KCE REPORT 237

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

GENE EXPRESSION PROFILING AND IMMUNOHISTOCHEMISTRY TESTS FOR PERSONALISED MANAGEMENT OF ADJUVANT CHEMOTHERAPY DECISIONS IN EARLY BREAST CANCER

A RAPID ASSESSMENT

KCE REPORT 237
HEALTH TECHNOLOGY ASSESSMENT

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

GENE EXPRESSION PROFILING AND IMMUNOHISTOCHEMISTRY TESTS FOR PERSONALISED MANAGEMENT OF ADJUVANT CHEMOTHERAPY DECISIONS IN EARLY BREAST CANCER

A RAPID ASSESSMENT

LORENA SAN MIGUEL, JOAN VLAYEN, CHRIS DE LAET

2015 www.kce.fgov.be

COLOPHON

Title:	Gene expression profiling and immunohistochemistry tests for personalised management of adjuvant chemotherapy decisions in early breast cancer – a Rapid Assessment
Authors:	Lorena San Miguel (KCE), Joan Vlayen (KCE), Chris De Laet (KCE)
Project coordinator:	Marijke Eyssen (KCE)
Reviewers:	Frank Hulstaert (KCE), Nancy Thiry (KCE)
External experts:	Bruce Poppe (UZ Gent), Roberto Salgado (Institut Jules Bordet; GZA), Christos Sotiriou (Institut Jules Bordet, ULB), Hans Wildiers (UZ Leuven), Jean-Jacques Cassiman (VLK), Jacques De Grève (UZ Brussel), Giuseppe Floris (UZ Leuven), Yves Guiot (UCL), Liesbeth Lenaerts (WIV – ISP), Patrick Neven (UZ Leuven), Nancy Van Damme (Stichting Kankerregister – Fondation Registre du Cancer), Hilde Engels (RIZIV – INAMI), Anouk Waeytens (RIZIV – INAMI),
External validators:	Report validated by Patrick Pauwels (UZA) and Mirjana Huic (Agency for Quality and Accreditation in Health Care and Social Welfare, Croatia). A third validator: Fatima Cardoso (Champalimaud Foundation, Portugal) had major remarks which are discussed in detail in Appendix 3 of this report.
Acknowledgements:	Stephan Devriese (KCE), Harlinde De Schutter (Stichting Kankerregister)
Other reported interests:	Membership of a stakeholder group on which the results of this report could have an impact.: Bruce Poppe (UZ Gent, UGent, Centrum Medische Genetica Gent), Jacques De Grève (BSMO) Owner of intellectual property rights (patent, product developer, copyrights, trademarks, etc.): Christos Sotiriou (brevet sur grade genomique) Participation in scientific or experimental research as an initiator, principal investigator or researcher: Fatima Cardoso (scientific director *Transbig* and principal investigator of the *Mindact Trial*), Patrick Neven Grants, fees or funds for a member of staff or another form of compensation for the execution of research: Fatima Cardoso (via EORTC for the running of the Mindact Trial), Patrick Neven Payments to speak, training remuneration, subsidised travel or payment for participation at a conference: Patrick Neven Presidency or accountable function within an institution, association, department or other entity on which the results of this report could have an impact: Bruce Poppe (UZ Gent, UGent, Centrum Medische Genetica Gent), Jean Jacques Cassiman (VLK), Patrick Neven, Jacques De Grève (BSMO)
Layout:	Ine Verhulst

Disclaimer:	• **The external experts were consulted about a (preliminary) version of the scientific report. Their comments were discussed during meetings. They did not co-author the scientific report and did not necessarily agree with its content.** • **Subsequently, a (final) version was submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. The validators did not co-author the scientific report and did not necessarily all three agree with its content.** • **Finally, this report has been approved by common assent by the Executive Board.** • **Only the KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of the KCE.**

Publication date:	13 January 2015
Domain:	Health Technology Assessment (HTA)
MeSH:	Breast Neoplasms; Gene Expression Profiling, Immunohistochemistry
NLM Classification:	WP 870
Language:	English
Format:	Adobe® PDF™ (A4)
Legal depot:	D/2015/10.273/13

How to refer to this document?

San Miguel L, Vlayen J, De Laet C. Gene expression profiling and immunohistochemistry tests for personalised management of adjuvant chemotherapy decisions in early breast cancer – a Rapid Assessment. Health Technology Assessment (HTA) Brussels: Belgian Health Care Knowledge Centre (KCE). 2015. KCE Reports 237. D/2015/10.273/13.

This document is available on the website of the Belgian Health Care Knowledge Centre.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
AE	Adverse Event
AJCC	American Joint Committee on Cancer
BCI	Breast Cancer Index
BCFI	Belgian Centre for Pharmacotherapeutical Interventions
CADTH	Canadian Agency for Drugs and Technologies in Health
CBIP	Belgian Centre for Pharmacotherapeutical Interventions
CE	Cost Effectiveness
CPI	Consumer Price Index
CU	Cost Utility
DNA	Deoxyribonucleic Acid
ER	Estrogen Receptor
FFPE	Formalin-fixed Paraffin-Embedded
GEP	Gene Expression Profiling
GGI	Genomic Grade Index
HTA	Health Technology Assessment
HER2	Human Epidermal Growth Factor Receptor 2
ICER	Incremental Cost Effectiveness Ratio
IHC	Immunohistochemistry
INAHTA	International Network of Agencies for Health Technology Assessment
LN	Lymph node
LYG	Life Years Gained
NCCN	National Comprehensive Cancer Network
NICE	National Institute for Health and Care Excellence
NPI	Nottingham Prognosis Index
PR	Progesterone Receptor
QALY	Quality Adjusted Life Year
QoL	Quality of Life
RCT	Randomised Controlled Trial

RS	Recurrence Score
RT-PCR	Real Time Reverse Transcription-Polymerase Chain Reaction
SR	Systematic Review
UICC	Union International Contre le Cancer

SCIENTIFIC REPORT

How to use this document?

This Scientific Report is not intended to be read as a stand-alone document, but as a complement to the Synthesis of this study. It gives a detailed account of the methods and results of each of the scientific building blocks underpinning the messages rendered in the Synthesis.

The context, problem description, as well as the discussion of the results and the conclusions are to be found in the Synthesis.

The Synthesis is published as a separate document on our web site. It can be accessed from the same referral page as the current document.

1 INTRODUCTION AND SCOPE

Gene expression profiling (GEP) and expanded immunohistochemistry (IHC) tests aim to improve decision-making relating to adjuvant chemotherapy for women with early breast cancer. The RIZIV/INAMI was contacted by breast cancer specialists to evaluate the options of reimbursement for these tests. After a pre-assessment of the literature, it became clear that no randomized trials are available yet, and the RIZIV/INAMI asked the KCE to provide an overview of the published reviews on the topic in order to have a supporting document during the discussions with clinicians.

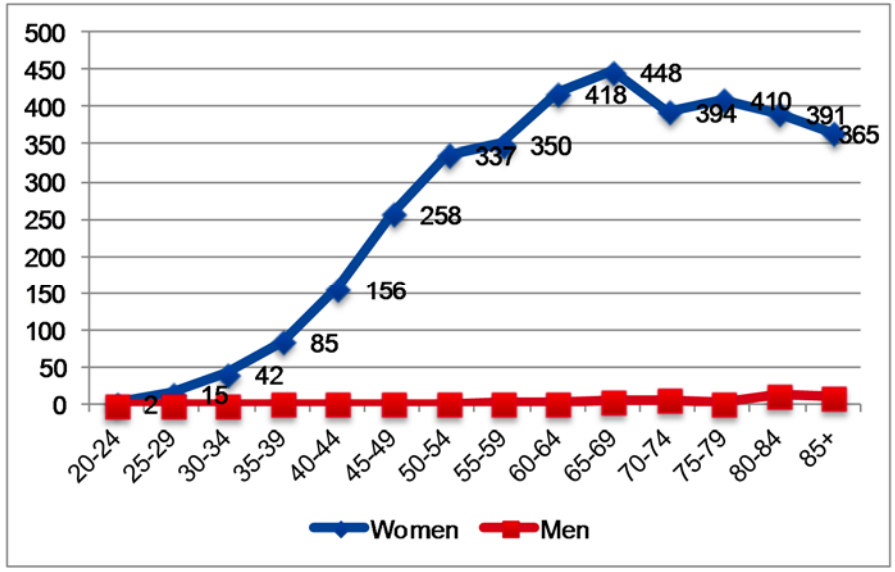
The aim of this report is to assess the clinical effectiveness and cost-effectiveness of selected GEP and IHC tests using the current evidence on analytical validity, clinical validity and clinical utility by doing a review of published systematic reviews and a review of economic analyses. Budgetary impact scenarios are also performed.

2 BREAST CANCER

2.1 Epidemiology

Breast cancer is the most commonly diagnosed cancer in women in Belgium and worldwide. In 2011 the Belgian Cancer Registry reported 10 490 cases of incident breast cancers in women and 75 in men.[1] The European age standardized rate of incident breast cancer in 2011 was 150/100 000 person years in women and 1/100 000 person years in men. Incidence increases markedly with age with a peak in the 65-69 year age category in women as shown in Figure 1. Mean age at diagnosis is 62 year in women.[1]

Figure 1 – Age and gender-specific incidence of breast cancer in Belgium in 2011 per 100 000 person years



Source: Belgian Cancer Registry[1]

2.2 Breast cancer classifications

Breast cancer can be classified using different schemes. Each of these aspects can influence treatment response and prognosis. Description of a breast cancer would optimally include several of these classification aspects, as well as other findings, such as signs found on physical examination. These classifications include histopathological type, grade, stage, receptor status and the presence or absence of specific genes as determined by genetic testing.[2] A detailed description of all these classifications is outside the scope of this report and we limit our description to the most relevant for this report.

TNM Stage Breast cancer classification

The American Joint Committee on Cancer (AJCC) and the Union Internationale Contre le Cancer (UICC) promote and jointly maintain TNM staging which is a two-step procedure.[3] The TNM staging has had different versions and is currently in its 7th edition.

The TNM system first classifies cancer by several factors, T for tumour, N for nodes and M for metastasis. Each of the TNM combinations correspond to a specific stage depending on the type of cancer.

For breast cancer those TNM combinations correspond to one of five stages:[4]

- Stage 0: Carcinoma in situ.
- Stage I to stage III: Higher numbers indicate more extensive disease: larger tumour size and/or spread of the cancer beyond the organ in which it first developed to nearby lymph nodes and/or tissues or organs adjacent to the location of the primary tumour.
- Stage IV: The cancer has spread to distant tissues or organs (metastases).

Receptor status breast cancer classification

The receptor status of breast cancer has traditionally been identified by immunohistochemistry (IHC), which stains the cells based on the presence of specific receptors. The receptors most commonly tested are the: estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2). Cancers were all three receptors are absent are often called triple negative.[5]

IHC remains the most common method of testing for receptor status, but DNA multi-gene expression profiles can categorize breast cancers into molecular subtypes that generally correspond to IHC receptor status.[5]

Receptor status is a critical assessment for all breast cancers as it determines the suitability of using personalised adjuvant treatments such as tamoxifen or trastuzumab.[2]

2.3 Prognosis

Breast cancer has overall a relatively good prognosis in Belgium with a five-year relative survival rate of 88.0% in women and 78.2 % in men and a ten-year relative survival (Flemish region only) of 78.9% and 61.9% respectively.[6] However, mortality is still considerable and this cancer accounts for approximately 20% of all cancer deaths in Belgian women. In 2009, 2 360 women (19 men) died from breast cancer in Belgium.[6, 7]

Relative survival is dependent of the stage of the tumour at the time of diagnosis. Five-year breast cancer survival in Belgium is 100% for TNM Clinical Stage 0, 99.4% for stage I and down to 28.0 % for Stage IV.[6]

More sophisticated tools based on several breast cancer classifications and patient characteristics have been developed, some of them computerized, such as the Nottingham Prognostic Index (NPI)[8, 9] and Adjuvant!Online.[10]

2.4 Guidelines for Breast Cancer Management

Treatment usually involves primary surgery (sometimes preceded by neoadjuvant therapy to downsize the tumour) to remove the primary tumour and any involved lymph nodes. It might be followed in a later phase by adjuvant therapy such as radiation therapy, endocrine therapy and/or chemotherapy with or without targeted biological therapy, all depending upon characteristics of tumour and patient.

Several scientific societies developed clinical practice guidelines for diagnosis and disease management. For Belgium we refer to the guidelines published by KCE developed in collaboration with the Belgian college of Oncology and last updated in 2013.[11]

The guidelines recommend, in addition to the basic triple assessment (clinical assessment, imaging and tissue sampling) the assessment of the hormonal receptors (ER, PR) and HER2 in primary invasive breast cancers. Neoadjuvant and adjuvant therapies after surgery should be decided depending on tumour and patient characteristics in a multidisciplinary team and taking the preferences of the patient into account.

3 TESTS TO GUIDE PERSONALISED MANAGEMENT OF BREAST CANCER

3.1 Personalised medicine in breast cancer

Personalised medicine is an often used but vague concept. Recently, Schleidgen et al. conducted a search of the use of this concept in scientific literature and proposes the following definition:[12] "*Personalised medicine seeks to improve stratification and timing of health care by utilizing biological information and biomarkers on the level of molecular disease pathways, genetics, proteomics as well as metabolomics.*"

Gene expression profiling (GEP) and immunohistochemistry tests (IHC) for breast cancer stratification are examples of personalised medicine.

The aim of these tests is to improve the targeting of adjuvant therapy by providing more accurate prognostic information for subgroups of patients.[5] The reasoning is that better knowledge of specific biological features of cancers may indicate an increased likelihood of rapid growth, metastasis risk and response to therapy may lead to a better risk stratification and to the selection of patients that are most likely to benefit from chemotherapy.[5, 13]

These tests are typically indicated for women with ER+ and LN– tumours together with information about grade, size etc. They are intended to give additional prognostic information and guide the decision on whether or not to offer chemotherapy. Some tests require samples to be sent away for central review following surgery which make cause a delay of up to a few weeks.[5]

3.2 Gene expression profiling (GEP)

Gene expression profiling tests assess the identity and number of messenger ribonucleic acid (mRNA) transcripts in a specific tissue sample.

Gene expression profiling gives information about the activity of genes that give rise to these mRNA transcripts. The mRNA levels are associated with the protein composition of the cells, and consequently to changes in the properties and functions of tissues and cells (both normal and malignant) in the body.

Various assays are used in the management of breast cancer. These assays investigate the expression of specific panels of genes, including real-time reverse transcription-polymerase chain reaction (RT-PCR) and deoxyribonucleic acid (DNA) microarrays. Many of these assays have been designed to measure the risk of cancer recurrence. Other uses of the assays include breast cancer subtyping using molecular classification systems,[14] predicting the likely benefit from certain types of therapy. A recent HTA assessed six of these GEP tests.[5]

Central testing is currently required for most of these tests, including OncotypeDX (USA), Mammaprint, H/I, and BluePrint (USA). Others can be performed at local laboratories: e.g. MapQuant DX and Endopredict. Both PAM50 and the Randox Breast Cancer array, could potentially be performed locally, but only if the necessary machine units are purchased.

3.3 Expanded Immunohistochemistry tests (IHC)

Immunohistochemistry tests measure protein synthesis levels in the tumour sample rather than mRNA or cDNA. IHC identifies the number of cancer cells synthesising specific proteins. The intensity of staining correlates with protein synthesis levels. Some of these tests offer the advantage of using existing immunohistochemical markers (ER, PR and HER2), which are routinely tested in early invasive breast cancer. The term 'expanded' is used to describe the fact that these tests are used in addition to the standard IHC receptor testing. However, differences in IHC values can occur caused by various reasons. The same recent HTA evaluated three of those IHC tests.[5] Some of these tests including IHC4 could be performed locally, while others like Mammostrat require central testing.

3.4 Tests addressed in this overview

Many tests exist and for practical reasons this review will be limited to specific tests only (Table 1). Test characteristics will be briefly described below.

