High[e]	Low	Unclear[c]	Low	High[e]

Abbreviation: RoBANS, Risk of Bias Assessment for Nonrandomized Studies.
[a]Possible risk of bias levels: low, high, and unclear.
[b]Unclear how patients were chosen and enrolled in study (e.g., if patient enrolment was consecutive).
[c]Unclear blinding of study assessors, but likely had limited impact on bias because gene expression profiling tests are objective tests.
[d]Incomplete data/selective reporting concerns where study did not report all preplanned or subgroup analyses.
[e]Some studies did not perform multivariate analyses to consider other potential confounding factors, or unclear how multivariables were chosen for the analysis.

Table A12: GRADE Evidence Profile for GEP Tests (EndoPredict, MammaPrint, Oncotype DX, and Prosigna)

Number of Studies (Design)	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Upgrade Considerations	Quality
Lymph-Node–Negative Population							
Prognostic Ability							
Freedom from distant recurrence: 20 studies (EndoPredict: 3 studies, MammaPrint: 5 studies, Oncotype DX: 8 studies, Prosigna: 5 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	No serious limitations	Undetected	None	⊕⊕⊕ Moderate
Disease-free survival: 6 studies (MammaPrint: 3 studies, Oncotype DX: 3 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕⊕ Low
Overall survival: 5 studies (MammaPrint: 1 study, Oncotype DX: 4 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕⊕ Low
Predictive Ability							
Freedom from distant recurrence: 2 RCTs (MammaPrint: 1 study, Oncotype DX: 1 study; 6 observational studies)	Serious limitations (−1)[a]	No serious limitations	Serious limitations (−1)[b]	No serious limitations	Undetected	None	⊕⊕ Low

Number of Studies (Design)	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Upgrade Considerations	Quality
Disease-free survival: 2 RCTs (MammaPrint: 1 study, Oncotype DX: 1 study) 2 observational studies: (Oncotype DX)	Serious limitations (−1)[a]	No serious limitations	Serious limitations (−1)[b]	No serious limitations	Undetected	None	⊕⊕ Low
Overall survival: 2 RCTs (MammaPrint: 1 study, Oncotype DX: 1 study) 1 observational study: (Oncotype DX)	Serious limitations (−1)[a]	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕⊕ Low
Changes in treatment decision-making: 13 observational studies (MammaPrint: 1 study; Oncotype DX: 10 studies, Prosigna: 2 studies)	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	None	⊕⊕ Low
Physician confidence: 6 observational studies (Oncotype DX: 4 studies, Prosigna; 2 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	No serious limitations	Undetected	None	⊕ Very low
Lymph-Node–Positive Population							
Prognostic Ability							
Freedom from distant recurrence: 19 studies (EndoPredict: 4 studies, Oncotype DX: 7 studies; Prosigna: 5 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕⊕ Low
Disease-free survival: 3 studies (Oncotype DX)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕ Very Low
Overall survival: 5 studies (MammaPrint: 1 study, Oncotype DX: 4 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕ Very Low
Predictive Ability							
Freedom from distant recurrence: 1 RCT (MammaPrint)	Serious limitations (−1)[a]	No serious limitations	Serious limitations (−1)[b]	No serious limitations	Undetected	None	⊕ Very low
Disease-free survival:1 observational study (Oncotype DX)	Serious limitations (−1)[a]	No serious limitations	Serious limitations (−1)[b]	No serious limitations	Undetected	None	⊕ Very low

Number of Studies (Design)	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Upgrade Considerations	Quality
Overall survival: 3 observational studies (Oncotype DX)	Serious limitations (−1)[a]	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕ Very low
Changes in treatment decision-making: 19 observational studies (EndoPredict: 4 studies; MammaPrint: 2 studies, Oncotype DX: 13 studies)	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	None	⊕⊕ Low
Physician confidence in treatment recommendations: 6 observational studies (MammaPrint: 1 study, Oncotype DX: 5 studies)	No serious limitations	No serious limitations	Serious limitations (−1)[b]	Serious limitations (−1)[c]	Undetected	None	⊕ Very Low

Abbreviations: GEP, gene expression profiling; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; RCT, randomized controlled trial.

[a]Ideal study design for predictive ability is RCT. Nonrandomized design of some included studies may have led to cofounding and additional biases.

[b]Heterogeneity concerns regarding included patient population and the generalizability of patients between studies.

[c]Studies that reported on overall survival were not powered to detect long-term differences.

[d]Uncertainty because of unclear reporting of how discussions regarding treatment discussions occurred. Other non-testing factors may have influenced treatment decision changes.

Appendix 7: Potentially Relevant Ongoing Studies

Table A13: List of Potentially Relevant Ongoing Studies from ClinicalTrials.gov

Clinical Trial No.	Study Official Name	GEP Test	Estimated Date of Completion
NCT03749421	Prospective study of the Prosigna assay on neoadjuvant clinical decision-making in women with HR+/HER2− breast cancer	Prosigna	January 2022
NCT01479101	Prospective neoadjuvant registry trial linking MammaPrint, subtyping, and treatment response: neoadjuvant breast registry—Symphony Trial (NBRST)	MammaPrint	January 2020
NCT00433589	MINDACT (Microarray In Node-Negative and 1 to 3 Positive Lymph Node Disease May Avoid Chemotherapy): a prospective, randomized study comparing the 70-gene signature with the common clinical-pathological criteria in selecting patients for adjuvant chemotherapy in breast cancer with 0 to 3 positive nodes	MammaPrint	March 2020
NCT01272037	A phase III, randomized clinical trial of standard adjuvant endocrine therapy ± chemotherapy in patients with 1–3 positive nodes, hormone receptor-positive and HER2-negative breast cancer with Recurrence Score (RS) of 25 or less. RxPONDER: a clinical trial treatment for positive-node, endocrine-responsive breast cancer	Oncotype DX	February 2022
NCT04016935	EndoPredict extended endocrine trial (EXET): evaluation of impact of EndoPredict on treatment decision	EndoPredict	September 2022
NCT03969121	Neoadjuvant hormonal therapy plus palbociclib in operable, hormone-sensitive, and HER2-negative primary breast cancer	EndoPredict	May 2021
NCT03503799	Reaching for Evidence-baSed Chemotherapy Use in Endocrine Sensitive Breast Cancer (RESCUE)	EndoPredict	May 2031
NCT01805271	Safety study of adding everolimus to adjuvant hormone therapy in women with high risk of relapse, ER+ and HER2− primary breast cancer, free of disease after receiving at least 1 year of adjuvant hormone therapy	EndoPredict	June 2031

Abbreviations: GEP, gene expression profiling; HER2, human epidermal growth factor receptor 2; HR, hormone receptor.

Appendix 8: Results of Applicability and Limitation Checklists for Studies Included in the Economic Literature Review

Table A14: Assessment of the Applicability of Studies Evaluating the Cost-Effectiveness of GEP Tests in Canadian Setting

Author, Year	Is the study population similar to the question?	Are the interventions similar to the question?	Is the health care system studied sufficiently similar to Ontario?	Were the perspectives clearly stated? If yes, what were they?	Are all direct effects included? Are all other effects included where they are material?	Are all future costs and outcomes discounted? If yes, at what rate?	Is the value of health effects expressed in terms of quality-adjusted life-years?	Are costs and outcomes from other sectors fully and appropriately measured and valued?	Overall Judgment[a]
Davidson et al, 2013[137]	Yes	Yes	Yes	Yes, third-party direct payer	Partially	Yes, 5%	Yes	Yes	Directly applicable
Hannouf et al, 2019[175]	Yes	Yes	Yes	Yes, Canadian public health care system	Partially	Yes, 1.5%	Yes	Yes	Directly applicable
Hannouf et al, 2014[142]	Partially yes (people with 1–3 LN+ breast cancer)	Yes	Yes	Yes, Canadian public health care system	Partially	Yes, 5%	Yes	Yes	Partially applicable
Hannouf et al, 2012[141]	Yes	Yes	Yes	Yes, Canadian public health care system	Partially	Yes, 5%	Yes	Yes	Directly applicable
HQO, 2010[129]	Yes	Partially[b]	Yes	Yes, Ontario public payer	Partially	Yes, 5%	Yes	Yes	Partially applicable
Lamond et al, 2012[153]	Yes	Yes	Yes	Yes, third-party direct payer	Partially	Yes, 3%	Yes	Yes	Directly applicable
Paulden et al, 2013[158]	Yes	Partially[b]	Yes	Yes, Ontario public payer	Partially	Yes, 5%	Yes	Yes	Partially applicable
Tiwana et al, 2013[168]	Yes	Partially[b]	Yes	Unclear, appeared to be payer perspective	Partially	Yes, 5%	Yes	Partially	Partially applicable
Tsoi et al, 2010, Canada[169]	Yes	Partially[b]	Yes	Yes, health care payer	Partially	Yes, 5%	Yes	Yes	Partially applicable

Abbreviations: GEP, gene expression profiling; HQO, Health Quality Ontario; LN, lymph node.
Note: Response options for all items were "yes," "partially," "no," "unclear," and "NA" (not applicable).
[a]Overall judgment may be "directly applicable," "partially applicable," or "not applicable."
[b]Compared with Adjuvant! Online, which is no longer available in practice.

Table A15: Assessment of the Limitations of Studies Evaluating the Cost-Effectiveness of GEP Tests in Canadian Setting

Author, Year	Does the model structure adequately reflect the nature of the health condition under evaluation?	Is the time horizon sufficiently long to reflect all important differences in costs and outcomes?	Are all important and relevant health outcomes included?	Are the clinical inputs[a] obtained from the best available sources?	Do the estimates of relative treatment effect match the estimates contained in the clinical sources?	Are all important and relevant (direct) costs included in the analysis?	Are the estimates of resource use obtained from best available sources?	Are the unit costs of resources obtained from best available sources?	Is an appropriate incremental analysis presented or can it be calculated from the reported data?	Are all important and uncertain parameters subjected to appropriate sensitivity analysis?	Is there a potential conflict of interest?	Overall Judgment[b]
Davidson et al, 2013[137]	Yes	Yes, lifetime horizon	Yes	Partially (actual clinical data from 150 patients from 2 participating BCCA centres)	NA	Yes	Yes (actual cost from 150 patients from 2 participating BCCA centres)	Yes	Yes	Yes	Unclear	Minor limitations
Hannouf et al, 2019[175]	Yes	Yes, lifetime horizon	Yes	From TransATAC study[52]	NA	Yes	Yes (local unit costs at London Regional Cancer Program, London, Ontario)	Yes	Yes	Yes	Unclear	Minor limitations
Hannouf et al, 2012[141]	Yes	Yes, lifetime horizon	Yes	Partially (NSABPB-14, NSABPB-20 and 7 years of follow-up data from the Manitoba Cancer Registry)[56,70]	NA	Yes	Yes (market price for the test, treatment costs from Manitoba health databases)	Yes	Yes	Yes	Unclear	Minor limitations
Lamond et al, 2012[153]	Yes	Yes, 25-year time horizon	Yes	Partially (NSABPB-20 and SWOG 8814)[56,90]	NA	Yes	Yes (local unit costs at the Queen Elizabeth II Health Sciences Centre, in Halifax, Nova Scotia)	Yes	Yes	Yes	Unclear	Minor limitations

Abbreviations: BCCA, British Columbia Cancer Agency; GEP, gene expression profiling; NSABPB, National Surgical Adjuvant Breast and Bowel Project clinical trial B; SWOG, Southwest Oncology Group trial.
Note: Response options for all items were "yes," "partially," "no," "unclear," and "NA" (not applicable).
[a]Clinical inputs include relative treatment effects, natural history, and utilities.
[b]Overall judgment may be "minor limitations," "potentially serious limitations," or "very serious limitations."

Table A16: Results of the Cost-Effectiveness of GEP Tests in Other Settings

Study ID	Comparator	Patient Group	Country	Perspective	Time Horizon	ICER/Economic Conclusion
			MammaPrint			
Lymph-Node–Negative						
Bonastre, et al, 2014[134]	AOL; chemotherapy for all	LN−	France	French national insurance scheme	10 years	Compared with AOL: €134,000/QALY Compared with chemotherapy for all: €84,000/QALY
Chen et al, 2010[135]	AOL	LN−	United States	Health care payer	Lifetime	USD$10,000/QALY
Exner et al, 2014[139]	Usual care	LN−, HR+, HER2−	Netherlands	NR	20 years	Dominant
Kondo et al, 2012[150]	Best practice (St. Gallen)	ER+, LN−, HER2−	Japan	Health care system, but presented as societal	10 years	USD$43,044/QALY
Retèl et al, 2010[162]	St Gallen; AOL	ER+, LN−	Netherlands	Health care payer	20 years	Compared with St Gallen: MammaPrint dominated Compared with AOL: €4,614/QALY
Retèl et al, 2013[160]	AOL	ER+, LN− (after local therapy)	Netherlands	Societal	20 years	70G-FFT: €5,247/QALY 70G-PAR: €6,200/QALY
Retèl et al, 2013[161]	AOL	ER+, LN−	Netherlands	Dutch health care	20 years	Dominant
Ward et al, 2013[172]	MammaPrint for all compared with current clinical practice; MammaPrint (NPI > 3.4) compared with current clinical practice	ER+, LN−, HER2−	United Kingdom	National Health Service and Personal Social Services	Lifetime (up to age 100 years)	MammaPrint for all: £12,240–£53,058/QALY MammaPrint for NPI > 3.4: £6,053–£29,569/QALY
Lymph-Node–Positive (or Mixed)						
NICE, 2018[36]	Current practice	Early-stage breast cancer	United Kingdom	National Health Service and Personal Social Services	Lifetime	For MINDACT trial population: £131,482/QALY For MINDACT mAOL high-risk subgroup: dominated For MINDACT mAOL low-risk subgroup: £414,202
Oestreicher et al, 2005[157]	Best practice	N ≥ 0 stage ≤ II pre-menopausal (LN+ 51%, ER+ 77%)	United States	Societal	Lifetime	USD$13,724/QALY (in favour of best practice)
Stein et al, 2016[166]	Chemotherapy for all	ER+, HER2-	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	£1,083/QALY

Study ID	Comparator	Patient Group	Country	Perspective	Time Horizon	ICER/Economic Conclusion
			Oncotype DX Compared With No Genomic Testing			
Lymph-Node–Negative						
Chandler et al, 2018[120]	Usual care	ER+, LN−, HER2-	United States	Societal	25 years	USD$188,125/QALY
Cosler et al, 2009[136]	Tamoxifen; chemotherapy and tamoxifen	ER+, LN−	United States	Health care payer	20 years	Compared with tamoxifen: dominated Compared with chemotherapy and tamoxifen: USD$4,432/QALY
Holt et al, 2013[143]	Usual care	ER+, LN−-1	United Kingdom	National Health Service	30 years	£6,232/QALY
Hornberger et al, 2005[145]	Usual care	ER+, LN−	United States	Societal	Lifetime	Dominant
Hornberger et al, 2011[144]	Best practice	ER+, LN−	United States	Health care payer	Lifetime	Dominant
Jahn et al, 2015[146]	AOL	HR+, LN−, HER2−	Austria	Societal perspective in line with the Austrian health care system	Lifetime	€5,978/QALY
Katz et al, 2015[147]	Usual care	HR+, LN−, HER2−	France	Societal perspective and health care payer perspective	30 years	Dominant
Klang et al, 2010[149]	Usual care	ER+, LN−	Israel	Health care payer	30 years	USD$10,770/QALY
Kondo et al, 2008[151]	Best practice (St. Gallen, NCCN guidelines)	HR+, LN−	Japan	Health care payer	Lifetime	Compared with St. Gallen guideline: USD$10,774/QALY Compared with NCCN guideline: USD$26,065/QALY
Kondo et al, 2011[152]	Best practice (St. Gallen)	ER+, LN−	Japan	Health care system although presented as societal	Lifetime (with assumptions about maximum survival after 10 1-year cycles)	USD$3,848/QALY
Lyman et al, 2007[154]	Tamoxifen; tamoxifen and chemotherapy	ER+, LN−	United States	Societal	Not reported, likely remaining life expectancy	Compared with tamoxifen: USD$4,432/QALY Compared with tamoxifen and chemotherapy: Oncotype dominant
Mislick et al, 2014[155]	Mammostrat vs. Oncotype DX	ER+. LN−	United States	Third party payer	10 years	USD$453,600/QALY (in favour of Mammostrat)

Study ID	Comparator	Patient Group	Country	Perspective	Time Horizon	ICER/Economic Conclusion
NICE, 2018[36]	Current practice	LN−, NPI ≤ 3.4; LN−, NPI > 3.4	United Kingdom	National Health Service and Personal Social Services	Lifetime	LN−, NPI ≤ 3.4: £122,725/QALY LN−, NPI > 3.4: dominated
Reed et al, 2013[159]	AOL	ER+, LN−	United States	United States health-system perspective and societal perspective	Lifetime (10 or to recurrence)	USD$10,788/QALY
Vataire et al, 2012[171]	Usual care	ER+, LN−, HER2−	France	Societal	30 years	Dominant
Wang et al, 2019[128]	No patients receiving chemotherapy	ER+, LN−, HER2−	United States	United States payer	Lifetime	Chemotherapy for those with Recurrence Score ≥ 31: USD$62,200/QALY Chemotherapy for those with Recurrence Score ≥ 18: USD$118,400/QALY
Ward et al, 2013[172]	Oncotype DX for all compared with current clinical practice; Oncotype DX for NPI > 3.4 compared with current clinical practice	ER+, LN−, HER2−	United Kingdom	National Health Service and Personal Social Services	Lifetime (up to age 100 years)	Oncotype DX for all: £29,502/QALY Oncotype DX for NPI > 3.4: £9,774/QALY
Ward et al, 2013[172]	Oncotype DX vs. IHC4 for all; Oncotype DX vs. IHC4 for NPI > 3.4	ER+, LN−, HER2−	United Kingdom	National Health Service and Personal Social Services	Lifetime (up to age 100 years)	Compared with IHC4 for all: £64,111/QALY Compared with IHC4 for NPI > 3.4: £31,125/QALY
Yamauchi et al, 2014[173]	Usual care	ER+, LN−	Japan	Societal	Lifetime	USD$6,368/QALY
Lymph-Node–Positive (or Mixed)						
Bargalló-Rocha et al, 2015[131]	Usual care	HR+, LN3, HER2−	Mexico	Instituto Mexicano del Seguro Social perspective	40 years	USD$1,914/LY
Blohmer et al, 2013[133]	Usual care	ER+, LN3, HER2−	Germany	Health care payer	30 years	Dominant
Hall et al, 2012[140]	Chemotherapy	ER+, LN+	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	£5529/QALY
Ibarrondo et al, 2018[123]	Usual care	HR+, size LN− or LN1mi, HER2−	Basque Country (Spain)	Health service perspective; Social perspective	Lifetime	Health service perspective with discount: €17,453/QALY Social perspective with discount: dominant

Study ID	Comparator	Patient Group	Country	Perspective	Time Horizon	ICER/Economic Conclusion
Kip et al, 2015[148]	Usual care	ER+, LN1	Netherlands	Dutch health care payer's perspective	30 years	€11,236/QALY
Kondo et al, 2011[152]	Best practice	ER+, LN+	Japan	Health care system although presented as societal	Lifetime (with assumptions about max survival after 10 1-year cycles)	USD$5,685/QALY
NICE, 2018[36]	Current practice	LN+ (1-3 nodes)	United Kingdom	National Health Service and Personal Social Services	Lifetime	Dominated
Stein et al, 2016[166]	Chemotherapy for all	ER+, HER2-	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	Dominant
Vanderlaan et al, 2011[170]	Best practice (NCCN guidelines)	ER+, LN+, HER2−	United States	US payer (managed care) perspective	30 years	Dominant
EndoPredict Compared to No Genomic Testing						
Lymph-Node–Negative						
NICE et al, 2018[36]	Current practice	LN−, NPI ≤ 3.4; LN−, NPI > 3.4	United Kingdom	National Health Service and Personal Social Services	Lifetime	LN−, NPI ≤ 3.4: £147,419/QALY LN−, NPI > 3.4: £46,788
Lymph-Node–Positive (or Mixed)						
Blank et al, 2015[132]	Best practice (German S3, St. Gallen, NCCN guidelines)	LN ≥ 0, ER+, HER2−	Germany	German health care system	Lifetime (50 years)	Dominant
Hinde et al, 2019[122]	Current practice	Age of 56.5, ER+, LN ≥ 0, HER2−, intermediate risk by PREDICT or NPI	United Kingdom	National Health Service	Lifetime	£26,836/QALY
NICE, 2018[36]	Current practice	LN+ (1–3 nodes)	United Kingdom	National Health Service and Personal Social Services	Lifetime	£21,458/QALY

Study ID	Comparator	Patient Group	Country	Perspective	Time Horizon	ICER/Economic Conclusion
Prosigna Compared With No Genomic Testing						
Lymph-Node–Negative						
NICE, 2018[36]	Current practice	LN−, NPI ≤ 3.4; LN−, NPI > 3.4	United Kingdom	National Health Service and Personal Social Services	Lifetime	LN−, NPI ≤ 3.4: £91,028/QALY LN−, NPI > 3.4: £26,058/QALY
Lymph-Node–Positive (or Mixed)						
NICE, 2018[36]	Current practice	LN+ (1-3 nodes)	United Kingdom	National Health Service and Personal Social Services	Lifetime	£28,731/QALY
Stein et al, 2016[a166]	Chemotherapy for all	ER+, HER2-	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	Prosigna subtype compared with chemotherapy for all: dominant Prosigna ROR_PT compared with chemotherapy for all: dominant
Head-to-Head Comparisons						
Retèl et al, 2012[163]	MammaPrint vs. Oncotype DX	ER+, LN−	Netherlands	Dutch health care perspective	20 years	MammaPrint dominant
Seguí et al, 2014[164]	MammaPrint vs. Oncotype DX	ER+, LN−, HER2−	Spain	the Spain national health care system perspective	Lifetime horizon	€1457/QALY (in favour of MammaPrint)
Stein et al, 2016[166]	Oncotype vs. MammaPrint	ER+, HER2−	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	Dominant
Stein et al, 2016a[166]	Oncotype vs. Prosigna subtype; Oncotype vs. Prosigna ROR_PT	ER+, HER2−	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	Oncotype vs. Prosigna subtype: £6,850/QALY Oncotype vs. Prosigna ROR_PT: £36,600/QALY
Stein et al, 2016a[166]	MammaPrint vs. Prosigna subtype; MammaPrint vs. Prosigna ROR_PT	ER+, HER2−	United Kingdom	National Health Service	Lifetime (up to maximum age 100 years)	Prosigna subtype: cost-saving Prosigna ROR_PT: dominant
Yang et al, 2012[174]	MammaPrint vs. Oncotype DX	ER+, LN−	United States	Third-party payer	10 years	MammaPrint: dominant

Abbreviations: AOL, Adjuvant! Online; ER, estrogen receptor; HR, hormone receptor; HER2, human epidermal growth factor 2; ICER, incremental cost-effectiveness ratio; IHC4, immunohistochemistry 4; LN, lymph node; mAOL, modified Adjuvant! Online; NCCN, National Comprehensive Cancer Network; NPI, Nottingham Prognostics Index; NR, not reported; QALY, quality-adjusted life-year; 70G-FFT, 70-gene-fresh frozen tissue; 70G-PAR,70-gene signature based on paraffin blocks.

[a]Two types of Prosigna were used in this analysis: Prosigna subtype, using Prosigna with intrinsic tumor subtypes including Luminal A, Luminal B, HER-2 enriched (HER-2E), and basal-like; Prosigna ROR_PT, Risk of Recurrence–weighted for proliferation score and tumour size.

Appendix 9: Parameters Used in Economic Model

Table A17: Parameters Used to Model Natural History and Impact of Treatment

Model Parameter	Mean (SE)	Value in DSA	Distribution	Reference
		EndoPredict[a]		
Lymph-Node–Negative				
High risk (10-year)	22.1% (3.47%)	20%–40%	Beta	Sestak et al, 2018[52]
Low risk (10-year)	6.6% (1.33%)	4%–10%	Beta	
Lymph-Node–Positive				
High risk (10-year)	30.3% (4.16%)	20%–40%	Beta	Sestak et al, 2018[52]
Low risk (10-year)	5.6% (4.97%)	4%–10%	Beta	
		MammaPrint		
Lymph-Node–Negative				
Clinical low, genetic low (5-year)[d]	2.4% (0.33%)	0–5%	Beta	Cardoso et al, 2016[51]
Clinical high, genetic low (5-year)[d]	6.1% (1.40%)	5%–20%	Beta	
Clinical low, genetic high (5-year) [d]	4.5% (1.53%)	2%–10%	Beta	
Clinical high, genetic high (5-year)[e]	9.1% (1.33%)	5%–20%	Beta	
Lymph-Node–Positive				
Clinical low, genetic low (5-year)[d]	2.4% (0.31%)	0–5%	Beta	Cardoso et al, 2016[51]
Clinical high, genetic low (5-year)[d]	4.4% (1.20%)	2%–10%	Beta	
Clinical low, genetic high (5-year) [d]	6.1% (1.76%)	5%–20%	Beta	
Clinical high, genetic high (5-year)[e]	9.4% (0.77%)	5%–20%	Beta	
		Oncotype DX[a]		
Lymph-Node–Negative				
Recurrence score of ≥ 26 (10-year)[b]	27.2% (6.1%)	20%–40%	Beta	Sestak et al, 2018[52]
Recurrence score of 10–25 (9-year)[c]	5.5% (0.50%)	2%–10%	Beta	Sparano et al, 2018[26]
Recurrence score of ≤ 10 (9-year)[c]	3.2% (0.70%)	0–5%	Beta	Sparano et al, 2018[26]
Lymph-Node–Positive				
High risk (10-year)	38.0% (11.25%)	30%–50%	Beta	Sestak et al, 2018[52]
Intermediate risk (10-year)	29.1% (6.17%)	20%–40%	Beta	
Low risk (10-year)	19.4% (4.34%)	10%–30%	Beta	

Model Parameter	Mean (SE)	Value in DSA	Distribution	Reference
		Prosigna[a]		
Lymph-Node–Negative				
High risk (10-year)	32.4% (5.20%)	20%–40%	Beta	Sestak et al, 2018[52]
Intermediate risk (10-year)	14.1% (2.91%)	10%–20%	Beta	
Low risk (10-year)	3.0% (1.07%)	0–5%	Beta	
Lymph-Node–Positive				
High risk (10-year)	30.7% (4.87%)	20%–40%	Beta	Sestak et al, 2018[52]
Intermediate risk (10-year)	20.7% (5.71%)	10%–20%	Beta	
Low risk (10-year)	3.0% (1.07%)	0–5%	Beta	

Abbreviations: DSA, deterministic sensitivity analysis; SE: standard error.
[a]Risk of developing a distant recurrence if receiving only hormone therapy.
[b]Reported from TransATAC study,[52] from the Oncotype DX high-risk group (Recurrence Score of ≥ 31).
[c]As reported in the TAILORx trial,[26] risk of developing a distant recurrence in 9 years; not directly used in the model calculation.
[d]As reported in the MINDACT trial,[51] not directly used in the model calculation; risk of developing a distant recurrence in 5 years; risk for people receiving only hormone therapy.
[e]As reported in the MINDACT trial,[51] this 5-year risk was for people with clinical and genetic high risk. All these people had received adjuvant chemotherapy. We used this risk, converted it to 10-year risk, and divided the 10-year risk with a relative risk of 0.76 to estimate the 10-year risk of developing a distant recurrence if people of this risk group had not received adjuvant chemotherapy.

Table A18: Cost Parameters Used to Model Natural History and Impact of Treatment

		Costs for Those Accepting Adjuvant Chemotherapy and Hormone Therapy, $				Costs for Those Accepting Hormone Therapy, $			
Cost Items	Proportion Incurred in the First Year[a]	Mean	SD	Estimated Monthly Cost in the First Year	Estimated Monthly Cost in the Second Year	Mean	SD	Estimated Monthly Cost in the First Year	Estimated Monthly Cost in the Second Year
Inpatient	90%	1,974	4,916	173.37[b]	74.13	844	3,565	26.49	11.33
Emergency department	90%	437	922	38.38[b]	13.26	151	334	5.86	2.03
Cancer clinic	90%	20,841	9,541	1830.43[b]	1212.20	13,802	8,590	279.66	185.21
Rehabilitation	90%	0	0	0.00	2.81	32	526	0.00	0.43
Complex continuing care	90%	27	394	2.37[b]	0.00	0	0	0.36	0.00
Home care	90%	1,596	1,649	140.17[b]	25.29	288	1,129	21.42	3.86
Physician billing	90%	6,642	3,904	583.35[b]	470.41	5,356	3,281	89.13	71.87
Chemotherapy	100%	3,129	7,695	559.81[c]	0.00	—	—	0.00	0.00
Supportive drugs	100%	1,603	3,501	286.79[c]	0.00	—	—	0.00	0.00
Endocrine therapy	60%	—	—	0.00	11.97[d]	223	375[e]	11.97	11.97

Abbreviation: SD, standard deviation.
[a]The assumed proportion of the costs incurred in the first year in the total cost of the 20 months of treatment after diagnosis.
[b]Divided by 11, because in our model, the first cycle was for risk classification and the treatment started in the second cycle. We estimated the monthly treatment cost by assuming the treatment cost in the first year was incurred within 11 months.
[c]Divided by 6, the adjuvant chemotherapy lasted for 6 cycles (6 months).
[d]Equal to the cost for those who received hormone therapy only.
[e]We assumed that patients used hormone therapy for 20 months.

Table A19: Hormone Therapy and Adjuvant Chemotherapy Regimens by Risk

Risk Classification	Hormone Therapy[a]	Chemotherapy[b]
High risk	Tamoxifen for 10 years or aromatase inhibitor for 7 years	6 cycles, FEC-D regimen First 3 cycles (fluorouracil, epirubicin, cyclophosphamide: 21 days) Day 1: IV at the hospital Days 2 to 21: rest days Next 3 cycles (docetaxel: 21 days) Day 1: IV at the hospital Days 2 to 21: rest days
Intermediate risk	Tamoxifen for 10 years or aromatase inhibitor for 7 years	4 cycles (docetaxel, cyclophosphamide: 21 days) Day 1: IV at the hospital Days 2 to 21: rest days
Low risk	Tamoxifen for 5 years or aromatase inhibitor for 5 years[c]	Same as those at intermediate risk, if receiving adjuvant chemotherapy
No risk classification	Tamoxifen for 7 years or aromatase inhibitor for 7 years[c]	Same as those at intermediate risk, if receiving adjuvant chemotherapy

Abbreviations: FEC-D, fluorouracil, epirubicin, cyclophosphamide, and docetaxel; IV, intravenous.
[a]In the reference case analysis, we assumed that all patients were postmenopausal.
[b]All people receiving chemotherapy would also receive granulocyte colony-stimulating factor for 8 days every cycle.
[c]For sensitivity analysis.

Table A20: Variables Used in the Scenario Analyses

Variable	Difference Between Scenario Analysis and Reference Case Analysis: Scenario Analysis	Difference Between Scenario Analysis and Reference Case Analysis: Reference Case Analysis	Reference
Comparison between tests	Cost-effectiveness of GEP tests compared with one another, using costs and QALYs estimated from reference case analysis	Cost-effectiveness of GEP tests compared with usual care	NA
Triage test before GEP tests	Usual care: using a clinical tool, modified AOL to classify people as low- and high risk	Usual care: no test, no risk classification	Sparano et al, 2018[26] Cardoso et al, 2016[261]
Premenopausal population	Age: 50 Hormone therapy: tamoxifen for 10 years Classification by Recurrence Score (0–15, 16–20, 21–25, 26–100)	Age: 58 Hormone therapy: tamoxifen for 7 years Classification by Recurrence Score (0–25, 26–100)	Assumption
LN+ population	*Probability of classification* High risk: 10.9% Intermediate risk: 31.7% Low risk: 57.4%	*Probability of classification* Recurrence score 26–100: 15.3% Recurrence score 11–25: 66.9% Recurrence score 0–10: 17.8%	Sestak et al, 2018[52] Torres 2018[103]
	10-year risk of distant recurrence High risk: 38.0% Intermediate risk: 29.1% Low risk: 19.4%	*10-year risk of distant recurrence* High risk: 27.2% Intermediate risk: 6.1% Low risk: 3.5%	
	Proportion of adjuvant chemotherapy High risk: 100% Intermediate risk: 78.3% Low risk: 28.9% Usual care: 79%	*Proportion of adjuvant chemotherapy*[84] High risk: 79.3% Intermediate risk: 32.9% Low risk: 4.1% Usual care: 38%	
Three-category Oncotype DX results	*Probability of classification* *High risk: 10.3%* *Intermediate risk: 26.4%* *Low risk: 63.3%*	*Probability of classification* Recurrence Score of 26–100: 15.3% Recurrence Score of 11–25: 66.9% Recurrence Score of 0–10: 17.8%	Sestak et al, 2018[52]
	10-year probability of distant recurrence High risk: 27.2% Intermediate risk: 16.7% Low risk: 5.9%	*10-year probability of distant recurrence* High risk: 27.2% Intermediate risk: 6.1% Low risk: 3.5%	
Status quo comparison	All people receive Oncotype DX through out-of-country program in the reference group	Usual care (no GEP tests) as the reference group	Assumption
Oncotype DX varied risk classification	*Probability of classification* High risk: 9.4% Intermediate risk: 32.9% Low risk: 57.7%	*Probability of classification* High risk: 10.3% Intermediate risk: 26.4% Low risk: 63.3%	Levine et al, 2016[84]

Variable	Difference Between Scenario Analysis and Reference Case Analysis		Reference
	Scenario Analysis	Reference Case Analysis	
Various chemotherapy acceptance	*Proportion of adjuvant chemotherapy* High risk: 100% Low risk: 0	*Proportion of adjuvant chemotherapy*[84] High risk: 79.3% Intermediate risk: 32.9% Low risk: 4.1%	Assumption
	3-category test High risk: 74% Intermediate: 17% Low risk: 0 *2-category test* High risk: 77% Low risk: 7%	*Proportion of adjuvant chemotherapy*[84] High risk: 79.3% Intermediate risk: 32.9% Low risk: 4.1%	NICE 2018[118]
Local recurrence	Probability: 10.5% of distant recurrence people developing local recurrence in the previous cycle Utility: −0.108 Cost: $8,397	No local recurrence	NICE 2018[118] Will et al, 2000[189]
Predictive benefit of tests	Different distant recurrence risk reduction recurrence across risk levels *Oncotype DX* No absolute risk reduction for those with a Recurrence Score of ≤ 10; monthly probabilities of distant recurrence for those with a Recurrence Score of 11–25, and ≥ 26 were converted from the reported 9-year risks *EndoPredict* Absolute risk reduction of 7.4% for high-risk people; 1.9% for low risk people[a]	Consistent relative risk reduction for different risk classifications	Sparano et al, 2018[26] Sestak et al, 2019[102]
Risk-dependent chemotherapy regimens	Cost for chemotherapy varies according to the regimens and risk levels	Same chemotherapy cost applied	Assumption, medication cost[b]

Abbreviations: AOL, Adjuvant! Online; EPclin, EndoPredict clinical score (a number between 1.1 and 6.2 that maps to a percentage Risk of Recurrence); GEP, gene expression profiling; LN, lymph node; NA, not applicable; NICE, National Institute for Clinical Excellence; QALY, quality-adjusted life-years.