Table 1– Overview of tests addressed in this study

	Method	N genes or markers
Oncotype DX	RT-PCR	21 genes
PAM50	RT-PCR	50 genes
MapQuant DX	RT-PCR	9 genes
H/I*	RT-PCR	3 genes
EndoPredict	RT-PCR	12 genes
Mammaprint	Microarray-based	70 genes
BluePrint	Microarray-based	80 genes
Randox Breast Cancer Array	Bio-chip technology	23 genes
Mammostrat	Immunohistochemistry	5 markers
NPI+	Immunohistochemistry	10 markers
IHC4	Immunohistochemistry	4 markers
uPA/PAI-1	ELISA	2 enzymes

**Test replaced by BCI (includes H/I and 5 prolifereation genes)*

3.4.1 GEP with real-time reverse transcription-polymerase chain reaction (RT-PCR)

3.4.1.1 Oncotype DX

OncotypeDX™ (Genomic Health Inc., Redwood City, CA, USA – http://www.oncotypedx.com) quantifies gene expression for 21 genes in breast cancer tissue using RT-PCR. It predicts the likelihood of recurrence in women of all ages with newly diagnosed stage I or II, ER+, LN– or LN+ (up to three nodes) breast cancer treated with tamoxifen. The test assigns the breast cancer a recurrence score (RS) and a risk category: low (RS < 18), intermediate (18 ≤ RS ≤ 30) or high (RS ≥ 31). The test also reports ER, PR and HER2 status.[5]

To quantify gene expression, RNA is extracted from formalin-fixed, paraffin-embedded (FFPE) tumour tissue and subjected to DNase I treatment. Total RNA content is measured and the absence of DNA contamination is verified. Reverse transcription is performed and is followed by quantitative TaqMan® (Roche Molecular Systems, Inc.) RT-PCR reactions in 384-well plates. The expression of each of 16 genes is measured in triplicate and then normalised relative to a set of five reference genes.[15]

The Oncotype DX test costs approximately US$3 400.[16]

3.4.1.2 PAM50

The PAM50 gene expression assay (ARUP Laboratories, Salt Lake City, UT, USA) identifies the major intrinsic biological subtypes of breast cancer. The current version of the test provides classification of breast cancer subtype and quantitative values for (gene/protein) ESR1/ER, PGR/PR, ERBB2/HER2, proliferation score and luminal score (ER pathway). The PAM50 Breast Cancer Intrinsic Classifier test is recommended for all patients diagnosed with invasive breast cancer, regardless of stage or ER status. It is FDA cleared. The cost of the test is of approximately US$3200.[5]

3.4.1.3 MapQuant DX

MapQuant DX™ (Ipsogen) quantifies gene expression for 8 genes in breast cancer tissue using RT-PCR. The output is the Genomic Grade Index, which divides histologically defined G2 tumours into low- or hig-grade. This test will be commercially available in 2015.

3.4.1.4 H/I

The Breast Cancer Index (BCI)℠ (bioTheranostics Inc., San Diego, CA, USA) is a RT-PCR assessment of the ratio of expression of two genes, HOXB13 and IL17BR, combined with the five gene Molecular Grade Index (MGI) and gives an indication of recurrence risk. The target population is those with ER+ and LN– early breast cancer. The BCI RS ranges from 0 to 10 and divides patients into three risk groups: low risk is defined as a score < 5, intermediate risk is a score of 5–6.3 and high risk is a score ≥ 6.4. Its price is of approximately US$3 200.[5]

3.4.1.5 EndoPredict

EndoPredict® (Sividon Diagnostics GmbH, Germany) is a multianalyte gene expression test to predict distant metastasis in ER-positive, HER2-negative breast cancer treated with endocrine therapy alone. The test is based on the combined analysis of 12 genes in formalin-fixed, paraffin-embedded tissue by RT-PCR. Relative gene expression levels are used to calculate the EndoPredict score ranging from 0 to 15. Patients with a score below or equal to 5 are classified as low risk for distant recurrence under endocrine therapy, those with a score above 5 as high risk. Its price is of approximately €1811.[17]Microarray- or biochip-based GEP

3.4.1.6 MammaPrint

MammaPrint® (Agendia, the Netherlands – http://www.agendia.com) is based on microarray technology and uses a 70-gene expression profile. MammaPrint is intended as a prognostic test for women of all ages, LN– and LN+ (up to three nodes positive), with a tumour size of ≤ 5.0 cm. MammaPrint is used to determine the risk of distant recurrence of early breast cancer.[5]

Mammaprint is performed on either formalin-fixed, paraffin-embedded (FFPE) tumour tissue or fresh tissue. Analysis is done in one of two central laboratories (the Netherlands and the USA). It is up to date one of the only two gene expression profiling test having FDA clearance.

Cost of the test is about € 2 675.[16]

3.4.1.7 BluePrint

BluePrint™ (Agendia, Amsterdam, the Netherlands) is used in addition to the MammaPrint test for molecular subtyping. It is an 80-gene microarray with a target population of patients with early-stage (stage I or II), LN– or LN+ (up to three nodes positive), ER+ or ER– breast cancer. BluePrint provides information on breast cancer subtype using three categories: basal-type, luminal-type and ERBB2-type cancers. It is offered at no additional cost over that of MammaPrint®[5]

3.4.1.8 Randox Breast Cancer Array

The Randox Breast Cancer Array (BCA) (Randox Laboratories, Crumlin, UK) is a complementary DNA (cDNA)-based expression biochip assay that aims to accurately define the clinical subtypes of breast cancer tumours before initiating treatment. The target population is all individuals with diagnosed breast cancer. No cost information was identified.

3.4.2 Immunohistochemistry

3.4.2.1 Mammostrat

The Mammostrat® test (Clarient Inc, US) uses five IHC markers (SLC7A5, HTF9C, p53, NDRG1 and CEACAM5) to stratify patients into risk groups to inform treatment decisions. These markers are independent of one another and do not directly measure either proliferation or hormone receptor status. The current version of the test provides classification into one of five breast cancer subtypes, and quantitative values for (gene/protein) ESR1/ER, PGR/PR, ERBB2/HER2, proliferation, and luminal score (ER pathway), along with a RS and category (low, moderate and high). Its cost is estimated to be between GBP1120 and GBP1620.

3.4.2.2 Nottingham Prognostic Index plus

NPI+ is a biomarker-based prognostic assay that integrates 10 predictive biomarkers of long-term survival and therapeutic response with existing clinical and molecular pathology knowledge to support individualised clinical decision-making. The assay is not commercialized yet.

3.4.2.3 IHC4

IHC4 assesses the levels of four key proteins (ER, PR, HER2 and Ki-67) in a breast cancer sample. The IHC4 score is calculated based on the percentage of cells positive for Ki-67 and PR (0–100%); the Histoscore for ER status (a measure of the percentage of cells positive multiplied by the intensity, range 0–300); and the tumour HER2 status, expressed as a binary measure (positive/negative). The final algorithm for IHC4 calculates a risk score for distant recurrence based on ER, PR, HER2 and Ki-67 in addition to classical clinical and pathological variables (composite risk score IHC4 + clinical). No risk category is given. IHC4 costs have been estimated to be approximately GBP150.[5]

3.4.3 Other

3.4.3.1 uPA/PAI-1 (Femtelle)

Urokinase plasminogen activator / plasminogen activator inhibitor 1 ELISA (uPA/PAI-1) is a registered enzyme-linked immunoassay (ELISA) kit (FEMTELLE) for the analysis of uPA/PAI-1 in fresh frozen tissue and is being provided by American Diagnostica Inc. (http://www.femtelle.de/en/). It is CE marked in Europe but for research use only in the USA. Other commercial ELISA kits for separate in-house analysis of uPA and/or PAI-1 are available from different suppliers. These also use samples other than tissue and are also used for indications other than cancer.[16]

Costs for FEMTELLE including preparation, shipping and analysis of samples in a qualified laboratory amount to €400. In house analysis with separate ELISA kits costs about €200.[16]

4 EFFICACY, EFFECTIVENESS AND SAFETY

4.1 Methodology

4.1.1 PICO question

4.1.1.1 Population

Women diagnosed with early invasive breast cancer. Some tests may only be used in a sub-population. For example, women with early-stage invasive breast cancer (stage I, II or III), lymph node negative or positive (up to 3), oestrogen receptor positive or negative and HER2 positive or negative.

4.1.1.2 Interventions

The assessment will include the following selected gene expression profiling tests and expanded immunohistochemistry tests:

- Oncotype DX
- PAM50
- MapQuant DX
- H/I (replaced by BCI)
- EndoPredict
- Mammaprint
- BluePrint
- Randox Breast Cancer Array
- Mammostrat
- NPI+
- IHC4
- uPA/PAI-1

4.1.1.3 Comparators

The comparator will be current clinical practice based on Belgian and international guidelines. This might include the use of other prognostic tools based on combinations with several pathological parameters such as tumour size, grade and lymph node status to predict survival, recurrence and the utility of adjuvant therapy in breast cancer.

4.1.1.4 Outcomes

- Analytic validity (i.e. the technical efficacy of the test),
- Clinical validity (i.e. the prognostic ability of the test),
- Clinical utility in relation to harm, impact on clinical decision making, evidence of improvement in outcomes (e.g. overall survival, quality of life) and health care costs.

4.1.2 Literature search

Efficacy and effectiveness were assessed based on a review of existing systematic reviews (SR). Reviews were identified through a systematic literature search in Medline and PreMedline (OVID), Embase, and the Cochrane library (CDSR, CENTRAL, DARE, HTA database). More details of these searches can be found in the appendix. Searches were run on June 27th 2014.

Reviews published in languages other than English, French or Dutch were excluded.

The search yielded 1 794 unique references. From these 1 679 were excluded based on title and abstract. The full-text of 115 papers was evaluated, and 13 reviews were finally included.

4.1.3 Quality appraisal and data extraction

For the quality appraisal of systematic reviews, the AMSTAR checklist was used (http://amstar.ca). Appraisal was done by one reviewer. Detailed results are reported in the appendix. Data extraction was done by the same reviewer. The evidence tables can also be found in appendix.

4.2 Results by test

4.2.1 GEP with real-time reverse transcription-polymerase chain reaction (RT-PCR)

4.2.1.1 Oncotype DX

In 2008, Marchionni et al. published a SR of the impact of three GEP tests on breast cancer outcomes. These tests included OncotypeDX, MammaPrint and H/I.[18] This SR was updated in 2010 by Smartt et al.[19] and MAS[20], and in 2013 by EUnetHTA[16] and Ward et al.[5] The latter identified 32 studies for Oncotype DX (Table 2). A systematic review and meta-analysis of fairly good quality by Carlson et al. [21] identified some additional more recent studies, as did the EUnetHTA review. Other systematic reviews of lower quality identified almost no additional studies.[22-26]

Most of the identified primary studies used a retrospective analysis of archived tumour samples together with a database of patient characteristics and prognostic information. Only three studies identified by Ward et al. stated that the design was prospective.[5] The majority of participants analysed in the studies were ER+, LN–, and the mean age was around 50–60 years. Most studies included a small number of participants (range 25–367) and follow-up was short or not reported for some. However, three studies analysed relatively large cohorts (at least thousand patients) with follow-up of up to 10 years.[27-29]. The methodological quality of the studies was assessed as low for most studies and only three received from the authors a positive assessment of at least 17 out of 21 methodological quality items.

Ward et al. concluded from this update for Oncotype DX that:[5]

- Analytical validity: no evidence in addition to the earlier reviews of Marchionni et al.[18] and Smartt et al.[19] There is limited evidence for the reproducibility of the tests across different samples of the same block and across samples from different blocks of the same tumour. Centralisation was considered to be a current strength of OncotypeDX with regard to reproducibility.

- Clinical validity (prognostic ability of the test): significant correlation of the Oncotype DX recurrence score with disease-free survival and overall survival. The recurrence score is shown to be a better predictor of distant recurrence at 10 years than traditional clinicopathological predictors. Key gaps relate to the stability of risk categories in populations other then ER+, LN– patients.
- Clinical utility: in earlier reviews the evidence on clinical utility is limited. In this review four studies present further evidence on the impact of Oncotype DX on clinical decision-making and report changes in breast cancer management post-Oncotype DX for between 31.5 and 38% of patients. However, these studies were judged by the authors as having a limited value due to study design an small sample size. Further supporting evidence is needed and key gaps relate to the extent to which the test added to the management of patients and the proportion of patients that would benefit from the test.

Conclusions (adapted from Ward et al.)

- **The evidence for Oncotype DX is more robust than the evidence for other tests. However, important evidence gaps are still present. This review has mainly identified studies supporting the prognostic ability (clinical validity) of the test. These studies are judged to be of moderate to high quality.**
- **No prospective studies reported on the impact of Oncotype DX on long-term outcomes such as overall survival, while four studies indicated that Oncotype DX leads to changes in decision making.**
- **Two studies on the predictive benefit of the test were identified, one for LN+ patients. The first evidence relating to improvements in quality of life and reductions in patient anxiety as a result of using the test has been reported, but this is based on small patient numbers and further evidence is required.**

Table 2 – Studies on Oncotype DX included in various systematic reviews

	Marchionni 2008	Smartt 2010	MAS 2010	Ward 2013	EUnetHTA 2013	Carlson 2013
Chang 2007	x	x	x			
Cobleigh 2005	x	x	x	x		
Cronin 2004	x	x	x	x		
Cronin 2007	x	x	x	x		
Esteva 2005	x	x	x	x		
Fan 2006	x	x	x			
Gianni 2005	x	x	x			
Habel 2006	x	x	x	x		
Hornberger 2005	x	x	x	x		
Lyman 2007	x	x	x			
Mina 2006	x	x	x			
Oratz 2007	x	x	x	x	x	x
Paik 2004	x	x	x	x		
Paik 2006	x	x	x	x		
Akashi-Tanaka 2009		x	x			
Asad 2008		x	x	x	x	x
Goldstein 2008		x	x	x		
Henry 2009		x		x	x	x
Kok 2009		x	x			
Li 2009		x		x		
Rayhanabad 2008		x		x	x	

	Marchionni 2008	Smartt 2010	MAS 2010	Ward 2013	EUnetHTA 2013	Carlson 2013
Wolf 2008		x	x	x		x
Lo 2007		x				
Shak 2009		x		x		
Erb 2007		x		x		
Gold 2009		x		x		
Ademuyiwa 2011				x	x	x
Albain 2010			x	x		
Bryant 2005				x		
Cuzick 2011				x		
Dowsett 2010			x	x		
Espinosa 2009			x			
Geffen 2009			x	x		
Holt 2011				x		x
Kelly 2010				x		x
Lo 2010			x	x	x	x
Mamounas 2010			x	x		
Tang 2010				x		
Tang 2011				x		
Toi 2010			x	x		
Yorozuya 2010				x		
de Boer 2011						x
Geffen 2011					x	x
Goodwin 2009						x

	Marchionni 2008	Smartt 2010	MAS 2010	Ward 2013	EUnetHTA 2013	Carlson 2013
Gregg 2009						x
Guth 2011						x
Hornberger 2011						x
Joh 2011					x	x
Kamal 2011					x	
Klang 2010						x
Lund 2012						x
Moinuddin 2009						x
Oratz 2011					x	
Partin 2011					x	x
Patel 2007						x
Rezai 2011						x
Richman 2011					x	
Tatarian 2011						x
Tzeng 2010					x	x

4.2.1.2 PAM50

Ward et al. identified six studies for the PAM50 test.[5] All of the studies had a retrospective design analysing archived tumour samples. In most studies more patients were LN+ and ER+ than LN– and ER–. The ages of the patients varied from a median age of 47.5 years in one paper to a median of 67 years in another. Most studies included a moderate number of tumour samples. Follow-up was around a median of 10 years in those that did report follow-up time. Overall, the risk of bias from these studies was judged to be moderate.

Ward et al. concluded that:[5]

- Analytical validity: the evidence is limited to two abstracts rated as low quality, providing a comparison of PAM50 against standard IHC measurements.
- Clinical validity (prognostic ability of the tests): four studies, two rated as high quality and two (abstracts) rated as low quality, were identified. These studies demonstrated that the intrinsic subtypes are (1) significantly associated with outcome, provided additional information to IHC approaches is given and standard clinicopathological measures are used, and (2) can identify a particularly low-risk group. They demonstrated that prognostic ability has been validated in external

cohorts. However, the population in most of the studies was LN+, with the exception of one study that assessed LN– patients. Therefore, the generalisability of these findings to LN–, ER+ patients is limited.

- Clinical utility: no studies were identified.

Arpino et al. identified three additional studies.[23] No assessment of the quality was included. One study included premenopausal women, a second study included postmenopausal ER+ women, while the third study included ER+ women with stage I-II breast cancer. In the first two studies, PAM50 results were predictive of tamoxifen and neoadjuvant exemestane benefit, respectively. In the third study, good agreement was found between PAM50 and Oncotype DX for prognostic risk assignment. No studies were found that evaluated the impact on treatment decisions.

Conclusions

- **The evidence for PAM50 is limited to studies supporting the prognostic ability (clinical validity) of the test. Most of the evidence is in node-positive patients.**
- **No studies reported on the impact of PAM50 on clinical management (clinical utility).**

4.2.1.3 MapQuant DX

Arpino et al. is the only systematic review that reported on MapQuant DX™.[23] No data were reported on the analytical validity. Arpino et al. reported on the clinical validity of the Genomic Grade Index (GGI), which is based on 97 genes associated with tumor differentiation and grade. MapQuant DX™ is a GGI signature based on nine genes. The GGI has been validated in at least three cohorts (including one cohort of 650 ER-positive patients who were untreated or only treated with tamoxifen),[23] and was found to be more closely associated with relapse-free survival than histological grade. Furthermore, the GGI appears to reclassify patients with histologic grade 2 tumors into two groups with high versus low risk of recurrence (HR 3.61, 95% CI 2.25-5.78; p< 0.001). One study reported that a high GGI is associated with increased sensitivity to neoadjuvant paclitaxel plus fluorouracil, adriamycin, and cyclophosphamide chemotherapy in both ER-negative and ER-positive patients, although it remains a predictor of worse survival in ER-positive patients only.[23] No data were reported on the clinical utility of the test.

Conclusions

- **The evidence for MapQuant DX is limited to studies supporting the prognostic ability (clinical validity) of the test.**
- **No studies reported on the impact of MapQuant DX on clinical management (clinical utility).**

4.2.1.4 H/I

Marchionni et al. and Smartt et al. included seven studies on the H/I ratio assay, all evaluating the clinical validity of the assay.[18, 19] Arpino et al. identified a very similar evidence base.[23] All these reviews were updated by Zhao et al. who included 11 studies (with 2958 participants) on the clinical validity of the assay in a meta-analysis.[30] Pooled results revealed that women with higher HOXB13-to-IL17BR expression ratio had significantly worse recurrence-free survival (HR=1.47; 95%CI 1.17-1.84; p=0.001) and non-significantly worse overall survival (HR=1.32; 95%CI 0.97-1.80; p=0.08). In the node-negative subgroup, meta-analysis of the resulting patient stratification demonstrated significantly different recurrence-free survival (HR=1.66; 95%CI 1.38-2.01; p<0.00001) and overall survival (HR=1.93; 95%CI 1.15-3.23; p=0.01). No data were reported on the analytical validity or clinical utility of the test.

Conclusions

- **The evidence for the H/I ratio assay is limited to studies supporting the prognostic ability (clinical validity) of the test.**
- **No studies reported on the impact of the H/I ratio assay on clinical management (clinical utility).**

4.2.1.5 EndoPredict

No systematic reviews reporting on the EndoPredict test were identified via our review.

4.2.2 Microarray- or biochip-based GEP

4.2.2.1 MammaPrint

Marchionni et al. and Smartt et al. included 13 studies on MammaPrint.[18, 19] In addition, Ward et al. included seven more recent studies (Table 3).[5] Other systematic reviews identified almost no additional studies.[16, 23, 25, 31]

Ward et al. concluded from this update for MammaPrint that:[5]

- Analytical validity: no evidence in addition to the earlier reviews of Marchionni et al.[18] and Smartt et al.[19] There were limited data on variability and reproducibility, with a limited number of patients and a moderate number of replications.
- Clinical validity (prognostic ability of the tests): the systematic reviews of Marchionni et al.[18] and Smartt et al.[19] identified a range of studies providing evidence on the prognostic ability of the test in heterogeneous populations. The evidence relating to the clinical validity of MammaPrint was not always conclusive nor supportive of the prognostic value of the test. Four studies suggested that the test could predict prognosis, one study failed to verify the prognostic utility of the test and in another the methods and results were at variance with those of other studies. The review of Ward et al. identified four additional studies that contain data on clinical validity. Of these, two were rated as high quality and two as moderate quality. These studies demonstrated that the MammaPrint score is a strong independent prognostic factor and may provide additional value to standard clinicopathological measures. The majority of the evidence suggests that the test is reliable at predicting outcome at 5 years. However, the population in all of these studies was relatively small (range 102-272). Follow-up was limited to only 5 years in two of the studies.
- Clinical utility: very limited evidence (one study) on clinical utility was identified by the reviews of Marchionni et al.[18] and Smartt et al.[19] Ward et al.[5] identified six additional studies. Of these, two were rated as high quality and four as moderate quality. Five of the six studies reported on how the MammaPrint test reclassifies patients into high- and low-risk groups compared with the risk assigned in current practice. These studies reported that there was a high level of discordance between MammaPrint and current practice, although the studies did not demonstrate how this would impact on treatment decisions. One study reported that the use of MammaPrint would result in altered treatment advice for 40% of patients, but this was based on the assumption that all patients classified as high risk would receive chemotherapy and no patients classified as low risk would receive chemotherapy rather than by providing evidence of actual changes in practice. All of the studies on clinical utility were based on small sample sizes.

Conclusions (adapted from Ward et al.)

- **The evidence base, in particular in relation to the prognostic ability of the test, is developing but is based on small sample sizes (≤ 272).**
- **The test appears to be prognostic at 5 years although the validity of the test to predict longer-term outcomes does not seem to have been established.**
- **It is not yet clear to what extent the use of the MammaPrint test will change the management of patients and to what extent chemotherapy would be offered to patients classified as having a good or a poor prognosis with MammaPrint.**
- **It is unclear to what extent MammaPrint risk groups are predictive of chemotherapy benefit or how the use of MammaPrint will improve patient outcomes through increases in disease-free and overall survival.**
- **The evidence for MammaPrint to date is mainly derived from premenopausal women, but younger women are more likely to be classified as having a poor prognosis using MammaPrint, which might overestimate the benefit of the test.**

Table 3 – Studies on MammaPrint included in various systematic reviews

	Arpino 2013	Lyman 2006	Marchionni 2008	Smartt 2010	Ward 2013	EUnetHTA 2013	MoH 2008	Malaysia
Ach 2007					x			
Bender 2009				x	x			
Bueno-de-Mesquita 2009	x			x	x	x		
Buyse 2006	x	x	x	x	x		x	
de Snoo 2009				x	x			
Espinosa 2005		x						
Gevensleben 2010					x	x		
Glas 2006			x	x	x			
Glas 2008				x	x			
Ishitobi 2010					x			
Knauer 2009				x	x			
Knauer 2010								
Kok 2010					x			
Kok 2012								
Kunz 2010								
Kunz 2011					x			
Mook 2009	x			x	x			
Mook 2010	x				x			
Na 2011					x			
Saghatchian 2009				x	x			
van de Vijver 2002	x	x	x	x	x		x	
van 't Veer 2002		x	x	x	x			
Wittner 2008	x			x	x			
Zarca 2009				x				

4.2.2.2 BluePrint

Ward et al. identified only one meeting abstract of a retrospective study (N=469) that related to the clinical validity of BluePrint[5]. The study was reported to have a high risk of bias. The authors showed that the developed multigene profile can classify breast tumours into luminal-, ERBB2- and basal-like subgroups.

No studies were found on the analytical validity or clinical utility by Ward et al.

None of the other included systematic reviews reported on BluePrint.

Conclusions

- **The evidence for BluePrint is limited to one meeting abstract supporting the prognostic ability (clinical validity) of the test.**
- **No studies reported on the impact of BluePrint on clinical management (clinical utility).**

4.2.2.3 Randox Breast Cancer Array

Ward et al. did not identify any relevant full peer-reviewed papers or meeting abstracts relating to the Randox assay.[5] However, supplementary evidence was provided by the manufacturer of the test, reporting on a patient cohort of 78 individuals. An overall agreement of 79% between hormonal status (primarily ER) and the multiplex biochip assay was reported. However, because of the lack of methodological detail and the small sample size, Ward et al. were unable to draw conclusions. No data were reported on the analytical validity or clinical utility of the test.

None of the other included systematic reviews reported on the Randox assay.

Conclusions

- **The evidence for the Randox assay is limited to one unpublished study supporting the prognostic ability (clinical validity) of the test.**
- **No studies reported on the impact of the Randox assay on clinical management (clinical utility).**

4.2.3 Immunohistochemistry

4.2.3.1 Mammostrat

Ward et al. identified three studies for the Mammostrat test, all containing data relating to clinical validity.[5] One of these studies also reported on clinical utility in terms of the predictive ability of the test by risk group. Arpino et al. did not find any additional studies.[23]

All studies were retrospective, but overall, the risk of bias was judged to be low. All three studies included more than 1000 patients.

Ward et al. concluded that:[5]

- Analytical validity: no studies were identified.
- Clinical validity (prognostic ability of the tests): the three studies suggested that Mammostrat can act as an independent prognostic tool for ER+, tamoxifen-treated breast cancer. The test has been validated in an external cohort. Although the evidence base for Mammostrat is relatively immature, these initial studies include a large sample size and appear to be of reasonable quality.
- Clinical utility: initial evidence suggested that low- and high-risk groups benefited from chemotherapy, with high-risk patients benefitting more than low-risk patients. The moderate-risk group did not appear to benefit. There was no published evidence on reclassification of risk groups compared with conventional risk classifiers, and no evidence on the impact of the test on decision-making.

Conclusions

- **The evidence for Mammostrat is mainly limited to studies supporting the prognostic ability (clinical validity) of the test. These studies include a large sample size and appear to be of reasonable quality.**
- **One study reported on clinical utility in terms of the predictive ability of the test by risk group. However, further evidence is required.**

4.2.3.2 Nottingham Prognostic Index plus

Ward et al. did not identify any relevant full peer-reviewed papers or meeting abstracts relating to the Nottingham Prognostic Index Plus test.[5] However, supplementary evidence (two draft full papers) was provided by the manufacturer of the test. Ward et al. concluded that the evidence base for NPI+ is currently insufficient to draw any firm conclusions regarding the analytic and clinical validity of the test, and as yet there is no available evidence on the clinical utility of the test.

Conclusions

- **The evidence for the NPI+ test is limited to two unpublished studies supporting the prognostic ability (clinical validity) of the test.**
- **No studies reported on the impact of the NPI+ test on clinical management (clinical utility).**

4.2.3.3 IHC4

Ward et al.[5] and Arpino et al.[23] identified the same study relating to the clinical validity of IHC4. This study was judged to be of high quality. The authors reported that the IHC4 score is a highly significant predictor of distant recurrence. This initial study included a large sample size and detailed the development of the test in one cohort of mainly LN- and HR+ patients (N=1125) and the external validation of the test in an independent cohort of ER+ women (N=786).

No studies were found on the analytical validity or clinical utility by Ward et al.

Conclusions

- **The evidence for IHC4 is limited to one large study supporting the prognostic ability (clinical validity) of the test.**
- **No studies reported on the impact of IHC4 on clinical management (clinical utility).**

4.2.4 Other

4.2.4.1 uPA/PAI-1 (Femtelle)

EUnetHTA assessed the uPA/PAI-1 test within the framework of a Core HTA Model for diagnostic technologies.[16] Due to the different assessment element questions they were not able to assess the analytical and clinical validity, but only the clinical utility. No RCTs or prospective cohort studies on uPA/PAI-1 were identified that assessed whether using this prognostic test (compared with standard/current practice or direct [head-to-head] comparison) to guide the use of adjuvant chemotherapy effectively improves long-term clinical outcomes such as overall survival and disease-specific survival. No studies were identified that addressed the question relating to the treatment choice with adjuvant therapy for uPA/PAI-1 tests. No studies were identified that assessed the effect on quality of life.[16]

Conclusion

- **No studies reported on the impact of uPA/PAI-1 on clinical management (clinical utility).**

5 SYSTEMATIC LITERATURE REVIEW OF ECONOMIC STUDIES

5.1 Introduction

This chapter provides an overview of published studies evaluating the use of Genetic profiling (GEP) and expanded inmunohistochemostry (IHC) tests in chemotherapy treatment decisions for early breast cancer patients from an economic perspective. The aim is to review the literature on the potential cost-effectiveness of these tests as an alternative to other risk stratification methods currently used for deciding whether to expose early breast cancer patients to chemotherapy.

5.2 Methods

5.2.1 Search strategy

A systematic search for relevant publications was carried out with the consultation of electronic reference databases up to 15/09/2014.

Medline (through OVID), EMBASE, Econlit (through OVID), NHSEED (CRD) and NHSHTA (CRD) were searched to retrieve primary full economic evaluations (studies comparing both costs and outcomes) and reviews of economic evaluations (i.e. secondary economic evaluations). An overview of the search strategy is given as an Appendix.

Furthermore, the websites of Health Technology Assessment (HTA) institutes listed on the INAHTA website (International Network of Agencies for Health Technology Assessment) and NICE (National Institute for Health and Care Excellence) were consulted to capture reports on the use of GEP or expanded IH tests in early breast cancer patients. No restrictions were imposed for language or time period.

5.2.2 Selection procedure

To identify potentially relevant studies for our analysis we first went through all titles and abstracts in order to exclude any obvious studies that did not match our research subject. All articles that appeared to be interesting, or for which there were some doubts, were read in full in order to select those relevant for inclusion in our review.

Reference lists of the selected primary and secondary economic evaluations found via our search were checked for additional references worth adding to our analysis.

Study selection was completed by one researcher but any doubts that came up during the exercise were discussed and solved in collaboration with a second reviewer.

All studies finally included in our review were critically appraised by using an in-house structured data extraction sheet based on the check list originally developed by Drummond et al.[32]. See appendix for a copy of the template used.

5.2.3 Selection criteria

All full economic evaluations looking at GEP or expanded IHC tests as prognostic tools for identifying patients most likely to benefit from chemotherapy treatment in early breast cancer were included in our review.

Cost descriptive analyses or cost comparisons not taking into consideration effectiveness were discarded. Similarly, publications in the form of letters, editorials or notes and abstracts were excluded, since these would not offer enough information to include them in our analysis and critically appraise their findings. An overview of the inclusion/exclusion criteria is given in Table 4.

Table 4 – Selection criteria for economic evaluations

Selection criteria	Inclusion criteria	Exclusion criteria
Population	Early invasive breast cancer patients	Regional and distant spread cancer
Intervention	GEP or expanded IHC tests	All other prognostic tools
Comparator	Standard practice	No prognostic tool excluded
Design	Full economic evaluations (primary or secondary)	Cost descriptive analysis, cost comparisons
Type of publication	Articles or reviews	Letters, editorials, notes, abstracts

GEP – Gene Expression Profiling; IHC –immunohistochemistry

Our search returned 387 citations, after eliminating duplicates. Of those, 328 did not meet our inclusion criteria based on a review of their title and/or abstract. Of the 59 citations left, 28 were excluded after reading their full text because of the study design (18), publication type (8) and intervention (2). Two further studies were excluded since their cost analysis description made reference to other evaluations already included in our review. This left us with 29 relevant studies to be included in our review. Further exploration of the references of the selected articles resulted in two additional references which were discarded, one for its design and the other because of the publication type.