[a]Selected from Sestak et al 2019,[102] based on EPclin score of 3 and 4 for low and high risk respectively. Because the 10-year risks of EPClin score of 3 and 4 were similar to the average 10-year risks of low and high risk groups in TransATAC study.[52]

[b]Ivan Tyono, email communication, May 23, 2019.

Table A21: Distributions Used in the Probabilistic Sensitivity Analysis

Variable	Distribution	Reference
Risk Classification Probabilities		
EndoPredict (low, high)	Beta	Sestak et al, 2018[52]
MammaPrint (low, high)	Dirichlet	Cardoso et al, 2016[261]
Oncotype DX (Recurrence Score of 0–10, 11–25, 26–100)	Dirichlet	Sparano et al, 2018[26]
Prosigna (low, intermediate, high)	Dirichlet	Sestak et al, 2018[52]
Proportion of Adjuvant Chemotherapy		
Different risk groups across GEP tests (chemotherapy or not)	Beta	Levine 2016,[84] Mittmann 2018[126]
Transition Probabilities		
From no distant recurrence to distant recurrence	Beta	Sparano et al, 2018[26] Sestak et al, 2018[52]
Chemotherapy toxicity–related death	Beta	Ludwig et al, 1989[182] Paulden et al, 2013[158]
From distant recurrence to death	Gamma	Leone et al, 2017[184]
Utility		
Recurrence-free in the first year with hormone therapy	Beta	Lidgren et al, 2007[185]
Recurrence-free in the first year with chemotherapy	Beta	Lidgren et al, 2007[185]
Distant recurrence	Beta	Lidgren et al, 2007[185]
Recurrence-free in subsequent years	Beta	Lidgren et al, 2007[185]
Cost		
Chemotherapy cost	Gamma	Mittmann et al, 2015[262]
Hormone therapy cost	Gamma	Mittmann et al, 2015[262]
Cost of chemotherapy toxicity–related death	Gamma	Paulden et al, 2013[158]
Distant recurrence treatment cost	Gamma	Will et al, 2000[189] Paulden et al, 2013[158]

Abbreviation: GEP, gene expression profiling.

Appendix 10: Results of the Economic Model

Table A22: Reference Case Analysis Results for Outcome and Costs, GEP Tests Versus Usual Care

Strategy	10-Year Distant Recurrence (Per 1,000 Persons)	10-Year Death From Breast Cancer (Per 1,000 Persons)	Providing Test (Per Person)	Providing Adjuvant Chemo (Per Person)	Incurred Before Distant Recurrence (Per Person)	Incurred After Distant Recurrence (Per Person)	Incurred Over Last 3 Months of Life (Per Person)
Usual care[a,b]	95	73	0	8,245	29,607	5,060	4,048
EndoPredict	91	70	2,964	5,360	30,246	4,762	3,811
Usual care[a,c]	77	59	0	8,245	29,880	4,148	3,316
MammaPrint	76	59	3,758	4,923	30,543	4,039	3,232
Usual care[a,d]	76	58	0	8,245	29,887	4,126	3,299
Oncotype DX	76	58	4,869	3,382	30,847	3,969	3,176
Usual care[a,b]	95	73	0	8,245	29,607	5,060	4,048
Prosigna	92	70	2,576	5,394	30,362	4,607	3,691

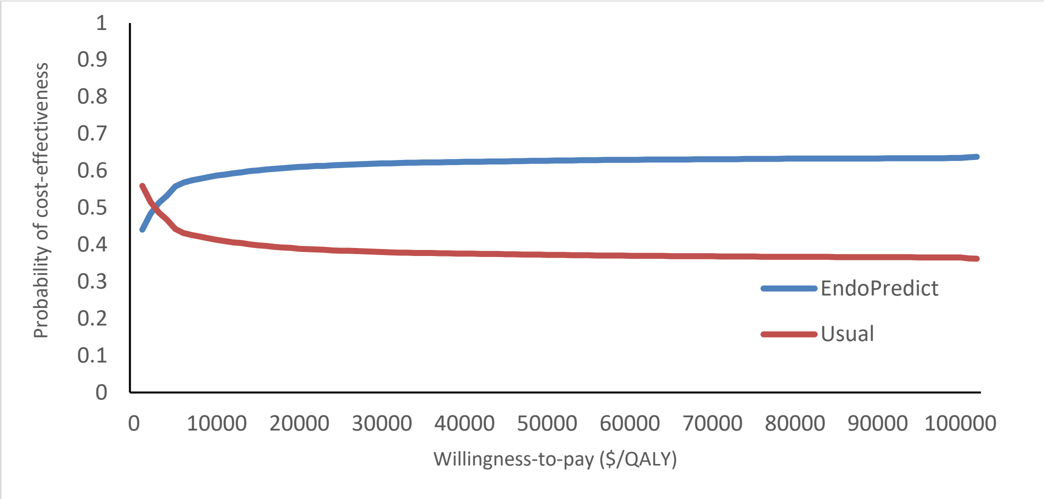
Abbreviations: CrI, credible interval; GEP, gene expression profiling; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.
[a]Reference group; usual care varies for each comparison.
[b]Usual care for EndoPredict and Prosigna represents the cost and outcomes for the study population in TransATAC study,[52] if they had not received EndoPredict or Prosigna.
[c]Usual care MammaPrint represents the cost and outcomes for the study population in MINDACT trial,[51] if they had not received MammaPrint.
[d]Usual care for Oncotype DX represents the cost and outcomes for the study population in TAILORx trial,[26] if they had not received Oncotype DX.

Table A23: Other Results of Scenario Analyses

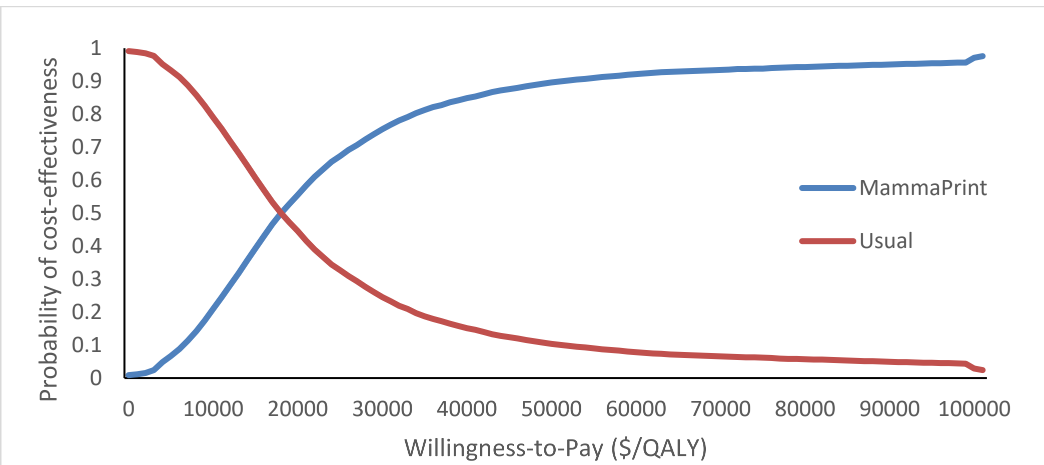
Scenario	Incremental Cost vs. Usual Care, $ (95% CrI)[a]	Incremental QALY vs. Usual Care, QALY (95% CrI)[a]	ICER, $/QALY[a]
Three-category Oncotype DX results	EndoPredict: NA MammaPrint: NA Oncotype DX: 1,478 (−1,164 to 4,078) Prosigna: NA	EndoPredict: NA MammaPrint: NA Oncotype DX: 0.06 (−0.63 to 0.74) Prosigna: NA	EndoPredict: NA MammaPrint: NA Oncotype DX: 26,460 Prosigna: NA
Status quo comparison	EndoPredict: NA MammaPrint: NA Oncotype DX: −541[b] Prosigna: NA	EndoPredict: NA MammaPrint: NA Oncotype DX: 0 Prosigna: NA	EndoPredict: NA MammaPrint: NA Oncotype DX: cost-saving Prosigna: NA
Oncotype DX varied risk classification	EndoPredict: NA MammaPrint: NA Oncotype DX: 2,033 (−486 to 4,569) Prosigna: NA	EndoPredict: NA MammaPrint: NA Oncotype DX: −0.04 (−0.71 to 0.64) Prosigna: NA	EndoPredict: NA MammaPrint: NA Oncotype DX: dominated[c,d] Prosigna: NA
Various chemotherapy acceptance (no chemotherapy for low-risk, 100% for high risk)	EndoPredict: 496 (−2,078 to 3,093) MammaPrint: 1,155 (497–1,764) Oncotype DX: 564 (−246 to 1,308) Prosigna: −252 (−1,142 to 590)	EndoPredict: 0.17 (−0.49 to 0.81) MammaPrint: 0.08 (0.03–0.13) Oncotype DX: 0.10 (0.01–0.21) Prosigna: 0.20 (0.11–0.31)	EndoPredict: 2,928 MammaPrint: 15,405 Oncotype DX: 5,894 Prosigna: dominant[e]
Various chemotherapy acceptance (for 2-category test: 7% for low risk, 77% for high risk; for 3-category test: 0 for low risk, 17% for intermediate risk, 77% for high risk)	EndoPredict: 426 (−2,188 to 2,972) MammaPrint: 1,166 (519–1,759) Oncotype DX: 1,029 (235–1,718) Prosigna: −1,490 (−2,348 to 740)	EndoPredict: 0.12 (−0.53 to 0.80) MammaPrint: 0.05 (0.01–0.09) Oncotype DX: 0.07 (0.00–0.16) Prosigna: 0.11 (0.04–0.21)	EndoPredict: 3,491 MammaPrint: 25,053 Oncotype DX: 15,439 Prosigna: dominant[e]
Local recurrence	EndoPredict: 173 (−2,477 to 2,781) MammaPrint: 901 (177–1,564) Oncotype DX: 681 (−158 to 1,434) Prosigna: −347 (−1,206 to 437)	EndoPredict: 0.12 (−0.54 to 0.79) MammaPrint: 0.05 (0.00–0.10) Oncotype DX: 0.07 (0.00–0.16) Prosigna: 0.15 (0.08–0.26)	EndoPredict: 1,402 MammaPrint: 19,683 Oncotype DX: 10,284 Prosigna: dominant[e]
Triage test for GEP tests	EndoPredict: NA MammaPrint: 2,165 (1,528–2,753) Oncotype DX: 1,624 (792–2,355) Prosigna: NA	EndoPredict: NA MammaPrint: −0.02 (−0.06 to 0.03) Oncotype DX: 0.05 (−0.02 to 0.16) Prosigna: NA	EndoPredict: NA MammaPrint: dominated Oncotype DX: 29,831 Prosigna: NA

Scenario	Incremental Cost vs. Usual Care, $ (95% CrI)[a]	Incremental QALY vs. Usual Care, QALY (95% CrI)[a]	ICER, $/QALY[a]
Predictive benefit of tests	EndoPredict: −103 (−2,808 to 2,613) MammaPrint: NA Oncotype DX: 272 (−1,034 to 1,469) Prosigna: NA	EndoPredict: 0.20 (−0.49 to 0.89) MammaPrint: NA Oncotype DX: 0.19 (−0.07 to 0.48) Prosigna: NA	EndoPredict: dominant[e] MammaPrint: NA Oncotype DX: 1,457 Prosigna: NA
Risk-dependent chemotherapy regimens	EndoPredict: 667 (−1,931 to 3,204) MammaPrint: 1,332 (755–1,831) Oncotype DX: 917 (205 –1,522) Prosigna: −61 (−791 to 589)	EndoPredict: 0.12 (−0.54 to 0.79) MammaPrint: 0.05 (0.00–0.10) Oncotype DX: 0.07 (0.00–0.17) Prosigna: 0.15 (0.08–0.26)	EndoPredict: 5,422 MammaPrint: 29,147 Oncotype DX: 13,871 Prosigna: dominant[e]

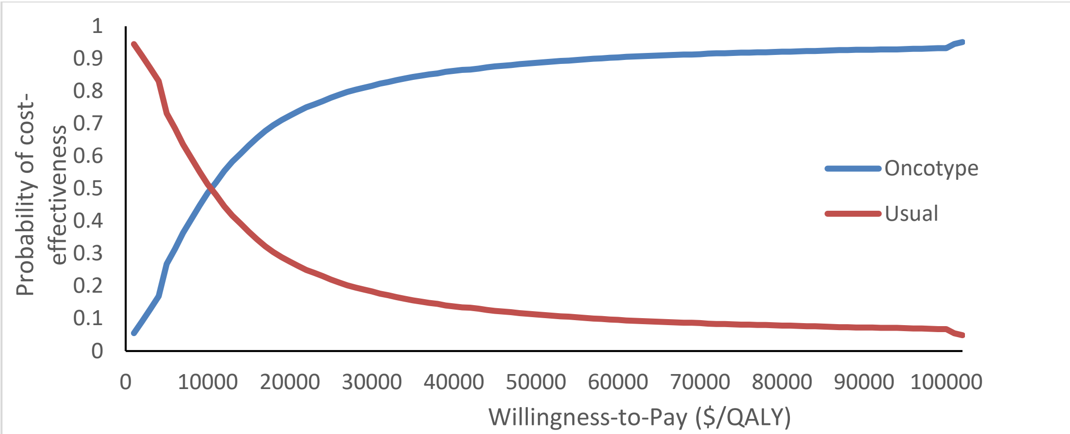
Abbreviations: CrI, credible interval; GEP, gene expression profiling; ICER, incremental cost-effectiveness ratio; NA, not applicable; QALY, quality-adjusted life-year.
[a]Reference group; usual care varies for each comparison.
[b]No credible interval because no probabilistic sensitivity analysis was conducted. The only difference between this scenario and the reference case was the Oncotype DX price difference.
[c]Dominated: more costly and less effective than usual care.
[d]Oncotype DX was interpreted as a 3-category test (low, intermediate, and high) in this scenario analysis.
[e]Dominant: less costly and more effective than usual care.