Our literature selection process is illustrated in a flow chart in an Appendix. Out of the 29 economic evaluations identified, one[5] consisted of HTA reports which included the development of original cost models and thus was included in our analysis. Two more[33, 34] consisted of systematic reviews of the economic literature and were only used for checking purposes, in order to ensure no primary economic evaluations had been missed from our review. This left us with 27 relevant economic evaluations.

5.3 Overview of economic evaluations

As shown in Table 5, eight studies were undertaken in Western Europe, with one of them performed in Germany,[35] three in the UK,[5, 36, 37] three in the Netherlands[38-41] and one in France.[42] No Belgian-specific studies were found.

Eight more studies were carried out in the USA[43-50], six in Canada,[51-56] three by the same author in Japan[57-59] and one in Israel.[60] Finally, there was one study undertaken in the USA and Netherlands.[38]

Twenty-three of the 27 studies dated from 2010 or later (see Table 5 for details), and nine[5, 35, 37, 39, 47, 48, 51, 53, 55] were published in 2013 or 2014, reflecting the importance that the topic has gained in the last years. All studies selected were model-based (decision-tree and/or Markov models).

Table 5 – Overview of economic evaluations on gene profiling (GEP) and expanded immunohistochemistry (IHC) tests in chemotherapy decisions for early breast cancer patients

Author	Year	Country	Type of economic evaluation	Perspective	Discount rate; both costs and outcomes (%)
Blomher[35]	2013	Germany	CUA/CEA	Healthcare payer	3%
Chen[43]	2010	USA	CUA/CEA	Healthcare payer	3%
Cosler[44]	2009	USA	CUA/CEA	Healthcare payer	NA
Davidson[51]	2013	Canada	CUA/CEA	Healthcare system	5%
Hall[36]	2012	UK	CUA/CEA	Healthcare system	3,50%
Hannouf[53]	2014	Canada	CUA/CEA	Healthcare system	5%
Hannouf[52]	2012	Canada	CUA	healthcare system	5%
Holt[37]	2013	UK	CUA/CEA	Healthcare system	3,50%
Hornberger[45]	2011	USA	CUA	Healthcare payer	3%
Hornberger[46]	2005	USA	CUA/CEA	Societal	3%
Klang[60]	2010	Israel	CUA	Healthcare payer	3%
Kondo[57]	2012	Japan	CUA/CEA	Healthcare system although presented as societal	3%
Kondo[59]	2011	Japan	CUA	Healthcare system although presented as societal	3%
Kondo[58]	2008	Japan	CUA/CEA	Healthcare payer	3%
Lamond[54]	2012	Canada	CUA	Healthcare system	3%
Mislick[47]	2014	USA	CUA/CEA	Healthcare payer	3% for costs only
Oestreicher[38]	2005	USA, Netherlands	CUA	Societal	3%
Paulden[55]	2013	Canada	CUA/CEA	Healthcare payer	5%

Reed[48]	2013	USA	CUA/CEA	Healthcare system and societal	3%
Retel[39]	2013	Netherlands	CUA/CEA	Healthcare system	4% for costs and 1,5% for effects
Retel[41]	2012	Netherlands	CUA/CEA	Healthcare system	4% costs and 1,5% for outcomes
Retel[40]	2010	Netherlands	CUA/CEA	Healthcare payer	costs:4%; benefits:1,5%
Tsoi[56]	2010	Canada	CUA/CEA	Healthcare system	5%
Valderlaan[49]	2011	USA	CUA	Healthcare payer	3%
Vataire[42]	2012	France	CUA/CEA	Societal	4%
Ward[5]	2013	UK	CUA	Healthcare system	3,50%
Yang[50]	2012	USA	CUA	Healthcare payer	3%

CEA: Cost effectiveness analysis, CUA: Cost utility analysis

5.3.1 Type of economic evaluation

Eleven of the studies performed cost-utility analyses[5, 38, 39, 45, 49, 50, 52-54, 59, 60] and expressed their outcomes in quality-adjusted-life-years (QALYs), while the remaining presented both cost-effectiveness and cost-utility results, presenting their clinical outcomes both in terms of QALYs and life-years-gained (LYG).

5.3.2 Time frame of analyses and discounting

Twelve studies included in this analysis looked at costs and outcomes over a patient's lifetime[5, 36, 38, 43, 45, 46, 51-53, 55, 56, 58] while a further 10 used a time horizon of 20-30 years.[35, 37, 39-42, 44, 49, 54, 60, 61] However, although a lifetime framework is considered the gold standard in economic evaluations, the relatively high age range at which breast cancer tends to be diagnosed may justify these more limited time frames. Only five studies looked at a time period of 10 years or less.[47, 48, 50, 57, 59]

Out of the 27 studies, all but one[44, 61] discounted costs and/or outcomes and gave details on the rates used, which reflected different national recommendations. Thirteen studies used a discount of 3% for both costs and outcomes,[35, 38, 43, 45, 46, 48-50, 54, 57-60] while a further study used the same rate but just for costs.[47] However, the three studies undertaken in the UK used a rate of 3,5%[5, 36, 37] following the recommendations of NICE and five Canadian studies used 5%[51-53, 55, 56] reflecting the recommendations by the Canadian Agency for Drugs and Technologies in Health (CADTH). The studies undertaken by Retel in the Netherlands[39-41] used 4% for costs and 1,5% for outcomes as advised by the Health Care Insurance Board (CVZ) in this country. Finally, a French study[42] used 4% in accordance by the recommendations of the Haute Autorité de Santé (HAS).

5.3.3 Perspective

Eleven studies were performed from a third party payer perspective,[35, 40, 43-45, 47, 49, 50, 55, 58, 60] while a further twelve presented their results from a healthcare system perspective.[5, 36, 37, 39, 41, 51-54, 56, 57, 59, 61] Only three studies used a societal perspective[38, 42, 46] taking into consideration productivity costs and one study provided both a healthcare system and a societal perspective.[48]

5.3.4 Population

The majority of the studies identified via our review modelled populations of women with early breast cancer, estrogen receptor positive (ER+) and lymph node negative (LN-). However, three studies focused purely on LN+ patients,[36, 49, 53] while a further five evaluations included in their modelling exercise some LN+ patients.[35, 37, 38, 54, 59] Out of the studies that mentioned the HER2 status of the hypothetical patient population, most included only HER2- patients, with only three explicitly mentioning the inclusion of HER2+ patients.[40, 58, 59] The mean age of the population varied, although most used a mean age between 50 and 60.

5.3.5 Intervention and comparator

Out of the 27 studies, 18 evaluated the cost-effectiveness of Oncotype DX versus standard practice, five focused on MammaPrint versus current practice,[38-40, 43, 57] while the HTA by Ward et al. included four original models assessing the cost-effectiveness of two gene expression profiling tests (i.e. Oncotype DX and MammaPrint) and two expanded immunohistochemistry tests (i.e. Mammostrat and IHC4).[5]

Comparators described as standard practice varied depending on the study and included “absence of test”, adjuvant Online! (AO), the Nottingham Prognosis Index (NPI) or international clinical guidelines such as those by the National Comprehensive Cancer Network (NCCN) or St Gallen.

Only two studies compared different gene profiling tests, more specifically MammaPrint versus Oncotype DX,[41, 50] while a further two evaluated Oncotype DX versus expanded immunohistochemistry tests, with Mislick et al. assessing Oncotype DX versus Mammostrat.[47] and Ward et al. looking at Oncotype DX versus IHC4[5].

5.3.6 Cost and outcome inputs

Different sources were consulted to derive costs. In addition to the published literature, hospital records, national administrative data and medical fees were also used. For the cost of tests most studies mentioned personal communication with the manufacturers of the specific test being evaluated as their costing source.

With regard to outcomes, all studies used for their models data from the published literature. For studies looking at Oncotype DX, the most common reference used for risk reductions was the study by Paik et al. published in 2006.[62] Evaluations on MammaPrint were also based on the literature, but different sources were consulted.

Quality of life (QoL) is an important factor to bear in mind when studying chronic conditions such as breast cancer, in which the QoL can be affected by the illness per se and its evolution, but also by other factors such as the potential adverse events (AEs) linked to chemotherapy treatment. QoL values for all studies were derived from the literature and often based on a study published in 2007 by Lidgren et al.[63] or from a less recent study by Earle et al. published in 2000.[64]

5.3.7 Modelling

All studies consisted of modelling exercises. Five of which[44, 45, 57-61] used the model structure originally designed by Hornberger et al. in 2005.[46] The structure remained simple, mostly considering four health states: no recurrence, recurrence without metastatic progression, recurrence with metastatic progression and death.

5.3.8 Results

5.3.8.1 Incremental costs

Table 6 shows the mean costs reported in the 27 studies included in our review. Comparisons between studies are difficult primarily because of the different costs borne in mind, differences in cost definitions, in standard practice and in prices as well as in the perspectives used for the analyses.

Although overall, most studies found that there was an incremental cost when using GEP or expanded IHC tests compared to standard practice, five evaluations comparing Oncotype DX with standard practice did find that gene expression profiling could result in savings.[35, 45, 46, 49, 65] In addition to these, Hannouf et al. found that Oncotype DX could be less costly when compared to no testing for premenopausal women but not for postmenopausal,[52] while a further study found Oncotype DX to be cheaper than using Tamoxifen and chemotherapy for all patients, but more expensive when compared to Tamoxifen alone.[44, 61] Two studies out of the six comparing MammaPrint to standard practice found it to be cost saving,[38, 39] and the model by Ward et al. on IHC4 concluded that its use would offer savings compared to current practice.[5]

When comparing different tests, a recent study by Mislick et al.[47] found Mammostrat to be a cheaper alternative compared to Oncotype DX. The two evaluations that looked at MammaPrint versus Oncotype DX showed that MammaPrint appeared to be a cheaper than Oncotype DX.[41, 50] Finally, the study by Ward et al. reported potential savings when using IHC4 compared to Oncotype DX.[5]

With regard to the reported prices for the tests, these ranged from a low of CAN$3 650 (€2 575)[56] to a high of US$4 500 (€3 500)[59] for Oncotype DX, from €2 675[39-41] to GBP2 675 (€3 402)[5] for MammaPrint and from GBP1 135 (€1 443) to US$2650 (€2 061) for Mammostrat. The model by Ward et al. (the only one covering IH4) mentioned costs of GBP150 (€191).[5]

Table 6 – Costs of gene expression profiling or expanded immunohistochemistry tests in early breast cancer

Author	Costing yr	Time horizon	Test	Population	Costs included	Cost source	Mean incremental cost
Versus standard practice							
Blomher 2013 Germany	2011	30	OT	LN0-3; ER+; HER2-	Test, chemo, follow-up, recurrence, productivity	Lit. and manufacturers	-€561
Cosler 2009 USA	NA	20	OT vs Tamox OT vs Tamox+chemo	LN-, ER+	Test, chemo, AEs, recurrence	Lit. and data from CMS	US$4 272 -US$2256
Davidson 2013 Canada	2010	Lifetime	OT	Stage I-II, LN-, ER+, HER2-	Test, chemo, AEs	manufacturers; chemo: BCCA's Systemic Therapy program; Other: Tsoi et al. 2010	CAN$2 188
Hall 2012 UK	2011	Lifetime	OT	LN+, ER+	Test, chemo, relapse, palliative care	UK NHS reference costs and BNF	GBP860
Hannouf 2012 Canada	2010	Lifetime	OT	Pre menop, ER+/PR+; LN- Post menop, ER+/PR+; LN-	Tests, chemo, monitoring and relapse	Provincial administrative database	-CAN$50/woman CAN$3 700/woman
Hannouf 2014 Canada	2012	Lifetime	OT	Post-menop, ER+/PR+, LN1-3	Test, surgery, radiation, chemo, endocrine therapy,	Provincial administrative database	CAN$36.2/woman

					serious AEs and surveillance		
Holt 2013 UK	2010	30	OT	ER+, LN-/LN+(1)	Test, adjuvant chemo, AEs, follow-up, recurrence	UK sources and hospital records	GBP888
Hornberger 2011 USA	NA	Lifetime	OT	LN-, ER+	Test, adjuvant chemo, AEs, recurrence; palliative care	Humana US insurance program	-US$1 160/test
Hornberger 2005 USA	2004	Lifetime	OT	LN-, ER+	Test, adjuvant chemo, AEs, follow-up, palliative care	Lit.	-US$2 028/patient
Klang 2010 Israel	NA	30	OT	LN-, ER+.	Pharmaceuticals and relapse	Healthcare organisation registries and manufacturer	US$1 828/patient tested
Kondo 2011 Japan	NA	10	OT	LN+/-, ER+, HER2 -/+ LN-, ER+, HER2 -/+	Test, follow-up, recurrence, palliative care	National medical care fee schedule	US$2700 US$2 407
Kondo 2008 Japan	NA	Lifetime	OT vs St Gallen OT vs NCCN	LN-, ER+	All direct treatment costs	National medical care fee schedule	US$2 550 US$2516
Lamond 2012 Canada	2011	25	OT	LN-; ER+ LN+; ER+	Test, chemo, AEs, monitoring, recurrence	Population-based cohort data and lit.	CAN$2585 CAN$864
Paulden 2013 Canada	2012	Lifetime	OT (high risk AO) OT (intermediate risk AO) OT (low risk AO) OT (all)	LN-, ER+, HER2-	Test, chemo, hormonal therapy, AEs, distant recurrence, follow-up, palliative care	Lit. (including Tsoi et al 2010)	CAN$230 CAN$330 CAN$1 890 CAN$2460

Reed 2013 USA	2011	10 or to recurrence	OT	LN-, ER+	Test, chemo, AEs, recurrence	Lit.	US$1741
Tsoi 2010 Canada	2008	Lifetime	OT	LN-, ER+, HER2	Test, chemo, hormonal treatment, AEs, recurrence, follow-up, palliative care	Lit., experts and manufacturers	CAN$4102
Valderlaan 2011 USA	2009	30	OT	LN+(1-3), ER+, HER2-	Test, chemo, AEs, recurrence, treatment	Centers for Medicare & Medicaid	-US$384/patient
Vataire 2012 France	2011	30	OT	LN-, ER+, HER2-	Test, chemo, AEs, recurrence, productivity	Laas et al. 2012	-€570
Ward 2013 UK	2010	Lifetime	OT (all) OT (NPI>3,4)	ER+, LN-, HER2-	Test, chemo, hormonal therapy, AEs, distant recurrence, local recurrence, palliative care	BNF, experts and NHS reference costs	GBP2593 GBP2127
Ward 2013 UK	2010	Lifetime	Mammostrat (all) Mammostrat (NPI>3,4)	ER+, LN-, HER2-	Test, chemo, hormonal therapy, AEs, distant recurrence, local recurrence, palliative care	BNF, experts and NHS reference costs	GBP1345 GBP1273
Chen 2010 USA	2007	Lifetime	MP	ER+, T1 or T2, LN-, HER2-	Test, chemo, recurrence, palliative care	Lit., vademecum and manufacturers	US$1332
Kondo 2012 Japan	NA	1-5, 6-10	MP	LN–, ER+, HER2-	Test, chemo, AEs, recurrence; palliative care	National medical care fee schedule	US$2571
Oestreicher 2005 USA, Netherlands	2003	Lifetime	MP	Pre-menop, TI-II, LN+51%, ER+77%	Test, chemo, recurrence	Lit.	-US$2882

Retel 2013 Netherlands	2005	20	MP	Grade II, ER+, LN-	Test, chemo, hormonal therapy, disease free survival, relapse, distant metastasis costs, death	Lit. (Lidgren et al. 2008)	-€2401
Retel 2010 Netherlands	2005	20	MP vs St Gallen MP vs Adjuvant Online!	LN-, ER+, HER2+/-	Test, chemo, hormonal therapy, trastuzumab, cost per health state, distant metastasis, death	Lidgren et al. 2008 and Dutch sources (Health Care Insurance Board 2006)	-€7430 €1130
Ward 2013 UK	2010	Lifetime	MP (all) MP (NPI>3,4)	ER+, LN-, HER2-	Test, chemo, hormonal therapy, AEs, distant recurrence, local recurrence, palliative care	BNF, experts and NHS reference costs	GBP3609 to GBP4119 GBP3997 to GBP5142
Ward 2013 UK	2010	Lifetime	IHC4 (all) IHC4 (NPI>3,4)	ER+, LN-, HER2-	Test, chemo, hormonal therapy, AEs, distant recurrence, local recurrence, palliative care	BNF, experts and NHS reference costs	-GBP175 -GBP492
Comparisons between tests							
Mislick 2014 USA	2009	10	OT vs Mammostrat	LN-, ER+	Test, chemo, AEs, recurrence, palliative care, death	Yang et al. 2012	US$2269
Ward et al 2013 UK	2010	Lifetime	OT vs IHC4 (all) OT vs IHC4 (NPI>3,4)	ER+, LN-, HER2-	Test, chemo, hormonal therapy, AEs, distant recurrence, local recurrence, palliative care	BNF, experts and NHS reference costs	-GBP2768 -GBP2619

Retel 2012 Netherlands	2010	20	MP vs OT	ER+, LN-	Test, other direct costs	Lidgren et al. 2008 + Dutch sources (Health Care Insurance Board 2006)	-€1475 to -€3941
Yang 2012 USA	2009	10	MP vs OT	ER+, LN-	Tests, chemo, AEs, recurrence, palliative care	Lit.	-US$6284

AEs: Adverse events, AO: Adjuvant Online!, BNF: British National Formulary, ER: Estrogen receptor, HER2: Human Epidermal growth factor Receptor 2, *LN: Lymph node, MP: MammaPrint, NCC: National Comprehensive Cancer Network, NPI: Nottingham Prognosis Index, OT: OncotypeDx*

5.3.8.2 Incremental outcomes

Table 7 shows the outcomes reported in the studies included in this review. Twelve evaluations looked at life years gained (LYG) in addition to QALYs as a key outcome. From these, all but one[46] reported positive outcomes linked to the use of GEP or expanded IHC tests compared to standard practice. Life years gained ranged from a maximum of 2,20[44] to a minimum of 0,05,[57] with a median reported gain of 0,16. Similarly, only two studies found a decrease in QALYs. The first one, published in 2005 on a premenopausal, mixed population (LN+51%; ER+77%)[38], quoted a loss of 0,22 QALYs when using MammaPrint compared to standard practice (NIH guidelines). The difference appeared to be due to the lower sensitivity assumed for MammaPrint to predict risk of recurrence (84%). The authors concluded that a sensitivity of 95% would be needed for MammaPrint to offer more QALYs than NIH guidelines. The second study reporting negative outcomes was the modelling exercise by Ward et al.[5] comparing Mammostrat versus standard practice. The authors reported a positive outcome for Mammostrat when used in all patients, but a negative one when used in NPI>3,4 only. It should be noted that this was the only evaluation comparing Mammostrat to standard practice and that the evidence for this test remains very limited.

All other evaluations reported positive outcomes with gains in QALYs that ranged from a high of 1,2[40] to a low of 0,04 for the modelling exercise comparing IHC4 to standard practice by Ward et al.[5] The median reported gain was of 0,14 QALYs.

The four studies assessing LN+ populations appeared to show lower QALY gains overall (0,06-0,16; median: 0,10) than those on LN- populations (0,04-1,2; mean:0,15). Nevertheless, their results remain positive towards testing.

With regard to the comparisons between tests, two out of the four evaluations favoured the use of Oncotype DX versus Mammostrat[47] and IHC4,[5] while the remaining two compared MammaPrint and Oncotype DX and reported QALY gains with MammaPrint.[41, 50]

5.3.8.3 Incremental cost-effectiveness ratios (ICERs)

Table 8 shows that, overall, the results appear favourable to the GEP tests when these are compared to standard practice, with 7/24 of the studies showing GEP test to be dominant (i.e. cheaper and more effective),[35, 39, 40, 45, 46, 49, 65] and a further study showing dominance for premenopausal women only.[52] The majority of the remaining studies displayed ICERs below €25 000, with only three presenting an ICER between €25000 and €50000.[52, 56, 57] One study reported the GEP test to be less efficacious but cheaper than standard practice (NIHCC).[38] Finally, the study by Ward et al.[5] which analysed two different GEP tests showed different results depending on the test analysed, the clinical data used or the approach followed (testing all patients versus NPI>3.4 only). Their findings appear particularly uncertain in the case of MammaPrint.

Six studies looked at MammaPrint,[38-41, 43, 57] while the remaining covering GEP tests versus standard practice focused on Oncotype Dx.

Ward et al. included in their analyses expanded IHC tests such as Mammostrat and IHC4,[5] and reported positive results for IHC4 versus

current practice (dominant), but negative results when modelling Mammostrat versus current practice (not cost-effective at a threshold of GBP20 000/QALY when testing all patients, and dominated when testing NPI>3,4 only).

Only four studies compared different tests, with two looking at MammaPrint versus Oncotype DX[41, 50], one studying Oncotype DX versus Mammostrat.[47] and a further one assessing Oncotype DX vs IHC4.[5] Amongst these, the two comparing MammaPrint versus Oncotype DX showed MammaPrint to be dominant (cheaper and more effective).The comparison of Oncotype DX versus Mammostrat showed the former to be more expensive while offering similar outcomes. Finally, the comparison by Ward et al. between Oncotype DX and IHC4 favoured IHC4.[5]

Table 7 – Outcomes of GEP or expanded IHC tests in early breast cancer

Author	Test	Population	Outcomes	Incremental LYG	Incremental QALYS
Versus Standard Practice					
Blomher 2013 Germany	OT	LN0-3; ER+; HER2-	LYG & QALYs	0,07	0,07
Cosler 2009 USA	OT vs Tamox OT vs Tamox+chemo	LN-, ER+	LYG & QALYs	2,2 No sig different	NA NA
Davidson 2013 Canada	OT	Stage I-II, LN-, ER+, HER2-,	LYG & QALYs	0,31	0,32
Hall 2012 UK	OT	LN+, ER+	LYG & QALYs	0,15	0,16
Hannouf 2012 Canada	OT	Pre menop, ER+/PR+; LN- Post menop; ER+/PR+; LN-	QALYs	NA NA	0,05 0,062
Hannouf 2014 Canada	OT	Post-menop,ER+/PR+, LN(1-3)	QALYs	NA	0,08
Holt 2013 UK	OT	ER+, LN-/LN+(1) ,	LYG & QALYs	0,16	0,14
Hornberger 2011 USA	OT	LN-, ER+	QALYs	NA	0,16
Hornberger 2005 USA	OT	LN-, ER+	LYG & QALYs	-0,0421	0,086

Klang 2010 Israel	OT	LN-, ER+.	QALYs	NA	0,17
Kondo 2011 Japan	OT	LN+/-, ER+, HER2 -/+ LN-, ER+, HER2 -/+	QALYs	NA NA	0,47 0,63
Kondo 2008 Japan	OT vs St Gallen OT vs NCCN	LN-, ER+	LYG & QALYs	0,221 0,083	0,237 0,097
Lamond 2012 Canada	OT	LN-; ER+ LN+; ER+	QALYs	NA NA	0,27 0,06
Paulden 2013 Canada	OT (high risk AO) OT (intermediate risk AO) OT (low risk AO) OT (all)	LN-, ER+, HER2-	LYG & QALYs	0,269 0,164 0,096 0,53	0,213 0,13 0,084 0,429
Reed 2013 USA	OT	LN-, ER+	LYG & QALYs	0,19	0,16
Tsoi 2010 Canada	OT	LN-, ER+, HER2	LYG & QALYs	0,064	0,065
Valderlaan 2011 USA	OT	LN+(1-3), ER+, HER2-	QALYs	NA	0,127
Vataire 2012 France	OT	LN-, ER+, HER2-	LYG & QALYs	0,15	0,14
Ward 2013 UK	OT (all) OT (NPI>3,4)	ER+, LN-, HER2-	QALYs	NA NA	0,08 0,22
Ward 2013 UK	Mammostrat (all) Mammostrat (NPI>3,4)	ER+, LN-, HER2-	QALYs	NA NA	0,05 -0,05
Chen 2010 USA	MP	ER+/-, T1 or T2, LN-, HER2-	LYG & QALYs	0,143	0,153
Kondo 2012 Japan	MP	LN–, ER+, HER2-	LYG & QALYs	0,05	0,06
Oestreicher 2005 USA, Netherlands	MP	Pre-menop, TI-II, LN+51%, ER+77%	QALYs	NA	-0,22

Retel 2013 Netherlands	MP	Grade II, ER+, LN-	QALYs	NA	0,61
Retel 2010 Netherlands	MP vs St Gallen MP vs AO	LN-, ER+, HER2+/-	LYG & QALYs	-0,26 0,2	1,2 0,24
Ward 2013 UK	MP (all) MP (NPI>3,4)	ER+, LN-, HER2-	QALYs	NA NA	0,29-0,08 0,66-0,18
Ward 2013 UK	IHC4 (all) IHC4 (NPI>3,4)	ER+, LN-, HER2-	QALYs	NA NA	0,04 0,13
Comparisons between tests					
Mislick 2014 USA	OT vs Mammostrat	LN-, ER+	LYG & QALYs	0,002	0,005
Ward et al 2013 UK	OT vs IHC4 (all) OT vs IHC4 (NPI>3,4)	ER+, LN-, HER2-	QALYs	NA NA	0,04 0,09
Retel 2012 Netherlands	MP vs OT	ER+, LN-	LYG & QALYs	0,14-0,40 depending on clinical source	0,08-0,31 depending on clinical source
Yang 2012 USA	MP vs OT	ER+, LN-	QALYs	NA	0,097

AO: Adjuvant Online!,ER: Estrogen receptor, HER2: Human Epidermal growth factor Receptor 2, *LN: Lymph node, LYG: life years gained, MP: MammaPrint, NCC: National Comprehensive Cancer Network, NPI: Nottingham Prognosis Index, OT: Oncotype Dx, QALYs: Quality Adjusted Life Years*

Table 8 – ICERs for GEP or expanded IHC tests in early breast cancer

Author	Test	Population	ICER	Prob. Of test being cost-effective
Versus standard practice				
Blomher 2013 Germany	OT	LN0-3; ER+; HER2-	OT dominates	87% for OT to dominate
Cosler 2009 USA	OT vs Tamox OT vs Tamox+chemo -all	LN-, ER+	OT dominates US$ 4 432/QALY	NA 70% for OT to be dominant
Davidson 2013 Canada	OT	Stage I-II, LN-, ER+, HER2-,	CAN$6 630 /QALY	94% at WTP CAN$50 000/QALY
Hall 2012 UK	OT	LN+, ER+	GBP5 529/QALY*	61% at WTP GBP30 000/QALY
Hannouf 2012 Canada	OT	Pre-menop, ER+/PR+; LN- Post-menop, ER+/PR+; LN-	OT dominates CAN$60 000/ QALY	54% at WTP CAN$100 000/QALY 62% at WTP CAN$100 000/QALY
Hannouf 2014 Canada	OT	Post-menop, ER+/PR+, LN 1-3	CAN$464/QALY	72% at WTP CAN$100 000/QALY
Holt 2013 UK	OT	ER+, LN-/LN+(1)	GBP6 232/QALY	99,6% at WTP GBP20 000/QALY
Hornberger 2011 USA	OT	LN-, ER+	OT dominates	81% for OT to dominate
Hornberger 2005 USA	OT	LN-, ER+	OT dominates	>67% for OT to dominate
Klang 2010 Israel	OT	LN-, ER+.	US$10 770/QALY	95% at WTP US$35 000/QALY
Kondo 2011 Japan	OT	LN+/-, ER+, HER2 -/+ LN-, ER+, HER2 -/+	US$5 685/QALY US$3 848/QALY	NA NA
Kondo 2008 Japan	OT vs St Gallen OT vs NCCN	LN-, ER+	US$10 774/QALY US$26 065/QALY	NA NA
Lamond 2012 Canada	OT	LN-; ER+ LN+; ER+	CAN$9 591/QALY CAN$14 844/ QALY	100% at WTP CAN$50 000/QALY
Lyman 2007 USA	OT/Tamox alone OT/Tamox+ chemo	LN-, ER+	US$1 944/LYG Similar LYG; cheaper	70% at WTP US$15 000/LYG
Paulden 2013 Canada	OT (high risk by AO) OT (intermediate risk by AO) OT (low risk by AO)	LN-, ER+, HER2-	CAN$1 111/QALY CAN $ 2526/QALY Can$22 440/QALY	99,82% at WTP CAN$50 000/QALY 99,4% at WTP CAN$50 000/QALY 78,46 at WTP CAN$50 000/QALY
Reed 2013 USA	OT	LN-, ER+	US$10 788/QALY	99% at WTP US$50 000/QALY

Tsoi 2010 Canada	OT	LN-, ER+, HER2	CAN$63 064/QALY	NA
Valderlaan 2011 USA	OT	LN+(1-3), ER+, HER2-	OT dominates	NA
Vataire 2012 France	OT	LN-, ER+, HER2-	OT dominates	NA
Ward 2013 UK	OT (all) OT (NPI>3,4)	ER+, LN-, HER2-	GBP29,502/QALY GBP 9 774/QALY	12,44% at WTP GBP20 000/QALY 91,56% at WTP GBP20 000/QALY
Ward 2013 UK	Mammostrat (all) Mammostrat (NPI>3,4)	ER+, LN-, HER2-	GBP27 731/QALY Dominated	36% at WTP GBP20 000/QALY 18% at WTP GBP20 000/QALY
Chen 2010 USA	MP	ER+, T1 or T2, LN-, HER2-	US$5 908/QALY	NA
Kondo 2012 Japan	MP	LN–, ER+, HER2-	US$43 044/QALY	NA
Oestreicher 2005 USA, Netherlands	MP	Pre-menop, TI-II, LN+51%, ER+77%	MP cheaper but less effective	95% for MP to be cheaper but less effective
Retel 2013 Netherlands	MP	Grade II, ER+, LN-	MP dominates	97% for MP to dominate
Retel 2010 Netherlands	MP vs St Gallen MP vs Adjuvant Online!	LN-, ER+, HER2+/-	MP dominates €4614/QALY	NA NA
Ward 2013 UK	MP (all) MP (NPI>3,4)	ER+, LN-, HER2-	GBP12 240-GBP53 058/QALY GBP6 053-GBP29 569	NA NA
Ward 2013 UK	IHC4 (all) IHC4 (NPI>3,4)	ER+, LN-, HER2-	IHC4 dominates IHC4 dominates	82%for IHC4 to dominate
Comparisons between tests				
Mislick 2014 USA	OT vs Mammostrat	LN-, ER+	More expensive and similar outcomes	NA
Ward et al 2013 UK	OT vs IHC4 (all) OT vs IHC4 (NPI>3,4)	ER+, LN-, HER2-	GBP64 111 GBP31 125	0,40% at WTP GBP20 000 18,60% at WTP GBP20 000
Retel 2012 Netherlands	MP vs OT	ER+, LN-	MP dominates	NA
Yang 2012 USA	MP vs OT	ER+, LN-	MP dominates	82% for MP to dominate

**For cut off point RS 18*
AO: Adjuvant Online!,ER: Estrogen receptor, HER2: Human Epidermal growth factor Receptor 2, *LN: Lymph node, LYG: life years gained, MP: MammaPrint, NCC: National Comprehensive Cancer Network, NPI: Nottingham Prognosis Index, OT: Oncotype Dx, QALYs: Quality Adjusted Life Years, WTP: willingness to pay*

5.3.9 Sensitivity analysis

Uncertainty is intrinsic to any economic evaluations and should therefore always be accounted for. All evaluations performed some kind of sensitivity analysis to assess the robustness of their results, although six of them limited their tests to one-way sensitivity analyses,[43, 49, 57-59, 65] and a further presented results of one and two-way sensitivity analyses but did not engage into a probabilistic evaluation of their results.[56]

The remaining studies undertook univariate and probabilistic sensitivity analyses, and although in some cases their analysis supported their main findings,[37, 42, 46, 54] overall there were important uncertainties present. Results appeared to be primarily sensitive to long-term recurrence and survival rates,[5, 36, 43, 45, 51-53, 56, 57, 59, 60] the distribution of risk scores, thresholds and chemotherapy decisions,[5, 38, 43, 51-53, 56] costs of chemotherapy[5, 38, 44, 45, 57, 60] and test costs.[36, 38, 51, 57, 59]

5.3.10 Conflict of interest

All 28 studies included in their manuscripts a declaration of conflict of interest. Only 10 out of them reported no conflict of interest.[5,44, 52-59] The existence of conflicts of interest may introduce a bias which could affect the validity of the study results, although there is, up to date, no hard evidence on this.