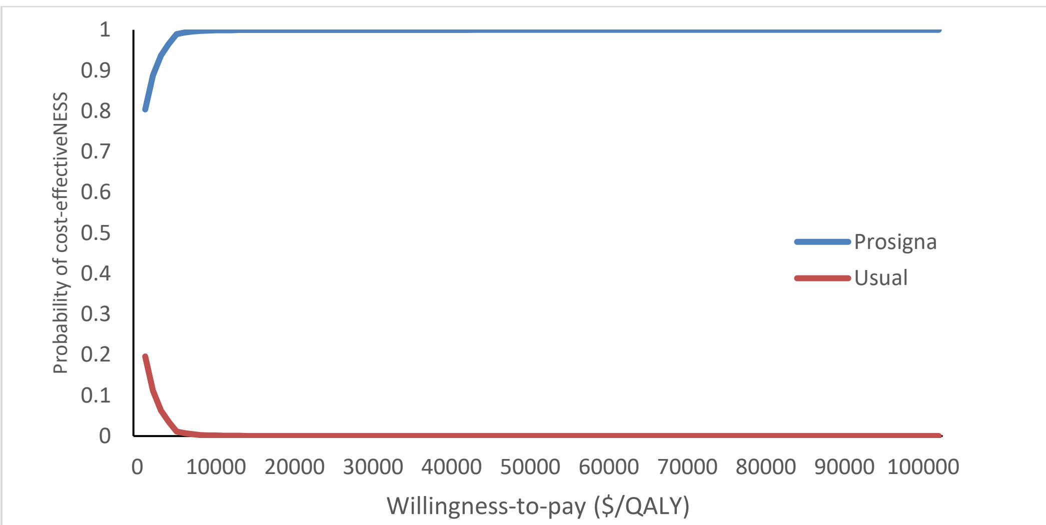
A. EndoPredict Versus Usual Care



B. MammaPrint Versus Usual Care

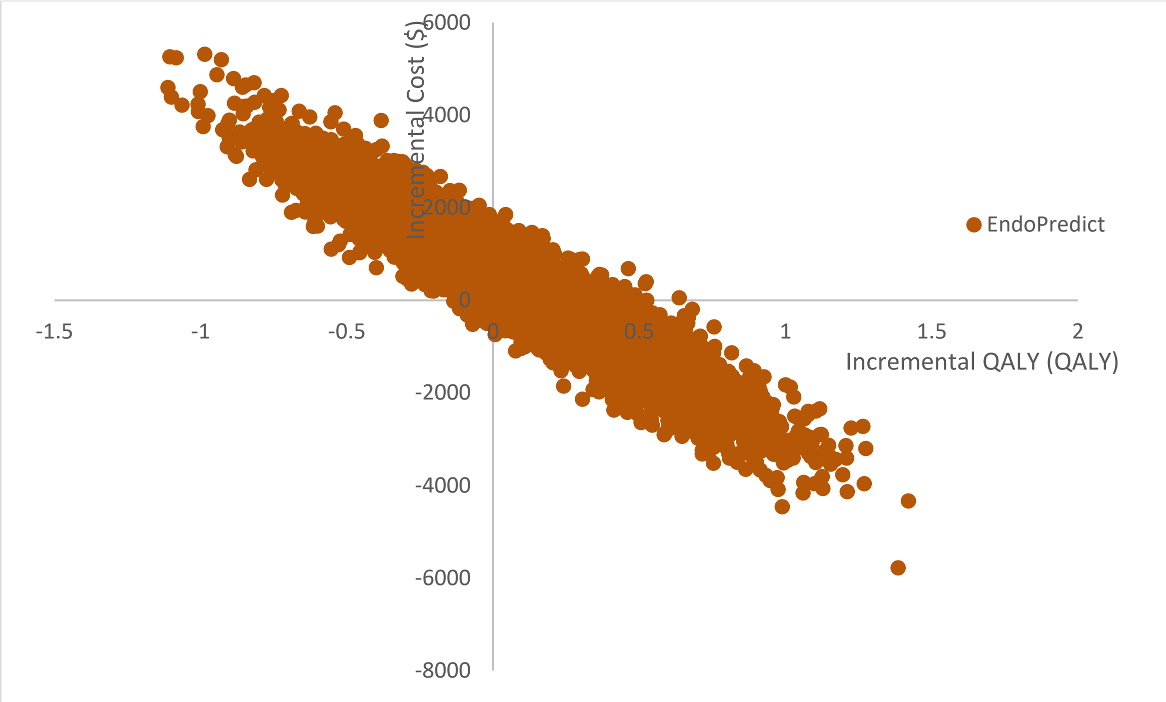


C. Oncotype DX Versus Usual Care

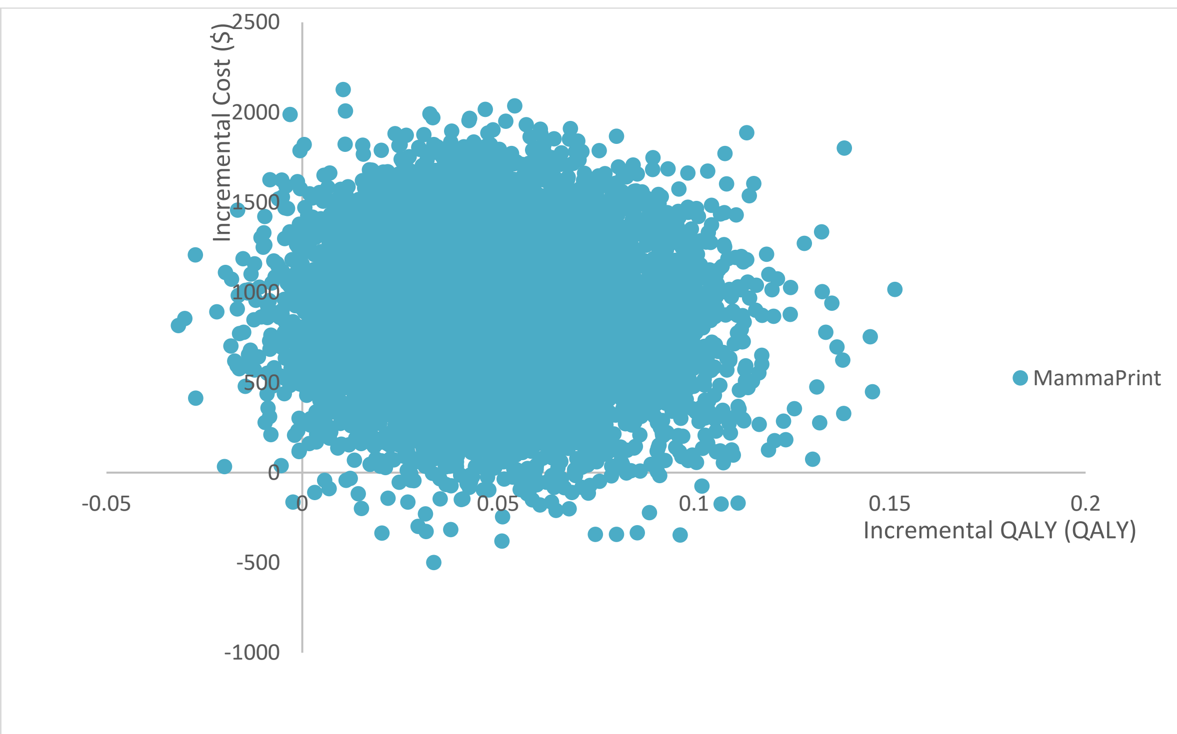


D. Prosigna Versus Usual Care

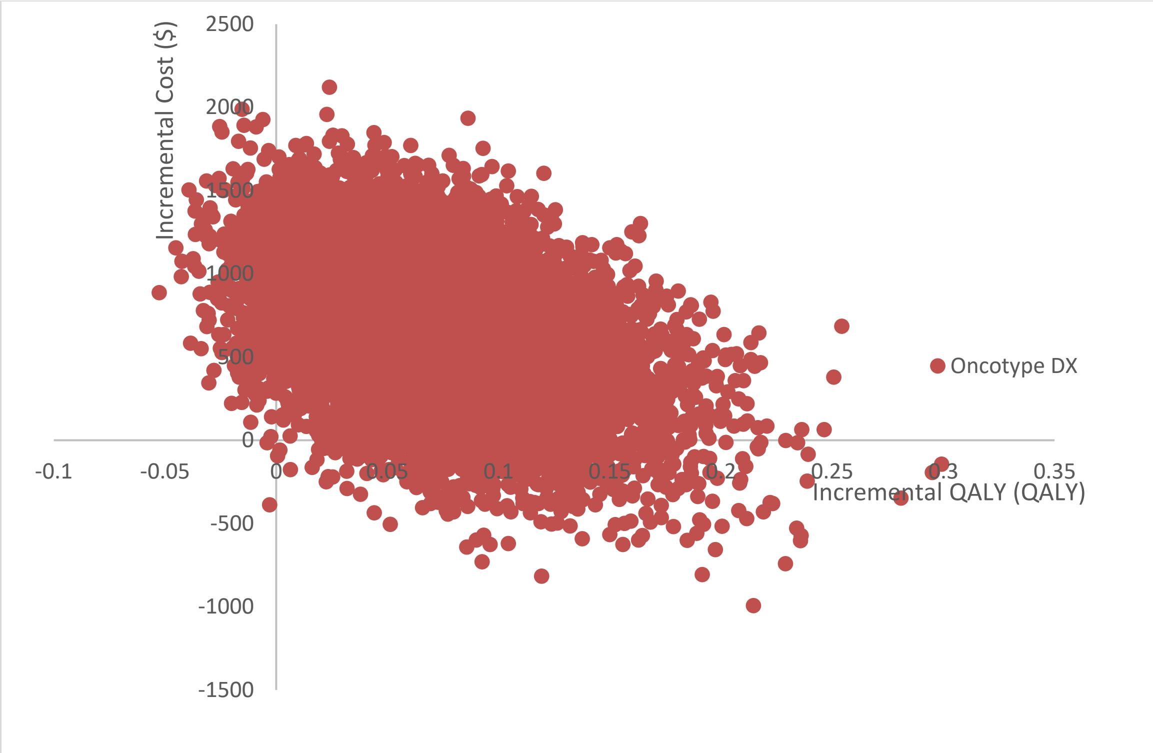
Figure A1: Cost-Effectiveness Acceptability Curves



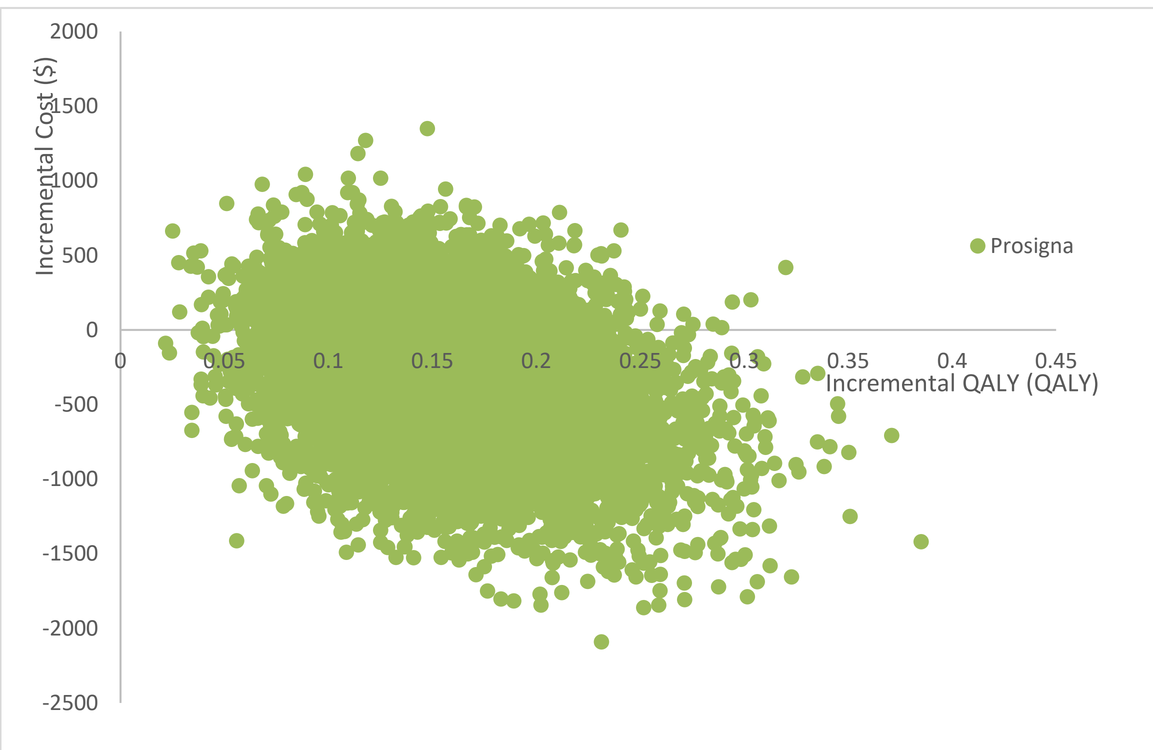
A. EndoPredict Versus Usual Care



B. MammaPrint Versus Usual Care

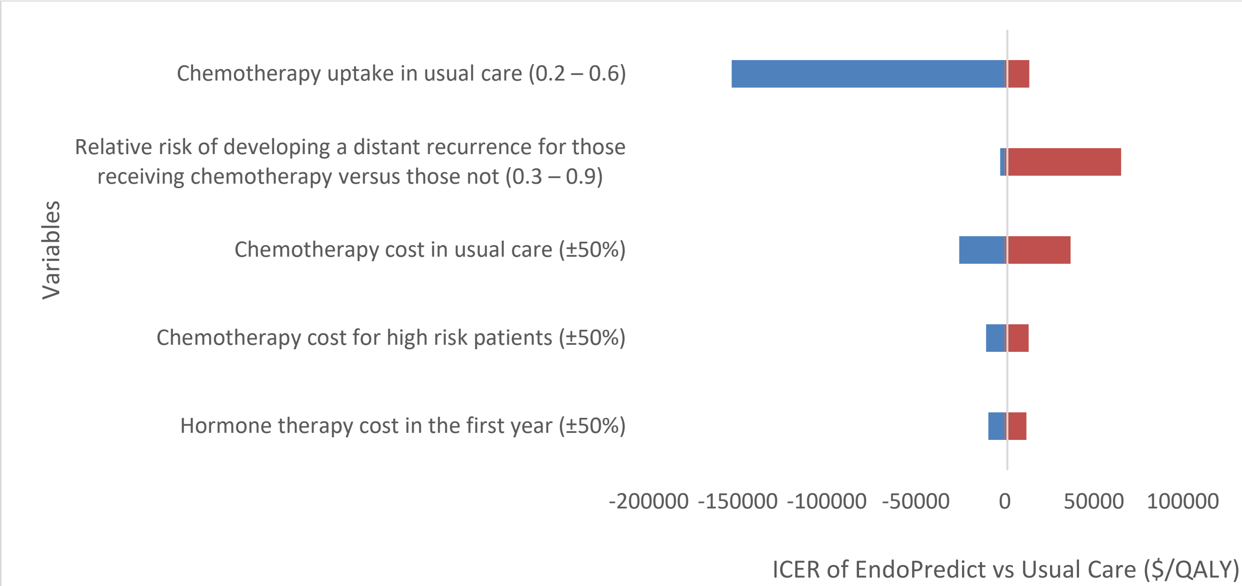


C. Oncotype DX Versus Usual Care

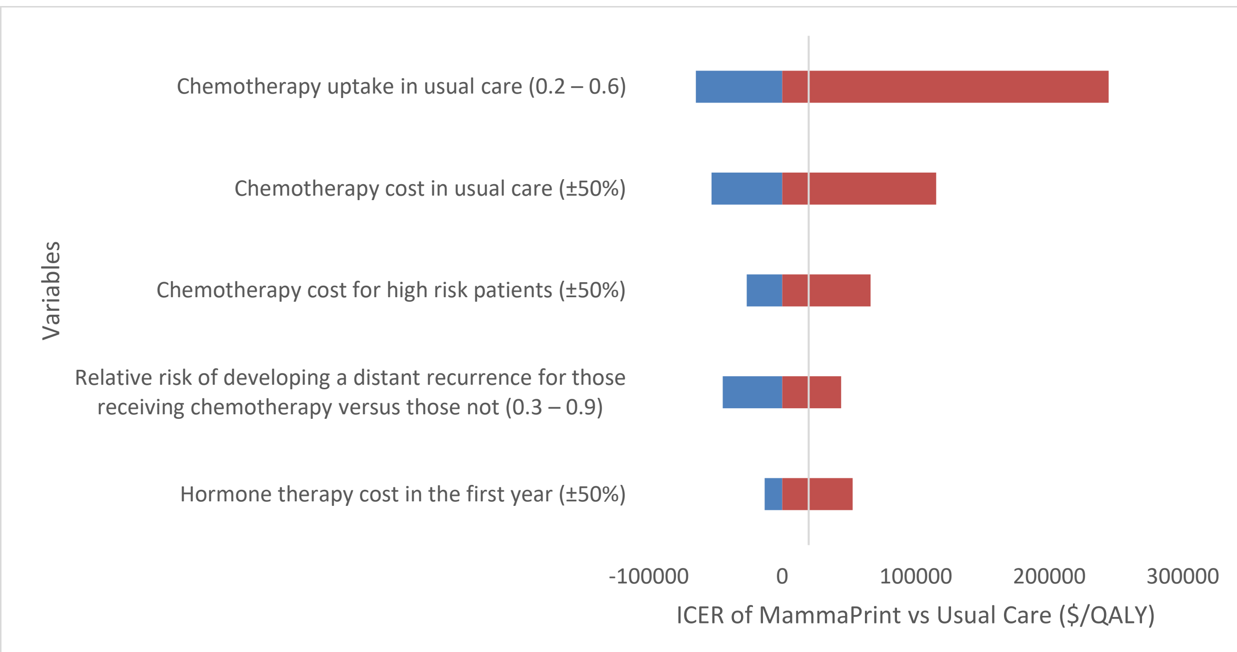


D. Prosigna Versus Usual Care

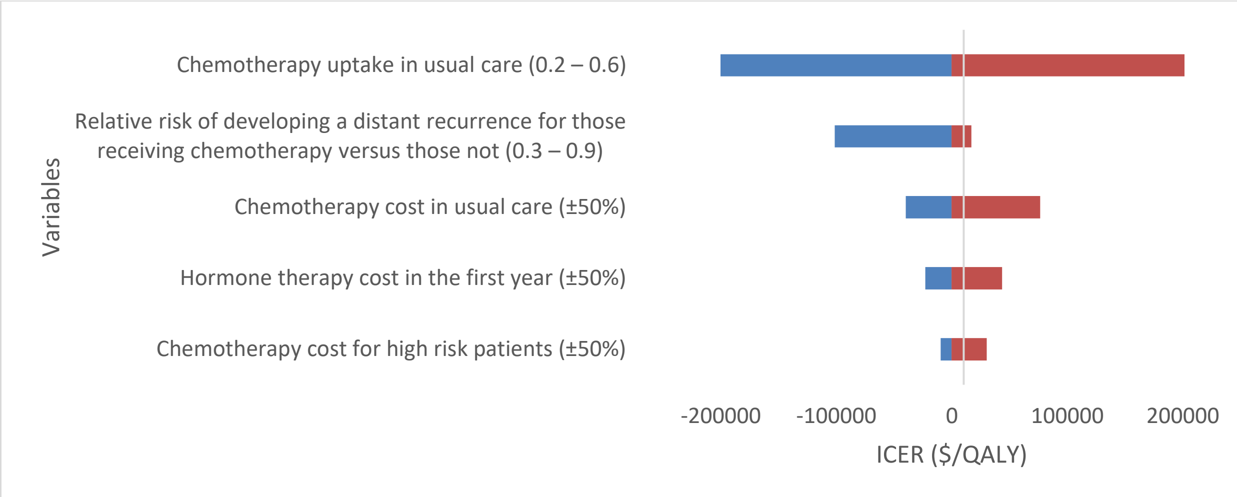
Figure A2: Incremental Cost-Effectiveness Scatter Plots



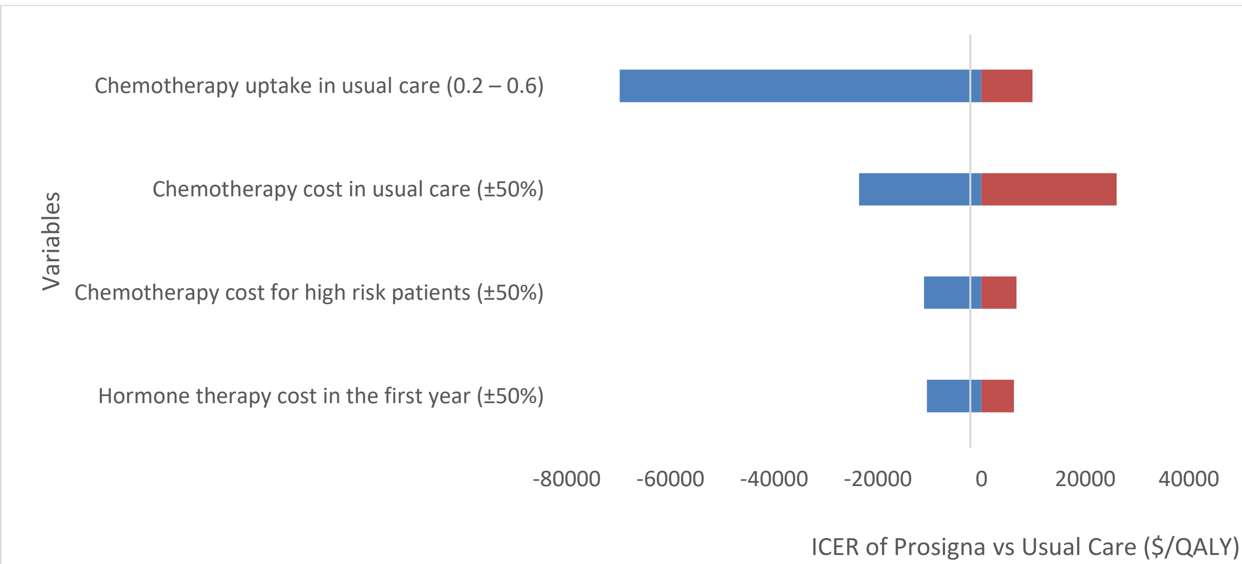
A. EndoPredict Versus Usual Care



B. MammaPrint Versus Usual Care



C. Oncotype DX Versus Usual Care



D. Prosigna Versus Usual Care

Figure A3: Tornado Plots for Incremental Cost-Effectiveness Ratios

Appendix 11: Budget Impact Analysis

Table A24: Per-Person Costs in Budget Impact Analysis

Test, Year	Test	Adjuvant Chemotherapy	Incurred Prior to Distant Recurrence	Incurred Following Distant Recurrence	Incurred Over Last 3 Months of Life	Total Cost
EndoPredict						
Year 1	$2,963.50	$5,383.82	$18,305.29	$82.78	$25.86	$26,761.25
Year 2	$0.00	$0.00	$3,649.07	$158.69	$89.47	$3,897.23
Year 3	$0.00	$0.00	$701.25	$194.79	$134.74	$1,030.79
Year 4	$0.00	$0.00	$624.23	$216.48	$162.67	$1,003.37
Year 5	$0.00	$0.00	$549.10	$228.79	$179.31	$957.21
MammaPrint						
Year 1	$3,757.53	$4,944.60	$18,434.84	$72.88	$23.38	$27,233.22
Year 2	$0.00	$0.00	$3,632.65	$133.78	$76.45	$3,842.88
Year 3	$0.00	$0.00	$703.77	$163.42	$113.53	$980.72
Year 4	$0.00	$0.00	$627.55	$181.32	$136.47	$945.34
Year 5	$0.00	$0.00	$552.98	$191.57	$150.21	$894.76
Oncotype DX						
Year 1	$4,868.59	$3,396.76	$18,865.81	$72.54	$23.19	$27,226.88
Year 2	$0.00	$0.00	$3,550.02	$133.83	$76.36	$3,760.21
Year 3	$0.00	$0.00	$703.96	$163.37	$113.51	$980.84
Year 4	$0.00	$0.00	$627.73	$181.05	$136.36	$945.14
Year 5	$0.00	$0.00	$553.16	$191.02	$149.92	$894.09
Prosigna						
Year 1	$2,576.00	$5,417.55	$18,296.82	$82.64	$25.35	$26,398.35
Year 2	$0.00	$0.00	$3,649.80	$162.58	$90.99	$3,903.37
Year 3	$0.00	$0.00	$700.93	$199.41	$137.87	$1,038.21
Year 4	$0.00	$0.00	$623.80	$220.96	$166.30	$1,011.06
Year 5	$0.00	$0.00	$548.64	$232.61	$182.78	$964.02
No Test[a]						
Year 1	$0.00	$8,281.09	$17,506.87	$72.50	$23.40	$25,883.86
Year 2	$0.00	$0.00	$3,811.32	$131.87	$75.55	$4,018.74
Year 3	$0.00	$0.00	$703.55	$161.26	$112.00	$976.82
Year 4	$0.00	$0.00	$627.41	$179.28	$134.78	$941.47
Year 5	$0.00	$0.00	$552.89	$189.88	$148.65	$891.43

Numbers may appear inexact due to rounding.
[a]Estimated based on the usual care group in the cost-effectiveness analysis comparing Oncotype DX and usual care.

Table A25: Budget Impact Scenario Analysis

Scenario	Budget Impact, $ Million[a]					
	Year 1	Year 2	Year 3	Year 4	Year 5	Total
EndoPredict monopoly	−0.27	−0.01	0.41	0.87	1.39	2.39
MammaPrint monopoly	1.68	1.86	2.20	2.55	2.91	11.21
Oncotype monopoly	1.65	1.49	1.81	2.13	2.45	9.53
Prosigna monopoly	−1.76	−1.59	−1.26	−0.88	−0.45	−5.94
Prosigna market share increasing to 45% while Oncotype decreases to 40% in 5 years, no change to EndoPredict and MammaPrint[b]	1.29	0.84	0.81	0.77	0.73	4.44
60% of the original price	−6.14	−6.79	−7.05	−7.29	−7.53	−34.80
70% of the original price	−4.28	−4.79	−4.90	−5.00	−5.09	−24.07
80% of the original price	−2.42	−2.79	−2.76	−2.71	−2.65	−13.34
90% of the original price	−0.57	−0.79	−0.61	−0.42	−0.21	−2.60
110% of the original price	3.15	3.21	3.68	4.16	4.66	18.86
120% of the original price	5.01	5.21	5.82	6.45	7.10	29.59
Clinical low risk only[c,d]	6.22	6.27	6.30	6.32	6.34	31.47
Clinical high risk only[c,e]	−12.31	−13.11	−13.25	−13.38	−13.51	−65.56

[a]In 2018 Canadian dollars.
[b]Prosigna has the lowest price in the reference case analysis.
[c]Assuming only Oncotype DX (the most expensive test) is funded.
[d]The scenario of clinical low risk people receiving tests and clinical high risk people not receiving tests compared with 40% of eligible people receiving tests.
[e]The scenario of clinical low-risk people not receiving tests and clinical high-risk people receiving tests compared with 40% of eligible people receiving tests.

Appendix 12: Letter of Information

LETTER OF INFORMATION

Health Quality Ontario[e] is conducting a review of gene expression profiling testing for people with early-stage breast cancer. The purpose is to understand whether this test should be publicly funded in Ontario.

An important part of this review involves gathering perspectives of patients and caregivers with experience with either gene expression profiling test or other current testing for early-stage breast cancer for treatment recommendation. They could have had the gene expression profiling test, recently or in the past or could be considering it in the future.

What Do You Need From Me

- ✓ Willingness to share your story
- ✓ 30 minutes of your time for a phone or videoconference
- ✓ Permission to audio- (not video-) record the interview

What Your Participation Involves

If you agree to share your experiences, you will be asked to have an interview with Health Quality Ontario[a] staff. The interview will last about 30 minutes. It will be held over the telephone or videoconference. With your permission, the interview will be audiotaped. The interviewer will ask you questions about your or your loved one's condition and your perspectives about treatment options in Ontario.

Participation is voluntary. You may refuse to participate, refuse to answer any questions or withdraw before or at any point during your interview. Withdrawal will in no way affect the care you receive.

Confidentiality

All information you share will be kept confidential and your privacy will be protected except as required by law. The results of this review will be published, however no identifying information will be released or published. Any records containing information from your interview will be stored securely until project completion. After the project completion, the records will be destroyed.

Risks to Participation

There are no known physical risks to participating. Some participants may experience discomfort or anxiety after speaking about their experience.

If you are interested, please contact us before May 31, 2019:

[e] Health Quality Ontario is now the Quality business unit at Ontario Health.

Appendix 13: Interview Guide

Introduction

Health Quality Ontario[f] is a provincial advisor to the Ministry of Health and the Ministry of Long-Term Care. We do a few things for the ministries, but one of the roles that we have is to conduct health technology assessments, which look at technologies and new health services. We review these technologies and health services for the consideration of public funding. If any of the questions seem to cause emotional distress or are uncomfortable, please let me know. You can feel free to either not answer the question or say as little as you like. Having said that, do you have any questions for me?

History of condition (early-stage breast cancer)	
Experience with early-stage breast cancer	

Lived experience with early-stage breast cancer

How is your day-to-day routine?	
What has been the impact and effect on quality of life?	
Did you see any sort of loss of independence?	
Did it have an impact on your loved ones/ caregivers, work, friends?	

Gene expression profiling testing

How did it meet or not meet your needs? How was it adequate or not adequate?	
What were the side effects?	
What were the benefits?	
What were the limitations and barriers?	
Were there issues related to cost, access, knowledge of health care system, etc.?	
Did it meet your needs for treatment?	
How was the conversation between you and the oncologist or other providers? Were you involved with the decision-making? If not, did you prefer being part of that decision-making?	

[f] Health Quality Ontario is now the Quality business unit at Ontario Health.

Receiving other types of tests prior to or after gene expression profiling to help with decision-making for treatment

How did it meet or not meet your needs? How was it adequate or not adequate?	
How long did you have to wait to receive it?	
What were the side effects?	
What were the benefits?	
What were the limitations and barriers?	
Were there issues related to cost, access, knowledge of health care system, etc.?	
Did it meet your needs for treatment?	
How was the conversation between you and the oncologist or other providers? Were you involved with the decision-making? If not, did you prefer being part of that decision-making?	

Treatment after receiving gene expression profiling test

How did it meet or not meet your needs? How was it adequate or not adequate?	
How long did you have to wait to receive it?	
What were the side effects?	
What were the benefits?	
What were the limitations and barriers?	
Were there issues related to cost, access, knowledge of health care system, etc.?	

Lived experience after receiving treatment

How is your day-to-day routine?	
What has been the impact and effect on quality of life?	
Did you see any sort of loss of independence?	
Did it have an impact on your loved ones/caregivers, work, friends?	
Do you feel more comfortable with it now as opposed to before?	

Barriers or challenges

Did you face any sort of barrier in terms of distance of travel? Accessibility of any services? Cost?	

REFERENCES

(1) Canadian Cancer Society. Risk factors for breast cancer [Internet]. Toronto (ON): The Society; 2018 [cited 2018 Aug 17]. Available from: http://www.cancer.ca/en/cancer-information/cancer-type/breast/risks

(2) Members of the Breast Expert Panel. AJCC cancer staging manual, eighth edition: part XI breast [Internet]. Chicago (IL): American College of Surgeons; 2018 [cited 2018 Aug 13]. Available from: https://cancerstaging.org/references-tools/deskreferences/Documents/AJCC%20Breast%20Cancer%20Staging%20System.pdf

(3) Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015;5(10):2929-43.

(4) Cho N. Molecular subtypes and imaging phenotypes of breast cancer. Ultrasonography. 2016;35(4):281-8.

(5) Canadian Cancer Society. Breast cancer statistics [Internet]. Toronto (ON): The Society; 2018 [cited 2018 Aug 13]. Available from: http://www.cancer.ca/en/cancer-information/cancer-type/breast/statistics

(6) Cancer Care Ontario. Ontario cancer statistics 2018 [Internet]. Toronto (ON): Cancer Care Ontario; 2018 [cited 2019 Aug 15]. Available from: https://www.cancercareontario.ca/sites/ccocancercare/files/assets/ANI_32628719_OCS2018_AODA_Mar2019_V02.pdf

(7) 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 Res Treat. 2017;165(3):709-20.

(8) Pan H, Gray R, Braybrooke J, Davies C, Taylor C, McGale P, et al. 20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years. N Engl J Med. 2017;377(19):1836-46.

(9) National Health Service. Predict [Internet]. Cambridge (UK): University of Cambridge; 2019 [cited 2019 May 9]. Available from: https://breast.predict.nhs.uk/tool

(10) Rakha EA, Soria D, Green AR, Lemetre C, Powe DG, Nolan CC, et al. Nottingham Prognostic Index Plus (NPI+): a modern clinical decision making tool in breast cancer. Br J Cancer. 2014;110(7):1688-97.

(11) Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat. 1992;22(3):207-19.

(12) Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol. 2001;19(4):980-91.

(13) Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. J Clin Oncol. 2011;29(32):4273-8.

(14) Marchionni L, Wilson RF, Marinopoulos SS, Wolff AC, Parmigiani G, Bass EB, et al. Impact of gene expression profiling tests on breast cancer outcomes [Internet]. Rockville (MD): Agency for Healthcare Research and Quality; 2008 [cited Jul 20. Available from: https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0007356

(15) Guler EN. Gene expression profiling in breast cancer and its effect on therapy selection in early-stage breast cancer. Eur J Breast Health. 2017;13(4):168-74.

(16) Tiberi D, Masucci L, Shedid D, Roy I, Vu T, Patocskai E, et al. Limitations of personalized medicine and gene assays for breast cancer. Cureus. 2017;9(3):e1100.

(17) Agendia. MammaPrint [Internet]. Amsterdam: Agendia; 2018 [cited 2018 Jul 27]. Available from: http://www.agendia.com/healthcare-professionals/breast-cancer/mammaprint

(18) Genomic Health Inc. Oncotype DX breast recurrence score: is your patient eligible? [Internet]. Redwood City (CA): Genomic Health Inc.; 2019 [cited 2019 May 9]. Available from: https://www.oncotypeiq.com/en-CA/breast-cancer/healthcare-professionals/oncotype-dx-breast-recurrence-score/is-your-patient-eligible

(19) NanoString Technologies. Prosigna [Internet]. Seattle (WA): NanoString Technologies; 2018 [cited 2018 Jul 27]. Available from: https://www.nanostring.com/diagnostics/prosigna

(20) Myriad Genetic Laboratories. Myriad EndoPredict technical specifications [Internet]. Salt Lake City (UT): Myriad Genetic Laboratories; 2018 [cited 2018 Aug 17]. Available from: https://s3.amazonaws.com/myriad-library/technical-specifications/EndoPredict+Tech+Specs.pdf

(21) Agendia. Understanding your results [Internet]. Amsterdam: Agendia; 2018 [cited 2018 Aug 17]. Available from: http://www.agendia.com/healthcare-professionals/breast-cancer/test-results

(22) Genomic Health Inc. Oncotype DX breast recurrence score: recurrence score result [Internet]. Redwood City (CA): Genomic Health Inc.; 2019 [cited 2019 May 9]. Available from: https://www.oncotypeiq.com/en-CA/breast-cancer/healthcare-professionals/oncotype-dx-breast-recurrence-score/interpreting-the-results

(23) NanoString Technologies. Use PAM50 to generate an individualized Prosigna score for every patient [Internet]. Seattle (WA): NanoString Technologies; 2018 [cited 2018 Aug 13]. Available from: https://www.nanostring.com/diagnostics/prosigna/technology/prosigna-algorithm

(24) NanoString Technologies. Package insert: Prosigna breast cancer prognostic gene signature assay [Internet]. Seattle (WA): NanoString Technologies; 2016 [cited 2018 Aug 17]. Available from: https://www.nanostring.com/application/files/8514/9090/9852/LBL_C02223_06_Package-Insert-Prosigna-Assay-US.pdf

(25) Myriad Genetic Laboratories. Test report sample [Internet]. Salt Lake City (UT): Myriad Genetic Laboratories; 2018 [cited 2018 Aug 17]. Available from: https://endopredict.com/low-test-report-sample

(26) Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N Engl J Med. 2018;379(2):111-21.