5.4 Discussion and conclusions

Despite the lack of randomized controlled trials up to date, looking at the clinical utility of these tests, there is a high level of consistency indicating that GEP or expanded IHC tests are likely to be cost-effective compared to standard practice. However, despite this consistency there is a number of important points worthwhile considering. We will get back to these in our general discussion section (section 7).

Sources of clinical data

The main limitation of the economic evaluations reviewed relates to the need to combine and model multiple sources of data given the fact that there are no studies following patients from the time of testing to final health outcomes. Such a combination increases the uncertainty surrounding model assumptions, since important factors which could play a role in the final results are not kept constant from one study to another (e.g. population characteristics, standard practice, etc). For Oncotype DX most of the models used data from Paik et al.[62] for risk reductions, but clinical sources used in models for other tests varied more.

Modelling/assumptions

Most of the studies included some assumptions not well backed-up with literature. In particular, some of the models assumed that all patients categorised as having a high risk of recurrence would be treated with chemotherapy, while none of the low risk patients would receive such treatment. Similarly the assumptions surrounding treatment with chemotherapy for those patients with an intermediate risk varied from one study to another. All of these factors have an important weight on the overall results since most of the savings coming from the testing alternatives are due to reductions in the amount of chemotherapy used.

The price of the test is another crucial factor to consider, as shown in the evaluation by Ward et al.[5] where at the original list price of Oncotype DX the authors concluded that testing would not be likely to be cost effective. This changed once the manufacturers offered a reduction in the price of the test under a confidential agreement (see NICE recommendations, http://www.nice.org.uk/guidance/dg10).

On the other hand, IHC4 is very likely to be cost-effective due to its low price but the highly limited evidence on its effectiveness makes it hard to draw any clear conclusions on this regard.

Transferability of results to the Belgian situation

The potential value of modelling the Belgian situation would be limited at present, given the lack of RCTs on the clinical utility of these tests. Furthermore, there are a number of assumptions that would need to be made regarding the baseline population receiving chemotherapy, the current risk stratification systems most commonly used and the proportion of patients for which risk is unclear under standard practice, as well as the proportion of patients who following a testing strategy would be treated with chemotherapy.

Conclusions

- **The available economic evaluations appear favorable to the GEP tests when compared to standard practice, with 7/24 of the studies showing GEP tests to be dominant (i.e. cheaper and more effective) and the majority of the remaining studies displaying ICERs <€25 000.**
- **Most evidence is available for Oncotype DX and Mammaprint.**
- **Economic evaluations on expanded IHC tests are limited to two studies on Mammostrat with uncertain results and one on IHC4, with positive results when comparing IHC4 to standard practice.**
- **There is also limited evidence (four studies only) comparing different tests.**
- **The main limitation of the models here reviewed is linked to the lack of clinical studies looking at the long-term consequences of testing, which made all evaluations rely on inputs from different sources which had to be modelled together resulting in important uncertainties.**

6 BUDGETARY IMPACT SCENARIOS

6.1 Background

The aim of this chapter is to offer an insight into the potential short-term budgetary impact that the introduction in Belgian clinical practice of GEP or IHC tests could bring, if these tests were reimbursed by the public insurance system for ER+, LN-, HER2-, early breast cancer patients.

As seen in the review of the clinical literature (see chapter 4), although there is evidence on the prognostic ability of these tests, there is at present a lack of reliable long-term data on their clinical utility. Nevertheless, there would be short term financial implications if these tests were reimbursed. On the one hand, the prices of the GEP tests are, at present, high. On the other hand, the economic evaluations reviewed (see chapter 5) have shown that these tests, may help to reduce chemotherapy treatment in those patients in which the risk of recurrence is unclear when stratified by means of current tools. Such avoidance could, in turn, off-set some of the extra costs of the test.

Avoiding chemotherapy treatment, would not simply save the pharmacological costs linked to the regimen used but would also avoid certain costs linked to its administration, as well as potential AEs and their consequent clinical management. Furthermore, it can also shorten or avoid the sick leave period commonly required to complete a full course of chemotherapy.

6.2 Methods

6.2.1 Data sources

6.2.1.1 Epidemiological data and health care utilisation data:

Data from the Belgian Cancer Registry was used. Epidemiological data was coupled with IMA data, allowing to identify the right population as well as information on the proportion of that population treated with chemotherapy. The target population is described as follows: ER+, LN-, HER2- patients following breast conserving surgery or mastectomy.

These data are only available for the year 2008 given the time consuming and resource intense process required to couple the two databases (the Belgian Cancer Registry and IMA databases). Despite this limitation, there

is no specific reason why the population should have changed in any significant way from 2008 to the present year and thus 2008 data were used as the basis of our estimations.

6.2.1.2 Cost data

For costs, the consultation of different sources resulted in data relating to different years. Costs were updated when necessary to the year 2013 using consumer price index data (CPI) data (Eurostats - http://epp.eurostat.ec.europa.eu) in order to offer a consistent costing approach.

6.2.1.3 Chemotherapy related costs

Commonly used chemotherapy regimens were discussed with a panel of experts in order to select those most frequently prescribed in Belgium for the specific population of interest. Once these were identified, pharmacological prices for these regimens were extracted from the CBIP/BCFI database (www.cbpi.be). The cheapest alternative for a specific product was selected for the purpose of our calculations.

6.2.1.4 For chemotherapy in day-care

Hospitals receive a maximum lump sum, covering costs linked to the administration of the therapy. This maximum lump sum varies from one hospital to another with a weighted mean of €131.25 per day in 2012, the latest year for which complete data is currently available.

6.2.1.5 Blood tests

A lump sum per day for clinical biology of €18.83 in 2012 was also included in the calculations to cover for the blood tests required to be taken before each chemotherapy cycle.

Day care and blood test costs are incurred in at every chemotherapy cycle and thus their weight on the overall costs depends not only on the specific treatment regimen used but also on the number of cycles required to complete a full course of treatment.

6.2.1.6 Other costs related to chemotherapy administration

Most often a central venous access line is used in order to facilitate the administration of the chemotherapy at each cycle. Hospitalisation for a day is required at an approximate 2010 per diem price of €388.16. Reported per diem hospitalisation costs are based on the 100% weighted average per diem price for all acute beds nationally and were extracted from the technical cell hospital data (https://tct.fgov.be). The price of the implant is €300 in the present year, while the placing fee is of €107. All costs related to the placement of the central venous access line are one-off costs.

6.2.1.7 Costs of prophylaxis or management of adverse events (AEs)

Only those most frequent early AEs following administration of a chemotherapy cycle which could have a significant impact on costs, were considered. Prophylactic treatment for nausea and vomiting was assumed to be administered for every patient receiving chemotherapy, although the specific products used varied form one chemotherapy regimen to another. Assumptions around the products used were based on the information received by the panel of Belgian experts consulted during this project.

An additional AE for which data was included in our estimations was neutropenia, a relatively common consequence of chemotherapy treatments. For this, the experts and the published evidence was consulted/checked in order to define rates of neutropenia in early breast cancer patients receiving chemotherapy treatment with or without prophylactic treatment.

6.2.1.8 Other costs

Chemotherapy-related hospitalisation was also considered. Rates from the literature were used and in particular those from Barcenas et al. 2014[66]

Transport costs were excluded from the analysis given the difficulties in obtaining reliable mean estimates in this regard (public reimbursement dependant on kilometres from and to the hospital where treatment takes place).

The prices of the tests were calculated from the literature and a mean price per test in euros was used in our base case calculations. Given that most economic evaluations focused on either Oncotype DX or MammaPrint, and the higher price as well as more evidence available for the former, the base case scenario focused on Oncotype DX. This was also thought to provide a conservative approach, since all other tests, including MammaPrint, appear to be commercially available at lower prices.

6.2.2 Perspective

The perspective considered is that of the healthcare payer, and in particular that of the RIZIV/INAMI. However, considering the importance of absenteeism in cancer patients, we also present estimations including sick-leave payments covered by the INAMI/RIZIV over the chemotherapy treatment period.

Patient co-payments are not considered in the calculations, but since most cancer treatment is covered by the public health insurance in Belgium, out of pocket costs should remain limited.

6.2.3 Time horizon

Budget calculations as opposed to cost-effectiveness evaluation focus on the short term, defined in this case as the period over which the chemotherapy treatment is usually provided. Thus, potential long term AEs such as heart failure or recurrence after months or years of disease free survival are not included in our analysis. In the context of breast cancer patients for whom the most important clinical benefits are often perceived in the long term, the calculations here included are likely to offer a sub estimation of the overall potential long-term savings that the tests could bring, assuming their clinical utility in the target population was proven. Given the current uncertainties on that regard, we have limited the analysis to a budget impact estimation. A full cost-effectiveness evaluation should be undertaken once the results from the MINDACT, GERICO 11 and the TAILOR-X trials or other RCTs become available.

6.2.4 Uncertainty

Because of the limited detailed data currently available, the scenario calculations performed were based on values subject to an important degree of uncertainty. Possible minimum and maximum values for each variable were identified, and their impact on the overall costs calculated in a simple one-way sensitivity analysis, altering each variable independently in order to identify those most likely to influence the overall results.

6.2.5 Scenarios studied

The availability of NPI data for the target population and the recent recommendations by NICE to limit GEP testing to patients with an intermediate risk score (3,4< NPI >5,4) (http://www.nice.org.uk/guidance/dg10) made it possible and appropriate to study two different scenarios:

1. in which all patients are tested independently of their NPI, or indeed any other current risk stratification systems
2. in which only those with 3,4<NPI<5,4 are tested and reclassified before chemotherapy treatment decisions are made.

6.2.6 Key assumptions made

A number of important assumptions were made:

1. Given the lack of complete data on the most commonly used combinations of chemotherapy agents currently used in ER+, LN-, HER2- early breast cancer patients, a panel of Belgian experts was consulted on this regard. From the discussion with the experts, two chemotherapy combinations were selected and thought to be the most representative. These included:
   - four cycles of the combination of epirubicin and cyclophosphamide (EC), followed by 12 weekly cycles of paclitaxel, assumed to be used in approximately 50% of the population
   - or four cycles of the combination docetaxel and cyclophosphamide (TC), assumed to be used in the other half of the population.

   For the first of these two regimens 10% of the population was assumed to receive dose dense chemotherapy (cycles separated by two-week periods as opposed to the three-week periods for conventional chemotherapy). For the second regimen (TC), this was estimated to be 40%. This estimate was kept relatively low given the nature of the population studied (LN- patients only).

   Although the experts though that the combination of FEC100 is likely to still be used in some Belgian centres or areas, they believed there has been (and continues to be) a shift from such regime to the above mentioned mainly because of improved tolerability profiles compared to FEC100[67].

2. The cost of blood tests can vary significantly depending on specific patient needs, so a minimum cost was included in our base case scenario in order to provide a conservative estimate. This cost represented the lump sum per day for clinical biology at €18.83. It was assumed that 20% additional blood tests would need to be performed for various clinical reasons (e.g. nadir control after first cycle if not dose dense, extra blood count for clinical problems).
3. Prophylaxis with anti-emetics to minimise potential nausea and vomiting was assumed to be used in all patients subject to chemotherapy treatment. However, actual products used vary depending on practices and chemotherapy regimens. An assumption was made that for patients receiving EC+ paclitaxel, anti-emetic treatment would consist on aprepitant on days 1 to 3 and short acting setron on day 1. The assumption for patients on the TC regime was that they would receive short acting setron on day 1.
4. The rates of neutropenia were based on the published literature and discussions with the experts. Twelve percent for the EC+ paclitaxel group and 15% for the TC group were used as estimations.[66] These rates applied for those not receiving prophylactic treatment, and were assumed to diminish to approximately 4% and 3% for the two regimens respectively when receiving prophylaxis.[68].
5. Rates of chemotherapy related hospitalisations within 6 months of chemotherapy initiation were calculated as the difference between hospitalisation due to neutropenia assumed for our model and the overall chemotherapy related hospitalisation rate mentioned in Barcenas et al.[66] for patients over and below 65 years of age.
6. Primary prophylactic treatment for neutropenia was assumed to be limited to patients receiving dose dense chemotherapy, following the recommendation of international guidelines[69]. From those receiving conventional chemotherapy regimens with cycles every 3 weeks, primary prophylactic treatment was assumed not to take place, while secondary prophylaxis was estimated to be provided to 20% of these patients on conventional regimens. This was based on expert opinion. Secondary prophylaxis was assumed to start from the second cycle, since neutropenia appears to develop mostly over the first chemotherapy cycles[68].
7. Assumptions on the sick-leave entitlement were based on the mean gross salary for a working woman in Belgium (statistics of the FPS Economy), and the proportion of the salary paid for by the RIZIV/INAMI (http://www.riziv.be) in case of sick leave. At present in Belgium, patients are entitled to a monthly payment of 60% of their gross salary for the first year of sick leave. An assumption was made that a mean sick-leave period of six months (the first of which is paid by the employer), would apply for 62% (Eurostats - http://epp.eurostat.ec.europa.eu) of those patients receiving chemotherapy, aged less than 65[70], although this could depend on regimen and more specifically, on whether dose dense regimens are applied versus longer, conventional treatments.

All key assumptions required for our estimations were tested in a univariate sensitivity analysis.

6.3 Results

6.3.1 Target population – size and frequency of chemotherapy treatment

The overall size of the target population (ER+, LN-, HER2-) in Belgium for the year 2008 amounted to 3266 patients, but NPI data was available for 2980 of them. Patients with more than one invasive tumour were excluded from the data sample, resulting in a population of 2671. From these, 2613 were treated with breast conserving surgery or mastectomy and thus represented the baseline population used in our calculations. With regards to age, 62% of them were younger than 65.

Based on the data from the Belgian Cancer Registry, only 15% of these patients received adjuvant chemotherapy in 2008. Looking at NPI levels, it is clear that the higher the NPI score the higher the probability of receiving chemotherapy (see table 9), while from those patients with NPI<3,4 only 6% received chemotherapy.

Table 9 – Chemotherapy use in target population in Belgium

	Number with no adj. chemo (%)	Number with adj. chemo (%)	Total
NPI ≤3,4	1684 (94)	112 (6)	1796
3,4<NPI≤5, 4	545 (67)	268 (33)	813
NPI>5,4	2 (50)	2 (50)	4
Total	2231 (85)	382 (15)	2613

It should be highlighted that these figures are likely to underestimate the overall number of ER+, LN-, HER2- breast cancer patients in Belgium, since the percentage of missing data for the database of female breast cancer patients concerning for example the ER status by immunohistochemistry was about 11.3% in 2008. Nevertheless, it remains the best and most complete source of Belgian epidemiological data up to date.

Data on risk distribution were extracted directly from the independent economic evaluation performed by Ward et al.[5] for Oncotype DX and were based on TransATAC data. Overall, there were 64,59% out of 706 patients classified as low risk, 26,20% as intermediate and 9,21% as high risk.

Applying these proportions to the Belgian population of ER+, LN- and HER2- patients, the following results are obtained:

- 1688 patients overall would be classified as low risk according to Oncotype DX
- 685 patients in Belgium would be classified as intermediate risk
- 241 Belgian patients would be considered as high risk patients.