(27) Agendia. Our science [Internet]. Amsterdam: Agendia; 2018 [cited 2018 Aug 17]. Available from: http://www.agendia.com/our-science

(28) United States Food and Drug Administration. Clinical Laboratory Improvement Amendments (CLIA) [Internet]. Silver Spring (MD): United States Food and Drug Administration; 2018 [cited 2018 Jul 27]. Available from: https://www.fda.gov/medicaldevices/deviceregulationandguidance/ivdregulatoryassistance/ucm124105.htm

(29) Chang MC, Souter LH, Kamel-Reid S, Rutherford M, Bedard P, Trudeau M, et al. Clinical utility of multigene profiling assays in early-stage breast cancer. Curr Oncol. 2017;24(5):e403-e22.

(30) Chang MC, Souter LH, Kamel-Reid S, Rutherford M, Bedard P, Trudeau M, et al. Clinical utility of multigene profiling assays in invasive early-stage breast cancer [Internet]. Toronto (ON): Cancer Care Ontario; 2016 [cited 2018 July 20]. Available from: https://www.cancercareontario.ca/sites/ccocancercare/files/guidelines/full/pebcMOAC-4f.pdf

(31) Henry NL, Somerfield MR, Abramson VG, Allison KH, Anders CK, Chingos DT, et al. Role of patient and disease factors in adjuvant systemic therapy decision making for early-stage, operable breast cancer: American Society of Clinical Oncology endorsement of Cancer Care Ontario guideline recommendations. J Clin Oncol. 2016;34(19):2303-11.

(32) 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. J Clin Oncol. 2017;35(24):2838-47.

(33) National Comprehensive Cancer Network. Breast cancer [Internet]. Fort Washington (PA): The Network; 2019 [cited 2019 Apr 17]. Available from: https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf

(34) Burstein HJ, Curigliano G, Loibl S, Dubsky P, Gnant M, Poortmans P, et al. Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. Ann Oncol. 2019;30(10):1541-57.

(35) Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015;26 Suppl 5:v8-30.

(36) National Institute for Health and Care Excellence (NICE). Tumour profiling tests to guide adjuvant chemotherapy decisions in early breast cancer [Internet]. London (UK): The Institute; 2018 [cited 2019 Mar 8]. Available from: https://www.nice.org.uk/guidance/dg34

(37) Institut national d'excellence en santé et en services sociaux (INESSS). Utilisation du test Oncotype DX aux fins de décision thérapeutique dans le contexte du traitement du cancer du sein infiltrant [Internet]. Montréal (QC): Gouvernement du Québec; 2016 [cited 2019 May 9]. Available from: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Oncologie/INESSS_Test_Oncotype_CancerduSein.pdf

(38) Institut national d'excellence en santé et en services sociaux (INESSS). Utilisation d'EndoPredict et de Prosigna dans les cas de cancer du sein invasif précoce [Internet]. Montréal (QC): Gouvernement du Québec; 2018 [cited 2019 May 9]. Available from: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Oncologie/INESSS_EndoPredict-et-Prosigna_vs_Oncotype.pdf

(39) Cancer Care Ontario. Clinical utility of multigene profiling assays in invasive early-stage breast cancer [Internet]. Toronto (ON): Cancer Care Ontario; 2019 [cited 2019 May 9]. Available from: https://www.cancercareontario.ca/en/content/clinical-utility-multigene-profiling-assays-invasive-early-stage-breast-cancer

(40) Government of Alberta. Health evidence reviews [Internet]. Edmonton (AB): Government of Alberta; 2019 [cited 2019 May 9]. Available from: https://www.alberta.ca/health-evidence-reviews.aspx

(41) McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40-6.

(42) Covidence systematic review software. Veritas Health Innovation. Melbourne (Australia). Available at: https://www.covidence.org/home.

(43) Greenland S. Quantitative methods in the review of epidemiologic literature. Epidemiol Rev. 1987;9:1-30.

(44) Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br Med J. 2011;343:d5928.

(45) 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. Ann Intern Med. 2019;170(1):51-8.
(46) Kim SY, Park JE, Lee YJ, Seo HJ, Sheen SS, Hahn S, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. J Clin Epidemiol. 2013;66(4):408-14.
(47) Schünemann H, Brożek J, Guyatt G, Oxman A, editors. GRADE handbook [Internet]. Hamilton (ON): GRADE Working Group; 2013 [cited 2017 Dec]. Available from http://gdt.guidelinedevelopment.org/app/handbook/handbook.html [Internet].
(48) Curtit E, Vannetzel JM, Darmon JC, 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. Breast. 2019;44:39-45.
(49) Wuerstlein R, Kates R, Gluz O, Grischke EM, 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 Res Treat. 2019;175(2):389-99.
(50) Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(6):e1000097.
(51) 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. N Engl J Med. 2016;375(8):717-29.
(52) Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C, et al. Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. JAMA Oncol. 2018;4(4):545-53.
(53) Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, et al. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. BMJ. 2015;350:h870.
(54) Filipits M, Nielsen TO, Rudas M, Greil R, Stoger H, Jakesz R, et al. The PAM50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. Clin Cancer Res. 2014;20(5):1298-305.
(55) Nitz U, Gluz O, Christgen M, Kates RE, Clemens M, Malter W, et al. Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial. Breast Cancer Res Treat. 2017;165(3):573-83.
(56) 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. J Clin Oncol. 2006;24(23):3726-34.
(57) Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol. 2013;31(22):2783-90.
(58) 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 Res Treat. 2015;149(2):439-48.
(59) 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 Res. 2017;19(1):120.
(60) 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. J Clin Oncol. 2015;33(8):916-22.

(61) Buus R, Sestak I, Kronenwett R, Denkert C, Dubsky P, Krappmann K, et al. Comparison of EndoPredict and EPclin with Oncotype DX recurrence score for prediction of risk of distant recurrence after endocrine therapy. J Natl Cancer Inst. 2016;108(11).

(62) Filipits M, Dubsky P, Rudas M, Greil R, Balic M, Bago-Horvath Z, et al. Prediction of distant recurrence using EndoPredict among women with ER+, HER2- node-positive and node-negative breast cancer treated with endocrine therapy only. Clin Cancer Res. 2019.

(63) Drukker CA, Bueno-de-Mesquita JM, Retel 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. Int J Cancer. 2013;133(4):929-36.

(64) Drukker CA, Nijenhuis MV, Bueno-de-Mesquita JM, Retel 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 Res Treat. 2014;145(3):697-705.

(65) 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 ultra-low risk over 2 decades. JAMA Oncol. 2017;3(11):1503-10.

(66) 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. N Engl J Med. 2002;347(25):1999-2009.

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

(68) 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. J Clin Oncol. 2010;28(11):1829-34.

(69) 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. J Clin Oncol. 2010;28(10):1677-83.

(70) 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. N Engl J Med. 2004;351(27):2817-26.

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

(72) 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;3:33.

(73) 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. Ann Oncol. 2014;25(2):339-45.

(74) Laenkholm AV, Jensen MB, Eriksen JO, Rasmussen BB, Knoop AS, Buckingham W, et al. PAM50 risk of recurrence score predicts 10-year distant recurrence in a comprehensive Danish cohort of postmenopausal women allocated to 5 years of

endocrine therapy for hormone receptor-positive early breast cancer. J Clin Oncol. 2018;36(8):735-40.
(75) 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.
(76) 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.
(77) Albanell J, Gonzalez A, Ruiz-Borrego M, Alba E, Garcia-Saenz JA, Corominas JM, 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. Ann Oncol. 2012;23(3):625-31.
(78) Kuijer A, Straver M, den Dekker B, van Bommel ACM, Elias SG, Smorenburg CH, et al. Impact of 70-gene signature use on adjuvant chemotherapy decisions in patients with estrogen receptor-positive early breast cancer: results of a prospective cohort study. J Clin Oncol. 2017;35(24):2814-9.
(79) Albanell J, Svedman C, Gligorov J, Holt SD, Bertelli G, Blohmer JU, 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. Eur J Cancer. 2016;66:104-13.
(80) Bargallo JE, Lara F, Shaw-Dulin R, Perez-Sanchez V, Villarreal-Garza C, Maldonado-Martinez H, et al. A study of the impact of the 21-gene breast cancer assay on the use of adjuvant chemotherapy in women with breast cancer in a Mexican public hospital. J Surg Oncol. 2015;111(2):203-7.
(81) de Boer RH, Baker C, Speakman D, Chao CY, Yoshizawa C, Mann GB. The impact of a genomic assay (Oncotype DX) on adjuvant treatment recommendations in early breast cancer. Med J Aust. 2013;199(3):205-8.
(82) Dieci MV, Guarneri V, Giarratano T, Mion M, Tortora G, De Rossi C, et al. First prospective multicenter Italian study on the impact of the 21-gene recurrence score in adjuvant clinical decisions for patients with ER-positive/HER2-negative breast cancer. Oncologist. 2018;23(3):297-305.
(83) Eiermann W, Rezai M, Kummel S, Kuhn T, Warm M, Friedrichs K, et al. The 21-gene recurrence score assay impacts adjuvant therapy recommendations for ER-positive, node-negative and node-positive early breast cancer resulting in a risk-adapted change in chemotherapy use. Ann Oncol. 2013;24(3):618-24.
(84) Levine MN, Julian JA, Bedard PL, Eisen A, Trudeau ME, Higgins B, et al. Prospective evaluation of the 21-gene recurrence score assay for breast cancer decision-making in Ontario. J Clin Oncol. 2016;34(10):1065-71.
(85) Lo SS, Mumby PB, Norton J, Rychlik K, Smerage J, Kash J, et al. Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. J Clin Oncol. 2010;28(10):1671-6.
(86) Loncaster J, Armstrong A, Howell S, Wilson G, Welch R, Chittalia A, et al. Impact of Oncotype DX breast recurrence score testing on adjuvant chemotherapy use in early breast cancer: real world experience in Greater Manchester, UK. [Erratum appears in Eur J Surg Oncol. 2017 Nov 23; PMID: 29174199]. Eur J Surg Oncol. 2017;43(5):931-7.
(87) Ozmen V, Atasoy A, Gokmen E, Ozdogan M, Guler N, Uras C, et al. Impact of Oncotype DX recurrence score on treatment decisions: results of a prospective multicenter study in Turkey. Cureus. 2016;8(3):e522.

(88) Hequet D, Callens C, Gentien D, Albaud B, Mouret-Reynier MA, 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.
(89) Wuerstlein R, Sotlar K, Gluz O, Otremba B, von Schumann R, Witzel I, et al. The West German Study Group Breast Cancer Intrinsic Subtype study: a prospective multicenter decision impact study utilizing the Prosigna assay for adjuvant treatment decision-making in estrogen-receptor-positive, HER2-negative early-stage breast cancer. Curr Med Res Opin. 2016;32(7):1217-24.
(90) Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol. 2010;11(1):55-65.
(91) Dubsky P, Brase JC, Jakesz R, Rudas M, Singer CF, Greil R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. Br J Cancer. 2013;109(12):2959-64.
(92) 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. Br J Cancer. 2015;112(8):1405-10.
(93) Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L, et al. Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study. Breast Cancer Res Treat. 2016;156(1):81-9.
(94) 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. J Clin Oncol. 2016;34(20):2341-9.
(95) 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. J Clin Oncol. 2016;34(20):2359-65.
(96) Mamounas EP, Liu Q, Paik S, Baehner FL, Tang G, Jeong JH, et al. 21-gene recurrence score and locoregional recurrence in node-positive/er-positive breast cancer treated with chemo-endocrine therapy. J Natl Cancer Inst. 2017;109(4).
(97) 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):526.
(98) Roberts MC, Miller DP, Shak S, Petkov VI. Breast cancer-specific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX recurrence score results in the SEER database. Breast Cancer Res Treat. 2017;163(2):303-10.
(99) Stemmer SM, Steiner M, Rizel S, Geffen DB, Nisenbaum B, Peretz T, et al. Clinical outcomes in ER+ HER2 -node-positive breast cancer patients who were treated according to the recurrence score results: evidence from a large prospectively designed registry. NPJ Breast Cancer. 2017;3:32.
(100) 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. J Natl Cancer Inst. 2013;105(19):1504-11.

(101) Jensen MB, Laenkholm AV, 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 Res. 2018;20(1):79.
(102) Sestak I, Martin M, Dubsky P, Kronenwett R, Rojo F, Cuzick J, et al. Prediction of chemotherapy benefit by EndoPredict in patients with breast cancer who received adjuvant endocrine therapy plus chemotherapy or endocrine therapy alone. Breast Cancer Res Treat. 2019;176(2):376-86.
(103) Torres S, Trudeau M, Gandhi S, Warner E, Verma S, Pritchard KI, et al. Prospective evaluation of the impact of the 21-gene recurrence score assay on adjuvant treatment decisions for women with node-positive breast cancer in Ontario, Canada. Oncologist. 2018;23(7):768-75.
(104) 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.
(105) Fallowfield L, Matthews L, May S, Jenkins V, Bloomfield D. Enhancing decision-making about adjuvant chemotherapy in early breast cancer following EndoPredict testing. Psychooncology. 2018;27(4):1264-9.
(106) Mokbel K, Wazir U, El Hage Chehade H, 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 Res. 2017;37(12):6863-9.
(107) Mokbel K, Wazir U, Wazir A, Kasem A, Mokbel K. The impact of Endopredict clinical score on chemotherapy recommendations in women with invasive ER(+)/HER2(-) breast cancer stratified as having moderate or poor prognosis by Nottingham Prognostic Index. Anticancer Res. 2018;38(8):4747-52.
(108) Cusumano PG, Generali D, Ciruelos E, Manso L, Ghanem I, Lifrange E, et al. European inter-institutional impact study of MammaPrint. Breast. 2014;23(4):423-8.
(109) Tsai M, Lo S, Audeh W, Qamar R, Budway R, Levine E, et al. Association of 70-gene signature assay findings with physicians' treatment guidance for patients with early breast cancer classified as intermediate risk by the 21-gene assay. JAMA Oncol. 2018;4(1):e173470.
(110) Evans CN, Brewer NT, Vadaparampil ST, Boisvert M, Ottaviano Y, Lee MC, et al. Impact of genomic testing and patient-reported outcomes on receipt of adjuvant chemotherapy. Breast Cancer Res Treat. 2016;156(3):549-55.
(111) 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. Br J Cancer. 2016;114(7):731-6.
(112) Leung RC, Yau TC, Chan MC, Chan SW, Chan TW, Tsang YY, et al. The impact of the Oncotype DX breast cancer assay on treatment decisions for women with estrogen receptor-positive, node-negative breast carcinoma in Hong Kong. Clin Breast Cancer. 2016;16(5):372-8.
(113) Martinez Del Prado P, Alvarez-Lopez I, Dominguez-Fernandez 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. Clinicoecon Outcomes Res. 2018;10:189-99.
(114) 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):265.

(115) Voelker HU, 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. Int J Breast Cancer. 2018;2018:2047089.

(116) ISRCTN Registry. Optimal personalised treatment of early breast cancer using multiparameter analysis [Internet]. London (UK): Springer Nature; 2019 [cited 2019 May 9]. Available from: http://www.isrctn.com/ISRCTN42400492

(117) 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 Treat Rev. 2018;62:74-90.

(118) National Institute for Health and Care Excellence. Tumour profiling tests to guide adjuvant chemotherapy decisions in people with breast cancer (update of DG10) [Internet]. London (UK): The Institute; 2018 [cited 2018 Oct 1st]. Available from: https://www.nice.org.uk/guidance/gid-dg10015/documents/diagnostics-assessment-report-2

(119) National Institute for Health and Care Excellence. Process and methods guides. Appendix I: Quality appraisal checklist–economic evaluations [Internet]. London (UK): The Institute; 2012 [cited 2016 Jan]. Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-i-quality-appraisal-checklist-economic-evaluations

(120) Chandler Y, Schechter CB, Jayasekera J, Near A, O'Neill SC, Isaacs C, et al. Cost effectiveness of gene expression profile testing in community practice. J Clin Oncol. 2018;36(6):554-62.

(121) Dinan MA, Wilson LE, Reed SD. Chemotherapy costs and 21-gene recurrence score genomic testing among Medicare beneficiaries with early-stage breast cancer, 2005 to 2011. J Natl Compr Canc Netw. 2019;17(3):245-54.

(122) Hinde S, Theriou C, May S, Matthews L, Arbon A, Fallowfield L, et al. The cost-effectiveness of EndoPredict to inform adjuvant chemotherapy decisions in early breast cancer. Health Policy Technol. 2019.

(123) Ibarrondo O, Alvarez-Lopez I, Freundlich F, Arrospide A, Galve-Calvo E, Gutierrez-Toribio M, et al. Probabilistic cost-utility analysis and expected value of perfect information for the Oncotype multigenic test: a discrete event simulation model. Gac Sanit. 2018;12:12.

(124) Lux MP, Nabieva N, Hildebrandt T, Rebscher H, Kummel S, Blohmer JU, et al. Budget impact analysis of gene expression tests to aid therapy decisions for breast cancer patients in Germany. Breast. 2018;37:89-98.

(125) Martel S, Lambertini M, Simon R, Matte C, Prady C. Adherence to guidelines in requesting oncotype DX in a publicly funded health care system. Curr Oncol. 2018;25(4):e311-e8.

(126) Mittmann N, Earle CC, Cheng SY, Julian JA, Rahman F, Seung SJ, et al. Population-based study to determine the health system costs of using the 21-gene assay. J Clin Oncol. 2018;36(3):238-43.

(127) Waintraub SE, McNamara D, Graham DMA, Pecora AL, Min J, Wu T, et al. Real-world economic value of a 21-gene assay in early-stage breast cancer. Am J Manag Care. 2017;23(12):e416-e20.

(128) Wang SY, Chen T, Dang W, Mougalian SS, Evans SB, Gross CP. Incorporating tumor characteristics to maximize 21-gene assay utility: a cost-effectiveness analysis. J Natl Compr Canc Netw. 2019;17(1):39-46.

(129) Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: an evidence-based and economic analysis. Ont Health Technol Assess Ser. 2010;10(23):1-57.

(130) Bacchi CE, Prisco F, Carvalho FM, Ojopi EB, Saad ED. Potential economic impact of the 21-gene expression assay on the treatment of breast cancer in Brazil. Rev Assoc Med Bras. 2010;56(2):186-91.
(131) Bargallo-Rocha JE, Lara-Medina F, Perez-Sanchez V, Vazquez-Romo R, Villarreal-Garza C, Martinez-Said H, et al. Cost-effectiveness of the 21-gene breast cancer assay in Mexico. Adv Ther. 2015;32(3):239-53.
(132) Blank PR, Filipits M, Dubsky P, Gutzwiller F, Lux MP, Brase JC, et al. Cost-effectiveness analysis of prognostic gene expression signature-based stratification of early breast cancer patients. Pharmacoeconomics. 2015;33(2):179-90.
(133) Blohmer JU, Rezai M, Kummel S, Kuhn T, Warm M, Friedrichs K, et al. Using the 21-gene assay to guide adjuvant chemotherapy decision-making in early-stage breast cancer: a cost-effectiveness evaluation in the German setting. J Med Econ. 2013;16(1):30-40.
(134) Bonastre J, Marguet S, Lueza B, Michiels S, Delaloge S, Saghatchian M. Cost effectiveness of molecular profiling for adjuvant decision making in patients with node-negative breast cancer. J Clin Oncol. 2014;32(31):3513-9.
(135) Chen E, Tong KB, Malin JL. Cost-effectiveness of 70-gene MammaPrint signature in node-negative breast cancer. Am J Manag Care. 2010;16(12):e333-42.
(136) Cosler LE, Lyman GH. Economic analysis of gene expression profile data to guide adjuvant treatment in women with early-stage breast cancer. Cancer Invest. 2009;27(10):953-9.
(137) Davidson JA, Cromwell I, Ellard SL, Lohrisch C, Gelmon KA, Shenkier T, et al. A prospective clinical utility and pharmacoeconomic study of the impact of the 21-gene recurrence score assay in oestrogen receptor positive node negative breast cancer. Eur J Cancer. 2013;49(11):2469-75.
(138) Epstein AJ, Wong YN, Mitra N, Vachani A, Hin S, Yang L, et al. Adjuvant chemotherapy use and health care costs after introduction of genomic testing in breast cancer. J Clin Oncol. 2015;33(36):4259-67.
(139) Exner R, Bago-Horvath Z, Bartsch R, Mittlboeck M, Retel VP, Fitzal F, et al. The multigene signature MammaPrint impacts on multidisciplinary team decisions in ER+, HER2- early breast cancer. Br J Cancer. 2014;111(5):837-42.
(140) Hall PS, McCabe C, Stein RC, Cameron D. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. J Natl Cancer Inst. 2012;104(1):56-66.
(141) Hannouf MB, Xie B, Brackstone M, Zaric GS. Cost-effectiveness of a 21-gene recurrence score assay versus Canadian clinical practice in women with early-stage estrogen- or progesterone-receptor-positive, axillary lymph-node negative breast cancer. BMC Cancer. 2012;12:447.
(142) Hannouf MB, Xie B, Brackstone M, Zaric GS. Cost effectiveness of a 21-gene recurrence score assay versus Canadian clinical practice in post-menopausal women with early-stage estrogen or progesterone-receptor-positive, axillary lymph-node positive breast cancer. Pharmacoeconomics. 2014;32(2):135-47.
(143) Holt S, Bertelli G, Humphreys I, Valentine W, Durrani S, Pudney D, et al. A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pNImi, ER-positive breast cancer in the U.K. Br J Cancer. 2013;108(11):2250-8.
(144) Hornberger J, Chien R, Krebs K, Hochheiser L. US insurance program's experience with a multigene assay for early-stage breast cancer. Am J Manag Care. 2011;17(5 Spec No):e194-202.

(145) Hornberger J, Cosler LE, Lyman GH. Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. Am J Manag Care. 2005;11(5):313-24.
(146) Jahn B, Rochau U, Kurzthaler C, Hubalek M, Miksad R, Sroczynski G, et al. Cost effectiveness of personalized treatment in women with early breast cancer: the application of OncotypeDX and Adjuvant! Online to guide adjuvant chemotherapy in Austria. SpringerPlus. 2015;4:752.
(147) Katz G, Romano O, Foa C, Vataire AL, Chantelard JV, Herve R, et al. Economic impact of gene expression profiling in patients with early-stage breast cancer in France. PLoS One. 2015;10(6):e0128880.
(148) Kip M, Monteban H, Steuten L. Long-term cost-effectiveness of Oncotype DX versus current clinical practice from a Dutch cost perspective. J Comp Eff Res. 2015;4(5):433-45.
(149) Klang SH, Hammerman A, Liebermann N, Efrat N, Doberne J, Hornberger J. Economic implications of 21-gene breast cancer risk assay from the perspective of an Israeli-managed health-care organization. Value Health. 2010;13(4):381-7.
(150) Kondo M, Hoshi SL, Ishiguro H, Toi M. Economic evaluation of the 70-gene prognosis-signature (MammaPrint) in hormone receptor-positive, lymph node-negative, human epidermal growth factor receptor type 2-negative early stage breast cancer in Japan. Breast Cancer Res Treat. 2012;133(2):759-68.
(151) Kondo M, Hoshi SL, Ishiguro H, Yoshibayashi H, Toi M. Economic evaluation of 21-gene reverse transcriptase-polymerase chain reaction assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer in Japan. Breast Cancer Res Treat. 2008;112(1):175-87.
(152) Kondo M, Hoshi SL, Yamanaka T, Ishiguro H, Toi M. Economic evaluation of the 21-gene signature (Oncotype DX) in lymph node-negative/positive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03). Breast Cancer Res Treat. 2011;127(3):739-49.
(153) Lamond NW, Skedgel C, Rayson D, Lethbridge L, Younis T. Cost-utility of the 21-gene recurrence score assay in node-negative and node-positive breast cancer. Breast Cancer Res Treat. 2012;133(3):1115-23.
(154) Lyman GH, Cosler LE, Kuderer NM, Hornberger J. Impact of a 21-gene RT-PCR assay on treatment decisions in early-stage breast cancer: an economic analysis based on prognostic and predictive validation studies. Cancer. 2007;109(6):1011-8.
(155) Mislick K, Schonfeld W, Bodnar C, Tong KB. Cost-effectiveness analysis of Mammostrat compared with Oncotype DX to inform the treatment of breast cancer. Clinicoecon Outcomes Res. 2014;6:37-47.
(156) Nerich V, Curtit E, Bazan F, Montcuquet P, Villanueva C, Chaigneau L, et al. [Economic assessment of the routine use of Oncotype DX assay for early breast cancer in Franche-Comte region]. Bull Cancer. 2014;101(7-8):681-9.
(157) Oestreicher N, Ramsey SD, Linden HM, McCune JS, van't Veer LJ, Burke W, et al. Gene expression profiling and breast cancer care: what are the potential benefits and policy implications? Genet Med. 2005;7(6):380-9.
(158) Paulden M, Franek J, Pham B, Bedard PL, Trudeau M, Krahn M. Cost-effectiveness of the 21-gene assay for guiding adjuvant chemotherapy decisions in early breast cancer. Value Health. 2013;16(5):729-39.
(159) Reed SD, Dinan MA, Schulman KA, Lyman GH. Cost-effectiveness of the 21-gene recurrence score assay in the context of multifactorial decision making to guide chemotherapy for early-stage breast cancer. Genet Med. 2013;15(3):203-11.

(160) Retel VP, Grutters JP, van Harten WH, Joore MA. Value of research and value of development in early assessments of new medical technologies. Value Health. 2013;16(5):720-8.
(161) Retel VP, Joore MA, Drukker CA, Bueno-de-Mesquita JM, Knauer M, van Tinteren H, et al. Prospective cost-effectiveness analysis of genomic profiling in breast cancer. Eur J Cancer. 2013;49(18):3773-9.
(162) Retel VP, Joore MA, Knauer M, Linn SC, Hauptmann M, Harten WH. Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. Eur J Cancer. 2010;46(8):1382-91.
(163) Retel VP, Joore MA, van Harten WH. Head-to-head comparison of the 70-gene signature versus the 21-gene assay: cost-effectiveness and the effect of compliance. Breast Cancer Res Treat. 2012;131(2):627-36.
(164) Segui MA, Crespo C, Cortes J, Lluch A, Brosa M, Becerra V, et al. Genomic profile of breast cancer: cost-effectiveness analysis from the Spanish National Healthcare System perspective. Expert Rev Pharmacoecon Outcomes Res. 2014;14(6):889-99.
(165) 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 Res Treat. 2015;153(3):573-82.
(166) 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 Technol Assess. 2016;20(10):xxiii-xxix, 1-201.
(167) Su KW, Hall J, Soulos PR, Abu-Khalaf MM, Evans SB, Mougalian SS, et al. Association of 21-gene recurrence score assay and adjuvant chemotherapy use in the medicare population, 2008-2011. J Geriatr Oncol. 2016;7(1):15-23.
(168) Tiwana KS, Smith A, Leggett L, MacKean G, Lorenzetti D, Clement F. Use of Oncotype DX for guiding adjuvant chemotherapy decisions in early stage invasive breast cancer patients in Alberta: a health technology assessment [Internet]. Calgary (AB): Alberta Health; 2013 [cited 2018 Oct 1]. Available from: https://open.alberta.ca/publications/use-of-oncotype-dx-for-guiding-adjuvant-chemotherapy-decisions
(169) Tsoi DT, Inoue M, Kelly CM, Verma S, Pritchard KI. Cost-effectiveness analysis of recurrence score-guided treatment using a 21-gene assay in early breast cancer. Oncologist. 2010;15(5):457-65.
(170) Vanderlaan BF, Broder MS, Chang EY, Oratz R, Bentley TG. Cost-effectiveness of 21-gene assay in node-positive, early-stage breast cancer. Am J Manag Care. 2011;17(7):455-64.
(171) Vataire AL, Laas E, Aballea S, Gligorov J, Rouzier R, Chereau E. [Cost-effectiveness of a chemotherapy predictive test]. Bull Cancer. 2012;99(10):907-14.
(172) Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P, et al. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: a systematic review and cost-effectiveness analysis. Health Technol Assess. 2013;17(44):1-302.
(173) Yamauchi H, Nakagawa C, Yamashige S, Takei H, Yagata H, Yoshida A, et al. Societal cost-effectiveness analysis of the 21-gene assay in estrogen-receptor-positive, lymph-node-negative early-stage breast cancer in Japan. BMC Health Serv Res. 2014;14:372.
(174) Yang M, Rajan S, Issa AM. Cost effectiveness of gene expression profiling for early stage breast cancer: a decision-analytic model. Cancer. 2012;118(20):5163-70.
(175) Hannouf MB, Zaric GS, Blanchette P, Brezden-Masley C, Paulden M, McCabe C, et al. Cost-effectiveness analysis of multigene expression profiling assays to guide adjuvant therapy decisions in women with invasive early-stage breast cancer. Pharmacogenomics J. Forthcoming 2019.

(176) Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health. 2013;16(2):231-50.
(177) Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada. 4th ed [Internet]. Ottawa (ON): The Agency; 2017 [cited 2018 Jan]. Available from: https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf
(178) Institute of Health Economics. Clinical effectiveness and cost-effectiveness of Oncotype DX and Prosigna genetic testing in early-stage breast cancer. 2019(In Press).
(179) Partridge AH, Rumble RB, Carey LA, Come SE, Davidson NE, Di Leo A, et al. Chemotherapy and targeted therapy for women with human epidermal growth factor receptor 2-negative (or unknown) advanced breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2014;32(29):3307-29.
(180) 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. Lancet. 2012;379(9814):432-44.
(181) Canadian Agency for Drugs and Technologies in Health (CADTH). Oncotype DX in women and men with ER-positive HER2-negative early stage breast cancer who are lymph node-positive: a review of clinical effectiveness and guidelines [Internet]. Ottawa (ON): The Agency; 2014 [cited 2018 Jul 20]. Available from: https://www.cadth.ca/sites/default/files/pdf/htis/mar-2014/RC0517_OncotypeDX_NodePos%20Final.pdf
(182) Ludwig Breast Cancer Study Group. Prolonged disease-free survival after one course of perioperative adjuvant chemotherapy for node-negative breast cancer. N Engl J Med. 1989;320(8):491-6.
(183) Azim Jr. HA, Michiels S, Zagouri F, Delaloge S, Filipits M, Namer M, et al. Utility of prognostic genomic tests in breast cancer practice: the IMPAKT 2012 Working Group Consensus Statement. Ann Oncol. 2013;24(3):647-54.
(184) Leone BA, Vallejo CT, Romero AO, Machiavelli MR, Perez JE, Leone J, et al. Prognostic impact of metastatic pattern in stage IV breast cancer at initial diagnosis. Breast Cancer Res Treat. 2017;161(3):537-48.
(185) Lidgren M, Wilking N, Jonsson B, Rehnberg C. Health related quality of life in different states of breast cancer. Qual Life Res. 2007;16(6):1073-81.
(186) Consumer Price Index, monthly, percentage change, not seasonally adjusted, Canada, provinces, Whitehorse and Yellowknife—health and personal care [Internet]. Ottawa: Bank of Canada; c2019 [cited 2018 Nov 15]. Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810000408
(187) Currency converter [Internet]. Ottawa: Bank of Canada; c2019 [cited 2019 Feb 1]. Available from: https://www.bankofcanada.ca/rates/exchange/currency-converter/
(188) Mittmann N, Beglaryan H, Liu N, Seung SJ, Rahman F, Gilbert J, et al. Examination of health system resources and costs associated with transitioning cancer survivors to primary care: a propensity-score-matched cohort study. J Oncol Pract. 2018:Jop1800275.
(189) Will BP, Berthelot JM, Le Petit C, Tomiak EM, Verma S, Evans WK. Estimates of the lifetime costs of breast cancer treatment in Canada. Eur J Cancer. 2000;36(6):724-35.
(190) de Bock GH, Putter H, Bonnema J, van der Hage JA, Bartelink H, van de Velde CJ. The impact of loco-regional recurrences on metastatic progression in early-stage breast cancer: a multistate model. Breast Cancer Res Treat. 2009;117(2):401-8.

(191) Newton EV, Grethlein SJ. Adjuvant therapy for breast cancer [Internet]. New York: MedSCAPE; c2018 [updated Feb 7, 2018; cited 2018 Nov 1]. Available from: https://emedicine.medscape.com/article/1946040-overview

(192) Masucci L, Torres S, Eisen A, Trudeau M, Tyono I, Saunders H, et al. Cost-utility analysis of 21-gene assay for node-positive early breast cancer. Curr Oncol. 2019;26(5):307-18.

(193) Tzeng JP, Mayer D, Richman AR, Lipkus I, Han PK, Valle CG, et al. Women's experiences with genomic testing for breast cancer recurrence risk. Cancer. 2010;116(8):1992-2000.

(194) Ontario population projections update, 2017-2041 [Internet]. Toronto (ON): Queen's Printer for Ontario; c2010 [updated 2018 June 25; cited 2018 Oct 19]. Available from: https://www.fin.gov.on.ca/en/economy/demographics/projections/table7.html

(195) Breast cancer statistics [Internet]. Toronto (ON): Canadian Cancer Society; c2018 [cited 2018 Oct 19]. Available from: http://www.cancer.ca/en/cancer-information/cancer-type/breast/statistics/?region=bc

(196) Fallahpour S, Navaneelan T, De P, Borgo A. Breast cancer survival by molecular subtype: a population-based analysis of cancer registry data. CMAJ Open. 2017;5(3):E734-E9.