According to Paik et al,[62] patients classified as low risk (RS<18) derived a minimal benefit from chemotherapy treatment (relative risk 1.31; 95% CI 0.46 to 3.78). Patients with intermediate risk showed no large benefit from chemotherapy (relative risk 0,61; 95%CI 0,24-1,59; p=0,39), but the uncertainty in the estimate did not exclude a clinical important benefit. While it was those classified as high risk which derived a significant benefit from chemotherapy treatment (relative risk 0,26; 95% CI 0,13 to 0,53; p<0,001).

Thus, if the GEP test was used in the Belgian population described in the Belgian Cancer Registry data of 2008, 241 patients would have had the most benefit from chemotherapy. A difference when compared to those who received chemotherapy in 2008 of 141 patients. However, given the uncertainties regarding patients with an intermediate risk and the benefit that they could derive from chemotherapy treatment, an assumption was made that some of them would still receive chemotherapy. Fifteen per cent was taken as the base case assumption reflecting the overall chemotherapy rate in Belgium for the entire ER+, LN-, HER2- patient population. Adding up those at high risk and 15% of those at intermediate risk there would be approximately 39 patients less receiving chemotherapy when using the test.

The overall cost calculations for pharmacological chemotherapy treatment in Belgium are shown in table 10, while other costs linked to the administration of chemotherapy treatment are presented in table 11. Finally table 12 gives the total costs of chemotherapy depending on the chemotherapy regimen.

With regard to the cost of the chemotherapy regimen, the combination of EC+paclitaxel appears to be the most expensive primarily because of the 12 cycles of paclitaxel at a drug price per cycle of €349. On the other hand, the TC regimen is less expensive primarily because of the limited number of cycles (4), despite Docetaxel having a price of €420. Following the assumption of half of the population being treated with EC+paclitaxel and the other half with TC this would result in a mean pharmacological cost of chemotherapy of €3 435.

Regarding other costs linked to chemotherapy administration, the central venous access has a relatively high cost (€824) since this is assumed to be placed on every patient. However, this is a one-off cost and thus is not linked to the number of cycles pursued. The costs of the administration in the hospital remain limited given the lump sum of €132 per cycle. Similarly the conservative approach taken regarding blood tests make their weight on the overall cost calculations very low. Supporting care costs include both prophylactic treatment for nausea and vomiting and neutropenia as well as costs linked to the management of neutropenia, including hospitalisation costs.

The cost of managing AEs may appear low but responds to the focus of our calculations on the short term (early AEs only). These include hospitalisation and treatment costs.

Table 10 – Mean costs of pharmacological chemotherapy regimens in Belgium

EC+ taxol	Dose mg/sqm	Dose mg*/patient	Price/mg	Drug price/cycle	N. of cycles	Drug price/chemo
Epirubicin	90	158	0,60	95	4	378
Cyclophosphamide	600	1050	0,07	78	4	311
Paclitaxel	175	306	1,14	349	12	4190
TC	Dose mg/sqm	Dose mg*/patient	Price/mg	Drug price/cycle	N. of cycles	Drug price/chemo
Docetaxel	75	131	3,20	420	4	1680
Cyclophosphamide	600	1050	0,07	78	4	311

* *surface area for a mean body weight of 70kg: 1,75m2 assumed ([71])*

Table 11 – Mean non- pharmacological chemotherapy costs in Belgium

Regimen	Central venous access	Supporting care (prophylaxis and early AEs incl. hospitalisation)	Administration	Blood tests
EC+paclitaxel	824	2103	1589	264
TC	824	2542	530	88

Table 12 – Mean overall chemotherapy costs in Belgium depending on regimen

Regimen	Drug costs - chemo	Administration costs	Clinical biology	Supporting care (prophylaxis and early AEs incl. hospitalisation)	Sick leave costs
EC+paclitaxel	4879	1589	264	2103	3457
TC	1991	530	88	2542	3457

Adding all these costs gives an approximate cost for adjuvant chemotherapy regimens of €7 817 without considering sick leave costs and a cost of €11 274 when the latter are included. These costs appear to be lower but not far from those recently reported for other EU countries[35, 71]. Table 13 shows the overall budget calculations for the two main scenarios: testing all patients versus testing only those with an intermediate risk according to NPI scores (3,4<NPI>5,4).

Table 13 – Budgetary impact - GEP tests

N. of ER+, LN-, HER2- early BC patients	2613	2613
Patients tested	All	3,4<NPI>5,4 only
% of patients tested	100%	31%
Cost of test	3128	3128
Total costs of testing	8173464	2543064
N. of patients with chemo w/o test	382	382
N. of patients with chemo with test	343	343
N. of patients in which chemo is avoided	39	39
Mean cost of chemo	7817	7817
Total cost of chemo w/o test	2986071	2986071
Total cost of chemo with test	10854675	5224275
Incremental budget with test w/o sick leave	7868603	2238203
Mean cost of chemo with sick leave	11274	11274
Incremental budget with test with sick leave	7733775	2103375

Overall, the test would add costs of €7 868 603 without sick leave costs or €7 733 775 including sick leave costs if used in all ER+, LN-, HER2- versus €2 238 203 without sick leave costs or €2 103 375 with sick leave, if testing is limited to only those with 3,4<NPI >5,4.

6.3.2 Sensitivity analysis

A number of theoretical assumptions used in our calculations were based on expert opinion as opposed to hard evidence (e.g. the percentage of patients using each regimen, the proportion of patients on dose dense chemotherapy or the proportion on prophylaxis for neutropenia). More importantly it is difficult to predict the percentage of patients classified as intermediate risk likely to be treated with chemotherapy. These make our estimates subject to important degrees of uncertainly that were borne in mind in our sensitivity analysis.

The main results from the sensitivity analysis are displayed in Table 14.

Table 14 – Sensitivity analysis - scenarios 1: testing all patients and scenario 2: testing NPI>3,4 versus current situation

	Baseline value	Minimum value	Maximum value	Overall impact (with baseline values)	Overall impact (with minimum values)	Overall impact (with maximum values)
Current practice vs Scenario 1 – Testing all ER+, LN-, HER2-						
without sick leave	**Baseline value**	**Min value**	**Max value**	**Budget impact – baseline**	**Budget impact - min**	**Budget impact - max**
Proportion on EC+taxol (in %)	50	30	70	7868603	7897339	7839868
Proportion on dose dense regimens (in %)	EC+P :10 ; TC :40	0	50	7868603	7906108	7830767
Proportion with intermediate risk treated with chemotherapy (in %)	15	0	20	7868603	7063458	8142196
with sick leave costs						
Proportion on EC+taxol (in %)	50	30	70	7733775	7762510	7705039
Proportion on dose dense regimens (in %)	EC+P :10 ; TC :40	0	50	7733775	7771280	7695939
Proportion with intermediate risk treated with chemotherapy (in %)	15	0	20	7733775	6583818	8128368
Length of sick leave (in months)	5	4	12	7733775	7760741	7545015
Current practice vs Scenario 2 – Testing NPI>3,4 only						
Without sick leave	**Baseline value**	**Min value**	**Max value**	**Budget impact – baseline**	**Budget impact – min**	**Budget impact - max**
Proportion on EC+taxol (in %)	50	30	70	2238203	2266939	2209468
Proportion on dose dense regimens (in %)	EC+P :10 ; TC :40	0	50	2238203	2275708	2200367
Proportion with intermediate risk treated with chemotherapy (in %)	15	0	20	2238203	1440875	2511796
with sick leave costs						
Proportion on EC+taxol (in %)	50	30	70	2103375	2132110	2074639
Proportion on dose dense regimens (in %)	10	0	50	2103375	2140880	2065539
Proportion with intermediate risk treated with chemotherapy (in %)	15	0	20	2103375	953418	2497968
Length of sick leave (in months)	5	4	12	2103375	2130341	1914615

Conclusions

- **Main drivers of the budgetary impact of GEP and expanded IHC tests are, aside from the cost of the test, the target population for testing, the proportion of those tested who will ultimately receive chemotherapy and the cost of chemotherapy.**
- **Our estimates show an approximate mean cost of chemotherapy in Belgium of €7 817 without considering sick leave costs and of €11 274 when the latter are included**
- **The test would add costs of €7 868 603 without sick leave costs or €7 733 775 including sick leave, if used in all ER+, LN-, HER2- versus €2 238 203 without sick leave costs or €2 103 375 with sick leave, if testing is limited to 3,4<NPI >5,4.**
- **Our results are subject to important uncertainties, such as the porportion of patients classified as having an intermediate risk after being tested who would receive chemotherapy.**

7 DISCUSSION AND LIMITATIONS

7.1 Clinical validity and utility

7.1.1 The evidence

The present overview of systematic reviews on the effectiveness of gene expression profiling and expanded immunohistochemistry tests for early breast cancer shows that most evidence is available for Oncotype DX (RT-PCR) and MammaPrint (micro-array GEP). In general, the evidence is mainly limited to their clinical validity (i.e. prognostic ability), and no RCTs appear to be available yet. For several tests (e.g. Oncotype DX, MammaPrint, Mammostrat), the evidence supporting their prognostic ability is quite strong, but this only gives indirect information about the clinical utility of these tests. Direct evidence (e.g. test-and-treat RCTs, comparative observational studies) evaluating the effect of management strategies incorporating these tests on clinical outcomes (i.e. survival, recurrence, etc.) is generally lacking. Would the GRADE system have been used in this report to assign a level of evidence to the conclusions (see KCE processes, http://processbook.kce.fgov.be/node/51), the indirectness of the evidence concerning patient-important outcomes such as survival would have immediately led to a downgrading to low or very low level evidence, even though the level of prognostic evidence is high in itself. However, the GRADE Working Group has not yet developed a method to evaluate prognostic evidence, so no grading system was used in the present report. If a RCT would have been available, the GRADE methodology for therapeutic interventions would have been used.

7.1.2 Limitations

An important limitation of our review of the clinical literature is the pragmatic approach chosen by doing a review of reviews. The rationale behind this choice is that from a pre-assessment of the literature it was clear that on the one hand high-quality recent reviews are available and on the other hand the evidence is currently limited to observational studies. Important consequences of our approach are that interpretation bias (i.e. interpretation of data by other reviewers) is introduced and that the most recent observational studies are not included. However, our estimation is that these studies would only have a small impact on the present conclusions. For

example, experts consulted during the current project mentioned the study of Drukker et al.[72], being one of the few prospective studies on this topic. This observational study prospectively evaluated the clinical validity of MammaPrint in a cohort of 427 patients with T1-3N0M0 breast cancer. No significant difference was found in five-year distant-recurrence-free interval between systematically untreated patients with a concordant low risk assessment and patients with a MammaPrint low-risk result even with a high-risk assessment by Adjuvant! Online (95.3% vs. 98.4%, p=0.29). No data on clinical utility in terms of effect on clinical outcomes are available from this study. However, the authors reported that in the 70-gene signature low-risk group 15% (33/219) of the patients received adjuvant chemotherapy, versus 81% (169/208) in the high-risk group. A limitation of the study, which was acknowledged by the authors, is that the treatment decisions were based on the (restrictive) Dutch guidelines of 2004 and doctor's and patients' preferences. Equality of prognosis between groups that did or did not receive adjuvant chemotherapy could therefore not be guaranteed (selection bias).

An ideal design to evaluate GEP and/or expanded IHC tests would be a RCT where an eligible population of women with early breast cancer is randomized to a treatment arm where the results of the test are not taken into account and a treatment arm where the results are taken into account for management decisions. However, such trials are costly and require a large sample size [73]. At least three RCTs with a somewhat different but acceptable design are currently ongoing (TAILORx, MINDACT, GERICO 11), but their first results are not expected before the end of 2015. In the TAILORx study, women with node-negative, estrogen-receptor positive HER2-negative breast cancer meeting 'standard criteria' for adjuvant chemotherapy are evaluated with Oncotype DX. Those with Oncotype DX low risk receive endocrine therapy alone, those with Oncotype DX high risk receive chemotherapy and endocrine therapy, and those with an intermediate risk (i.e. score of 11-25) are randomized to endocrine therapy alone or a combination of adjuvant chemotherapy and endocrine therapy. Similarly, in the French GERICO 11 study, elderly women with node-negative, estrogen-receptor positive HER2-negative T1-3 breast cancer and a high risk based on the Genomic Grade are randomized to hormonal therapy or a combination of adjuvant chemotherapy and hormonal therapy. Finally, in the MINDACT study, women with node-negative or up to 3 positive nodes T1-3N0-1M0 breast cancer have their risk assessed by MammaPrint and a modified version of Adjuvant Online (which includes HER-2 status). Those with high risk by both risk assessment methods receive chemotherapy (and endocrine therapy if ER+) and those with low risk by both methods do not receive chemotherapy and receive endocrine therapy if ER+. Those with discordant results by both methods are randomized to method of risk assessment, i.e. to follow what MammaPrint indicates (high risk to receive chemotherapy and low risk not to receive) or follow what Adjuvant Online indicates (high risk to receive chemotherapy and low risk not to receive).

Once the results from these trials are made public, an update of this report would be recommended. The status of the present report should therefore be considered preliminary.

An additional limitation of our review was the fact that only one reviewer quality appraised the included reviews and performed the data extraction.

7.2 Cost-effectiveness

7.2.1 The evidence

Despite the lack of randomized controlled trials up to date, looking at the clinical utility of these tests, we decided to still review the economic evidence available up to date for two reasons. First, there is a need to highlight current data gaps for Belgium. For example, there are at present limited or no data on the baseline population receiving chemotherapy, the current risk stratification systems most commonly used and the proportion of patients for which risk is unclear under standard practice, as well as the proportion of patients who - following a testing strategy - would be treated with chemotherapy. Such data could be captured by means of a registry and would facilitate a better understanding of the role and value of these tests in the Belgian context.

Second, there is a growing body of cost-effectiveness studies in this field of high methodological quality that deserves some attention and interpretation. In this regard, there is a high level of consistency which appears to indicate that GEP or expanded IHC tests are likely to be cost-effective compared to standard practice and should help to reduce current chemotherapy levels by better targeting the treatment and limiting it to those most likely to benefit

from it. However, such results should be interpreted with great caution in view of the important limitations identified during our review.

7.2.2 Limitations

The main limitation relates to the need to combine and model multiple sources of data given the fact that there are no studies following patients from the time of testing to final health outcomes. Such combinations increase, as already mentioned on section 5 of our report, the uncertainty surrounding the model assumptions, since important factors which could play a role in the final results are not kept constant from one study to another (e.g. population characteristics, standard practice, etc). Furthermore, the importance of certain variables such as population characteristics or standard practices on the overall results makes the extrapolation and generalizability of those results to the Belgian context, not advisable.

The grey literature was excluded from our study.

7.3 Budgetary impact

7.3.1 Preliminary estimations

Our budgetary estimations show that limiting testing to a subpopulation of patients for which the risk of recurrence is unclear according to current methods of stratification (e.g. NPI score) is likely to be an economically more attractive option. Nevertheless, the introduction of these tests in clinical practice is not likely to bring in savings in the short term from a health insurance perspective. This is expected, given that budgetary impact models focus on the short term and, as such, rely completely on the cost of testing and chemotherapy spared. Thus, the cost of distant recurrence or long-term adverse events, important cost drivers in economic evaluations on cancer, are not included. A cost-effectiveness model taking into consideration these important factors would ideally be fed with data from the ongoing prospective RCTs on Oncotype DX, MammaPrint and Genomic Grade (TAILOR-X, MINDACT, and GERICO 11) combined with data from Belgian registries on patient characteristics and standard practices.

Our analysis is carried out from a healthcare payer perspective, although estimations taking into consideration sick-leave over the chemotherapy treatment period were also explored. If a societal perspective would have been preferred, full productivity costs would have been included, and the overall results would have been more positive towards testing, since following chemotherapy has been shown to be a crucial predictor of rates of return to work in primary breast cancer patients[74, 75]. On this same line, Broeckx et al.[76] found that productivity costs account for as much as 86% of the total mean costs of breast cancer in Belgium (€107 456 over a period of 6 years).