(197) Breast cancer in men [Internet]. Toronto (ON): Canadian Cancer Society; c2018 [cited 2018 Nov 20]. Available from: http://www.cancer.ca/en/cancer-information/cancer-type/breast/breast-cancer/breast-cancer-in-men/?region=on

(198) Smith A, Farrah K. Gene expression profiling tests for breast cancer: a rapid qualitative review [Internet]. Toronto (ON): Canadian Agency for Drugs and Technologies in Health; 2019 [cited 2019 May 9]. Available from: https://www.cadth.ca/sites/default/files/rr/2019/RC1099%20Gene%20Profiling%20BrCa%20Final.pdf

(199) Selva A, Sola I, Zhang Y, Pardo-Hernandez H, Haynes RB, Martinez Garcia L, et al. Development and use of a content search strategy for retrieving studies on patients' views and preferences. Health Qual Life Outcomes. 2017;15(1):126.

(200) O'Neill SC, Brewer NT, Lillie SE, Morrill EF, Dees EC, Carey LA, et al. Women's interest in gene expression analysis for breast cancer recurrence risk. J Clin Oncol. 2007;25(29):4628-34.

(201) Brewer NT, Edwards AS, O'Neill SC, Tzeng JP, Carey LA, Rimer BK. When genomic and standard test results diverge: implications for breast cancer patients' preference for chemotherapy. Breast Cancer Res Treat. 2009;117(1):25-9.

(202) Brewer NT, Tzeng JP, Lillie SE, Edwards AS, Peppercorn JM, Rimer BK. Health literacy and cancer risk perception: implications for genomic risk communication. Med Decis Making. 2009;29(2):157-66.

(203) Martin M, Gonzalez-Rivera M, Morales S, de la Haba-Rodriguez J, Gonzalez-Cortijo L, Manso L, et al. Prospective study of the impact of the Prosigna assay on adjuvant clinical decision-making in unselected patients with estrogen receptor positive, human epidermal growth factor receptor negative, node negative early-stage breast cancer. Curr Med Res Opin. 2015;31(6):1129-37.

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

(205) Retel VP, Bueno-de-Mesquita JM, Hummel MJ, van de Vijver MJ, Douma KF, Karsenberg K, et al. Constructive technology assessment (CTA) as a tool in coverage with evidence development: the case of the 70-gene prognosis signature for breast cancer diagnostics. Int J Technol Assess Health Care. 2009;25(1):73-83.

(206) Leggett LE, Lorenzetti DL, Noseworthy T, Tiwana S, Mackean G, Clement F. Experiences and attitudes toward risk of recurrence testing in women with breast cancer: a systematic review. Breast Cancer Res Treat. 2014;144(3):457-65.

(207) Yanes T, Willis AM, Meiser B, Tucker KM, Best M. Psychosocial and behavioral outcomes of genomic testing in cancer: a systematic review. Eur J Hum Genet. 2019;27(1):28-35.

(208) DeFrank JT, Carey LA, Brewer NT. Understanding how breast cancer patients use risk information from genomic tests. J Behav Med. 2013;36(6):567-73.

(209) DeFrank JT, Salz T, Reeder-Hayes K, Brewer NT. Who gets genomic testing for breast cancer recurrence risk? Public Health Genomics. 2013;16(5):215-22.

(210) Panattoni L, Lieu TA, Jayasekera J, O'Neill S, Mandelblatt JS, Etzioni R, et al. The impact of gene expression profile testing on confidence in chemotherapy decisions and prognostic expectations. Breast Cancer Res Treat. 2019;173(2):417-27.

(211) Friese CR, Li Y, Bondarenko I, Hofer TP, Ward KC, Hamilton AS, et al. Chemotherapy decisions and patient experience with the recurrence score assay for early-stage breast cancer. Cancer. 2017;123(1):43-51.

(212) Gligorov J, Pivot XB, Jacot W, Naman HL, Spaeth D, Misset JL, et al. Prospective clinical utility study of the use of the 21-gene assay in adjuvant clinical decision making in women with estrogen receptor-positive early invasive breast cancer: results from the SWITCH study. Oncologist. 2015;20(8):873-9.

(213) Kurian AW, Bondarenko I, Jagsi R, Friese CR, McLeod MC, Hawley ST, et al. Recent trends in chemotherapy use and oncologists' treatment recommendations for early-stage breast cancer. J Natl Cancer Inst. 2018;110(5):493-500.

(214) Lillie SE, Brewer NT, O'Neill SC, Morrill EF, Dees EC, Carey LA, et al. Retention and use of breast cancer recurrence risk information from genomic tests: the role of health literacy. Cancer Epidemiol Biomarkers Prev. 2007;16(2):249-55.

(215) Lipkus IM, Vadaparampil ST, Jacobsen PB, Miree CA. Knowledge about genomic recurrence risk testing among breast cancer survivors. J Cancer Educ. 2011;26(4):664-9.

(216) Murciano-Goroff YR, McCarthy AM, Bristol MN, Groeneveld P, Domchek SM, Motanya UN, et al. Uptake of BRCA 1/2 and Oncotype DX testing by medical and surgical oncologists. Breast Cancer Res Treat. 2018;171(1):173-80.

(217) Ngoi N, Lee SC, Hartman M, Khin LW, Wong A. Interest and attitudes of patients, cancer physicians, medical students and cancer researchers towards a spectrum of genetic tests relevant to breast cancer patients. Breast. 2013;22(1):47-52.

(218) Patil D, Issa AM. Factors affecting the adoption and use of gene expression profiling by oncologists in clinical practice. Per Med. 2015;12(1):33-42.

(219) Richman AR, Tzeng JP, Carey LA, Retel VP, Brewer NT. Knowledge of genomic testing among early-stage breast cancer patients. Psychooncology. 2011;20(1):28-35.

(220) Seror V, Marino P, Bertucci F, Mancini J, Extra JM, Ferrero JM, et al. Breast cancer patients' views on the use of genomic testing to guide decisions about their postoperative chemotherapy. Public Health Genomics. 2013;16(3):110-7.

(221) Sulayman N, Spellman E, Graves KD, Peshkin BN, Isaacs C, Schwartz MD, et al. Psychosocial and quality of life in women receiving the 21-gene recurrence score assay: the impact of decision style in women with intermediate RS. J Cancer Epidemiol. 2012;2012:728290.

(222) Yamauchi H, Nakagawa C, Takei H, Chao C, Yoshizawa C, Yagata H, et al. Prospective study of the effect of the 21-gene assay on adjuvant clinical decision-making in Japanese women with estrogen receptor-positive, node-negative, and node-positive breast cancer. Clin Breast Cancer. 2014;14(3):191-7.

(223) Barham L. Public and patient involvement at the UK National Institute for Health and Clinical Excellence. Patient. 2011;4(1):1-10.

(224) Messina J, Grainger DL. A pilot study to identify areas for further improvements in patient and public involvement in health technology assessments for medicines. Patient. 2012;5(3):199-211.

(225) Ontario Health Technology Advisory Committee Public Engagement Subcommittee. Public engagement for health technology assessment at Health Quality Ontario—final report from the Ontario Health Technology Advisory Committee Public Engagement Subcommittee [Internet]. Toronto (ON): Queen's Printer for Ontario; 2015 Apr [cited 2018 Apr 30]. Available from: http://www.hqontario.ca/Portals/0/documents/evidence/special-reports/report-subcommittee-20150407-en.pdf

(226) Kvale S. Interviews: an introduction to qualitative research interviewing. Thousand Oaks (CA): Sage; 1996.

(227) Kuzel AJ. Sampling in qualitative inquiry. In: Miller WL, Crabtree BF, editors. Doing qualitative research. Thousand Oaks (CA): Sage; 1999. p. 33-45.

(228) Morse J. Emerging from the data: cognitive processes of analysis in qualitative research. In: Morse J, editor. Critical issues in qualitative research methods. Thousand Oaks (CA): Sage; 1994. p. 23-41.

(229) Patton MQ. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks (CA): Sage; 2002.

(230) Health Technology Assessment International. Introduction to health technology assessment [Internet]. Edmonton (AB): Health Technology Assessment International; 2015 [cited 2018 Apr 30]. Available from: http://www.htai.org/fileadmin/HTAi_Files/ISG/PatientInvolvement/v2_files/Resource/PCISG-Resource-Intro_to_HTA__KFacey__Jun13.pdf

(231) Strauss AL, Corbin JM. Grounded theory research: procedures, canons, and evaluative criteria. Qual Sociol. 1990;13(1):3-21.

(232) Strauss AL, Corbin JM. Grounded theory methodology: an overview. In: Denzin NK, Lincoln YS, editors. Handbook of qualitative research. Thousand Oaks (CA): Sage; 1994. p. 273-85.

(233) NVivo qualitative data analysis software. QSR International. Doncaster, Victoria (Australia). Available at: https://www.qsrinternational.com/nvivo/home.

(234) Alberta Health Services. Adjuvant systemic therapy for early stage (lymph node negative and lymph node positive) breast cancer [Internet]. Edmonton (AB): Alberta Health Services; 2018 [cited 2018 Jul 20]. Available from: https://www.albertahealthservices.ca/assets/info/hp/cancer/if-hp-cancer-guide-adjuvant-systemic-therapy-breast.pdf

(235) BC Cancer. Laboratory services [Internet]. Vancouver (BC): British Columbia Provincial Health Services Authority; 2019 [cited 2019 May 9]. Available from: http://www.bccancer.bc.ca/health-professionals/clinical-resources/laboratory-services

(236) Nova Scotia Health Authority. Oncotype DX breast cancer test [Internet]. Halifax (NS): Nova Scotia Health Authority; 2019 [cited 2019 May 9]. Available from: http://www.cdha.nshealth.ca/nova-scotia-cancer-care-program-6

(237) Eisen A, Fletcher GG, Gandhi S, Mates M, Freedman OC, Dent SF, et al. Optimal systemic therapy for early female breast cancer [Internet]. Toronto (ON): Cancer Care Ontario; 2014 [cited 2018 Jul 20]. Available from: https://www.cancercareontario.ca/en/guidelines-advice/types-of-cancer/331

(238) Eisen A, Fletcher GG, Gandhi S, Mates M, Freedman OC, Dent SF, et al. Optimal systemic therapy for early breast cancer in women: a clinical practice guideline. Curr Oncol. 2015;22(Suppl 1):S67-81.

(239) Colomer R, Aranda-Lopez I, Albanell J, Garcia-Caballero T, Ciruelos E, Lopez-Garcia MA, et al. Biomarkers in breast cancer: a consensus statement by the Spanish Society of Medical Oncology and the Spanish Society of Pathology. Clin Transl Oncol. 2018;20(7):815-26.
(240) 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. J Clin Oncol. 2019;37(22):1956-64.
(241) Duffy MJ, 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). Eur J Cancer. 2017;75:284-98.
(242) Broder MS, Sing AP. A systematic review of the impact of molecular diagnostics on treatment decisions for patients with breast cancer. J Oncopathol. 2013;1(2):131-43.
(243) Carlson JJ, Roth JA. The impact of the Oncotype Dx breast cancer assay in clinical practice: a systematic review and meta-analysis. Breast Cancer Res Treat. 2013;141(1):13-22.
(244) 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]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2014 [cited 2018 Jul 20]. Available from: https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0073303/pdf/PubMedHealth_PMH0073303.pdf
(245) Canadian Agency for Drugs and Technologies in Health (CADTH). Oncotype DX in women and men with ER-positive, HER2-negative early stage breast cancer who are lymph node negative: a review of clinical effectiveness and guidelines [Internet]. Ottawa (ON): The Agency; 2014 [cited 2018 Jul 20]. Available from: https://cadth.ca/sites/default/files/pdf/htis/apr-2014/RC0524%20Oncotype%20DX%20for%20node%20negative%20patients%20Final.pdf
(246) Augustovski F, Soto N, Caporale J, Gonzalez L, Gibbons L, Ciapponi A. Decision-making impact on adjuvant chemotherapy allocation in early node-negative breast cancer with a 21-gene assay: systematic review and meta-analysis. Breast Cancer Res Treat. 2015;152(3):611-25.
(247) 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 [Internet]. Brussels: Belgian Health Care Knowledge Centre (KCE); 2015 [cited 2018 Jul 25]. Available from: https://kce.fgov.be/sites/default/files/atoms/files/KCE_237_Gene%20expression%20profiling_Report_0.pdf
(248) Marrone M, Stewart A, Dotson WD. Clinical utility of gene-expression profiling in women with early breast cancer: an overview of systematic reviews. Genet Med. 2015;17(7):519-32.
(249) Institute for Quality and Efficiency in Health Care. Biomarker-based tests to support the decision for or against adjuvant systemic chemotherapy in primary breast cancer [Internet]. Cologne (Germany): The Institute; 2016 [cited 2019 May 9]. Available from: https://www.iqwig.de/en/projects-results/projects/non-drug-interventions/d-projekte/d14-01-biomarker-based-tests-for-the-decision-for-or-against-adjuvant-systemic-chemotherapy-in-primary-breast-cancer.6097.html
(250) Issa AM, Chaudhari VS, Marchant GE. The value of multigene predictors of clinical outcome in breast cancer: an analysis of the evidence. Expert Rev Mol Diagn. 2015;15(2):277-86.

(251) Canadian Agency for Drugs and Technologies in Health (CADTH). Gene expression tests for women with early stage breast cancer: a review of clinical utility and cost-effectiveness [Internet]. Ottawa (ON): The Agency; 2017 [cited 2018 Jul 20]. Available from: https://www.cadth.ca/sites/default/files/pdf/htis/2017/RC0934%20Mlc%20Tests%20for%20Breast%20Cancer%20Final.pdf

(252) Harnan S, Tappenden P, Cooper K, Stevens J, Bessey A, Rafia R, et al. Tumour profiling tests to guide adjuvant chemotherapy decisions in people with breast cancer (update of DG10) [Internet]. London (UK): University of Sheffield; 2017 [cited 2019 May 9]. Available from: https://www.nice.org.uk/guidance/dg34/documents/diagnostics-assessment-report-2

(253) Scope A, Essat M, Pandor A, Rafia R, Ward SE, Wyld L, et al. Gene expression profiling and expanded immunohistochemistry tests to guide selection of chemotherapy regimens in breast cancer management: a systematic review. Int J Technol Assess Health Care. 2017;33(1):32-45.

(254) European Network for Health Technology Assessment (EUnetHTA). Added value of using the gene expression signature test MammaPrint for adjuvant chemotherapy decision-making in early breast cancer [Internet]. Diemen (Netherlands): The Network; 2018 [cited 2018 Oct 4]. Available from: https://www.eunethta.eu/wp-content/uploads/2018/01/EUnetHTA_assessment_mammaprint_final.pdf

(255) King V, Mosbaek C, Carson S, Lazur B, Leof A, Liu R, et al. Gene expression profile testing of cancer tissue [Internet]. Olympia (WA): Washington State Health Care Authority; 2018 [cited 2018 Jul 20]. Available from: https://www.hca.wa.gov/assets/program/gene-expression-final-rpt-20180220_0.pdf

(256) Oregon Health Authority. Health Evidence Review Commission (HERC) coverage guidance: genome expression profiling for breast cancer [Internet]. Salem (OR): The Authority; 2018 [cited Oct 4 2018]. Available from: https://www.oregon.gov/oha/HPA/CSI-HERC/EvidenceBasedReports/CG%20-%20Breast%20Cancer%20Gene.pdf

(257) Dubsky P, Brase JC, Jakesz R, Rudas M, Singer CF, Greil R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. Br J Cancer. 2013;109(12):2959-64.

(258) Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol. 2010;11(1):55-65.

(259) 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. N Engl J Med. 2015;373(21):2005-14.

(260) Gnant M, Sestak I, Filipits M, Dowsett M, Balic M, Lopez-Knowles E, et al. Identifying clinically relevant prognostic subgroups of postmenopausal women with node-positive hormone receptor-positive early-stage breast cancer treated with endocrine therapy: a combined analysis of ABCSG-8 and ATAC using the PAM50 risk of recurrence score and intrinsic subtype. Ann Oncol. 2015;26(8):1685-91.

(261) Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst. 2006;98(17):1183-92.

(262) Mittmann N, Stout NK, Lee P, Tosteson ANA, Trentham-Dietz A, Alagoz O, et al. Total cost-effectiveness of mammography screening strategies [Internet]. Ottawa (ON): Statistics Canada; 2015 [cited 2018 Nov 2nd]. Available from: https://www150.statcan.gc.ca/n1/pub/82-003-x/2015012/article/14295-eng.htm

About Us

This health technology assessment was produced by the Quality business unit at Ontario Health, the government agency that when fully established will be responsible for ensuring all Ontarians receive high-quality health care where and when they need it.

For more information about Ontario Health, visit ontariohealth.ca.

About the Ontario Health Technology Advisory Committee

How to Obtain Reports From the Ontario Health Technology Assessment Series

Disclaimer

Ontario Health (Quality)
130 Bloor Street West, 10th Floor
Toronto, Ontario
M5S 1N5
Tel: 416-323-6868
Toll Free: 1-866-623-6868
Fax: 416-323-9261
Email: EvidenceInfo@hqontario.ca
www.hqontario.ca

ISSN 1915-7398 (online)
ISBN 978-1-4868-4215-5 (PDF)

© Queen's Printer for Ontario, 2020

The copyright for all Ontario Health publications is owned by the Queen's Printer for Ontario. Materials may be reproduced for commercial purposes only under a licence from the Queen's Printer. For further information or to request a licence to reproduce content, please contact:

Senior Copyright Advisor
Publications Ontario
416-326-5153
copyright@ontario.ca