Our budget analysis focuses on Oncotype DX, given that it is the test for which there is up to date more evidence, but also because it is the most expensive test according to the literature reviewed. Bearing in mind these factors the illustrative calculations here included are likely to represent an overestimation of the real impact that the introduction of these tests are likely to have in Belgium.

7.3.2 Limitations

Our estimations are no exempt of limitations and in particular, there is a crucial factor that should be highlighted. This relates to patients classified as having an intermediate risk according to Oncotype DX. Although the clinical evidence from Paik et al[62]. does not appear to show a large benefit from treating this risk group with chemotherapy (relative risk 0,61; 95%CI 0,24-1,59; p=0,39), it is highly unlikely that none of them would be considered for such treatment. Furthermore, the low rates of chemotherapy, used for the overall target population, (i.e. 15% of ER+, LN-, HER2- early breast cancer patients in 2008 in Belgium), make the proportion of these intermediate risk patients to be treated with chemotherapy a determinant cost factor. If a large proportion of these patients was treated with chemotherapy, then the amount of chemotherapy offered overall may not diminish when compared to a strategy without GEP testing. Indeed, the chemotherapy treated population could even increase. This would be in contrast with the general findings from the published evidence in which a decrease in chemotherapy was consistently reported. In this regard, it should be stressed that the departing chemotherapy rates used in the economic evaluations summarised in section 5 for ER+, LN- populations were noticeably higher, (from 26%[54] to 69%[52]), than those shown in the data from the Belgian Cancer Registry, which makes reducing current chemotherapy rates in this country more challenging than in others where the approach appears to be more aggressive.

Despite the negative short term budgetary impact that an increase in the number of patients receiving chemotherapy could have, it is important to highlight that, if the tests helped to identify patients at high risk of recurrence and likely to benefit from chemotherapy treatment, who would otherwise not have been identified and treated, this could in turn still bring savings in the long-term as it can be illustrated with an example from the published literature in this topic with Holt et al. 2013 concluding that Oncotype DX is cost-effective but finding that its use was linked to an increase in terms of budgetary impact when compared to not testing patients.

Further limitations include the need to use expert opinion in order to populate our model and estimate the cost of these tests in the Belgian context. There are at present no real Belgian data available on specific chemotherapy regimens for this specific population, and according to the discussions held with the experts, clinical practice differs greatly from one centre to another. Thus, our calculations provide a very simplified vision of what truly happens in clinical practice. The data used to identify the size of the target population as well as the proportion of them currently receiving adjuvant chemotherapy dates from 2008 and these could have changed to a certain extent in the last years.

At present, there is no agreement either on the proportion of the target population for which it is difficult to make decisions regarding chemotherapy treatment. During our discussions with experts, a proportion of 10% was mentioned, but this was thought to differ from one centre to another depending on their current risk stratification systems or their chemotherapy treatment approach amongst other factors. In our estimations we used the NPI tool as an objective measure of patients with an intermediate risk of recurrence, but whether this is commonly used in Belgium and how much influence it has on current treatment decisions is to this date unknown.

All of the above mentioned limitations highlight the importance of considering our calculations as an approximation to the real impact that these tests are likely to have in budgetary terms. Nevertheless, they offer a good insight into the current data gaps that would still need to be filled to ensure the Belgian situation can be modelled and studied in enough detail. Those data gaps have been taken into consideration at the time of drafting our recommendations to policy makers, health care providers and researchers.

7.4 Other limitations

Overall, the rapid nature of this review has as a further implication, i.e. the limited consideration of the ethical aspects that the use of these tests may bring. The extent to which patients may be willing to accept the test results, as a fundamental part of treatment decisions, as well as the inequalities that could potentially arise if these tests are only available to patients who can afford them, are important factors that should be further explored and debated.

APPENDICES

APPENDIX 1. CLINICAL REVIEW

Appendix 1.1. Search Strategies

Appendix 1.1.1. Medline

1 exp Breast Neoplasms/ (217787)

2 exp mammary neoplasms/ (19431)

3 "Neoplasms, Ductal, Lobular, and Medullary"/ (68)

4 exp breast/ (31803)

5 exp neoplasms/ (2563894)

6 4 and 5 (17610)

7 (breast$ adj5 (neoplasm$ or cancer$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or dcis or ductal or infiltrat$ or intraductal$ or lobular or medullary)).mp. (260793)

8 (mammar$ adj5 (neoplasm$ or cancer$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or dcis or ductal or infiltrat$ or intraductal$ or lobular or medullary)).mp. (34739)

9 1 or 2 or 3 or 6 or 7 or 8 (281662)

10 MammaPrint.mp. (93)

11 70-gene.mp. (411)

12 gene70.mp. (0)

13 gene?seventy.mp. (1)

14 seventy?gene.mp. (0)

15 amsterdam profile.mp. (0)

16 Oncotype.mp. (194)

17 21-gene.mp. (432)

18 gene21.mp. (3)

19 gene?twentyone.mp. (0)

20 twentyone?gene.mp. (0)

21 GHI Recurrence score.mp. (0)

22 GHI-RS.mp. (3)

23 92-gene.mp. (27)

24 gene92.mp. (0)
25 gene?ninetytwo.mp. (0)
26 ninetytwo?gene.mp. (0)
27 (RT-PCR adj5 GHI Recurrence score).mp. (0)
28 (RT-PCR adj5 "21").mp. (193)
29 Randox.mp. (151)
30 Blueprint.mp. (2093)
31 80-gene.mp. (50)
32 gene80.mp. (0)
33 gene?eighty.mp. (0)
34 eighty?gene.mp. (0)
35 PAM50.mp. (51)
36 50-gene.mp. (81)
37 gene50.mp. (0)
38 gene?fifty.mp. (0)
39 fifty?gene.mp. (0)
40 breast bioclassifier.mp. (1)
41 Breast Cancer Index.mp. (21)
42 Breast cancer gene expression ratio.mp. (2)
43 2-gene.mp. (13253)
44 Two-gene-index.mp. (3)
45 2-gene-index.mp. (1)
46 Two?gene.mp. (0)
47 gene?two.mp. (1)
48 H?I.mp. (22827)
49 H:I.mp. (12166)
50 5-gene.mp. (1524)
51 gene5.mp. (3)
52 gene?five.mp. (1)
53 five?gene.mp. (1)
54 7-gene.mp. (734)
55 seven-gene.mp. (216)
56 gene7.mp. (2)
57 gene?seven.mp. (0)
58 Theros.mp. (8)
59 Biotheranostics.mp. (8)
60 HOXB13$.mp. (129)
61 homeobox?13$.mp. (0)
62 interleukin?17B$.mp. (0)
63 IL17BR.mp. (17)
64 mammostrat.mp. (8)
65 five-biomarker-assay.mp. (0)
66 IHC4.mp. (17)
67 NPI+.mp. (1328)
68 Nottingham prognostic index plus.mp. (1)
69 Nottingham prognostic index +.mp. (242)
70 MapQuant*.mp. (8)
71 97-gene.mp. (25)
72 BCA.mp. (1290)
73 BCI.mp. (1425)
74 Endopredict.mp. (10)
75 exp Gene Expression Profiling/ (82181)
76 GEP.mp. (862)
77 GGI.mp. (94)
78 genomic grade index.mp. (24)
79 Femtelle.mp. (1)
80 ("uPA" and "PAI-1").mp. (1162)
81 exp Immunohistochemistry/ (515886)
82 or/10-81 (636752)
83 9 and 82 (23973)

84 meta-analysis.mp,pt. or review.pt. or search:.tw. (2050983)

85 83 and 84 (1669)

Appendix 1.1.2. Embase

'breast cancer'/exp OR 'breast tumor'/de OR (breast*:ab,ti AND (neoplasm*:ab,ti OR cancer*:ab,ti OR tumor*:ab,ti OR tumour*:ab,ti OR carcinoma*:ab,ti OR adenocarcinoma*:ab,ti OR sarcoma*:ab,ti OR dcis:ab,ti OR ductal:ab,ti OR infiltrat*:ab,ti OR intraductal*:ab,ti OR lobular:ab,ti OR medullary:ab,ti)) OR (mammar*:ab,ti AND (neoplasm*:ab,ti OR cancer*:ab,ti OR tumor*:ab,ti OR tumour*:ab,ti OR carcinoma*:ab,ti OR adenocarcinoma*:ab,ti OR sarcoma*:ab,ti OR dcis:ab,ti OR ductal:ab,ti OR infiltrat*:ab,ti OR intraductal*:ab,ti OR lobular:ab,ti OR medullary:ab,ti)) AND (mammaprint:ab,ti OR '70 gene':ab,ti OR gene70:ab,ti OR (amsterdam:ab,ti AND profiel:ab,ti) OR oncotype:ab,ti OR '21 gene':ab,ti OR gene21:ab,ti OR ghi:ab,ti OR '92 gene':ab,ti OR gene92:ab,ti OR 'rt pcr':ab,ti OR randox:ab,ti OR blueprint:ab,ti OR '80 gene':ab,ti OR gene80:ab,ti OR pam50:ab,ti OR '50 gene':ab,ti OR gene50:ab,ti OR (breast:ab,ti AND bioclassifier:ab,ti) OR (breast:ab,ti AND cancer:ab,ti AND index:ab,ti) OR (breast:ab,ti AND cancer:ab,ti AND gene:ab,ti AND expression:ab,ti AND ratio:ab,ti) OR '2 gene':ab,ti OR '5 gene':ab,ti OR gene5:ab,ti OR '7 gene':ab,ti OR gene7:ab,ti OR theros:ab,ti OR biotheranostics:ab,ti OR hoxb13*:ab,ti OR homeobox*:ab,ti OR il17br:ab,ti OR mammostrat:ab,ti OR 'five biomarker assay':ab,ti OR ihc4:ab,ti OR npi:ab,ti OR (nottingham:ab,ti AND prognostic:ab,ti AND index:ab,ti) OR mapquant*:ab,ti OR '97 gene':ab,ti OR bca:ab,ti OR bci:ab,ti OR endopredict:ab,ti OR gep:ab,ti OR ggi:ab,ti OR (genomic:ab,ti AND grade:ab,ti AND index:ab,ti) OR femtelle:ab,ti OR (upa:ab,ti AND 'pai 1':ab,ti) OR 'gene expression profiling'/exp OR 'immunohistochemistry'/exp) AND ([cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim) AND ([article]/lim OR [article in press]/lim OR [review]/lim)

Appendix 1.1.3. Cochrane Library

#1 MeSH descriptor: [Breast Neoplasms] 1 tree(s) exploded

#2 MeSH descriptor: [Mammary Neoplasms, Animal] 1 tree(s) exploded

#3 MeSH descriptor: [Neoplasms, Ductal, Lobular, and Medullary] 1 tree(s) exploded

#4 MeSH descriptor: [Breast] 1 tree(s) exploded

#5 MeSH descriptor: [Neoplasms] 1 tree(s) exploded

#6 #4 and #5

#7 (breast* and (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)):ti,ab

#8 (mammar* and (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or adenocarcinoma* or sarcoma* or dcis or ductal or infiltrat* or intraductal* or lobular or medullary)):ti,ab

#9 #1 or #2 or #3 or #6 or #7 or #8

#10 (MammaPrint or 70-gene or gene70 or (amsterdam and profiel) or oncotype or 21-gene or gene21 or GHI or 92-gene or gene92 or RT-PCR or randox or blueprint or 80-gene or gene80 or PAM50 or 50-gene or gene50 or (breast and bioclassifier) or (breast and cancer and index) or (breast and cancer and gene and expression and ratio) or 2-gene or 5-gene or gene5 or 7-gene or gene7 or theros or biotheranostics or HOXB13* or homeobox* or IL17BR or mammostrat or five-biomarker-assay or IHC4 or NPI or (Nottingham and prognostic and index) or MapQuant* or 97-gene or BCA or BCI or endopredict or GEP or GGI or (genomic and grade and index) or femtelle or (uPA and PAI-1)):ti,ab

#11 (gene and expression and profiling):ti,ab

#12 MeSH descriptor: [Immunohistochemistry] 1 tree(s) exploded

#13 #10 or #11 or #12

#14 #9 and #13

Appendix 1.2. Quality appraisal of included systematic reviews

Author	1. Was an 'a priori' design provided?		2. Was there duplicate study selection and data extraction?		3. Was a comprehensive literature search performed?		4. Was the status of publication (i.e. grey literature) NOT used as an		5. Was a list of studies (included and excluded) provided?		6. Were the characteristics of the included studies provided?		7. Was the scientific quality of the included studies assessed and documented?		8. Was the scientific quality of the included studies used appropriately in formulating		9. Were the methods used to combine the findings of studies appropriate?		10. Was the likelihood of publication bias assessed?		11. Was the conflict of interest stated?	
Arpino 2013	N	No clear protocol, no clear inclusion criteria	?	Not reported	Y	Medline (1966-2012), Cancerlit (1966-2012), and Embase (1990-2012); ASCO abstracts; references of reviews and articles Key words provided	N	English only; no grey literature	N	Not provided	N	No tables provided	N	Partly assessed, but not documented	N		NA		N		N	
Carlson 2013	?	Unclear if established before conduct of review	?	Unclear for selection, yes for data extraction	Y	Medline, Embase; ASCO & SABCS abstracts; references Key words provided	N	English only; no grey literature	N	No list of excluded studies	Y		N		N		Y		Y		N	
Kelly 2010	?	Unclear if established before conduct of review	?	Not reported	Y	Medline, Embase; ASCO & SABCS abstracts; references Key words provided	N	Unclear language restrictions; no grey literature	N	Not provided	N	No tables provided	N		N		NA		N		N	
Kuderer 2009	?	Unclear if established before conduct of review	?	Unclear for selection, yes for data extraction	Y	Medline, Embase, Cochrane Library, DARE; references Key words provided	N	No language restriction, but no grey literature searched	N	No list of excluded studies	Y	Although unclear which test was used exactly	N		N		Y		Y		N	
Lyman 2006	?	Unclear if established before conduct of review	?	Unclear for selection, yes for data extraction	Y	Medline, Embase, Cochrane Library, DARE, Dissertation Abstracts; references Key words provided	Y		N	No list of excluded studies	Y		N		N		Y		Y		N	
Marchionni 2008	Y	See full AHRQ report	Y		Y	Medline, Embase, CDSR, CENTRAL, Cinahl; manufacturers; hand searching Key words provided	N	English only; unpublished studies were searched	Y	See full AHRQ report	Y		Y		Y		NA		N		Y	
Smartt 2010	Y		?	Not reported	Y	Medline, Embase, CDSR, CENTRAL, Cinahl; conference abstracts Key words provided	N	English only; no grey literature	N	No list of excluded studies	Y		Y		Y		NA		N		N	
Ward 2013	Y		Y		Y	Medline, Embase, CENTRAL, CDSR, DARE, HTA, BIOSIS previews, Web of Science; references; manufacturers; experts	Y		Y		Y		Y		Y		NA		N		N	
Zhao 2014	?	Unclear if established before conduct of review	Y		Y	Medline, Embase, CNKI; references Key words provided	N	No language restriction, but no grey literature searched	N	No list of excluded studies	Y		N		N		Y		Y		N	
MAS 2010	Y		?	Yes for selection, unclear for data extraction	Y	Medline, Embase, Cinahl, Cochrane Library, INAHTA; references; experts	N	English only; no grey literature	N	No list of excluded studies	Y		Y		Y		NA		N		N	
EUnetHTA 2013	Y		Y		Y	Medline, Embase, Cochrane Library, Cinahl, CRD databases; trial registries; reference lists	N	English only	Y	No list of excluded studies, but information on some excluded studies in Appendix 6	Y		N	No RCTs or prospective cohort studies were found, and the other study designs were not quality apparaised	Y		NA		N	No formal assessment of publication bias, although publication bias was suspected	Y	
MoH Malaysia 2008	?	Unclear if established before conduct of review	?	Not reported	Y	Medline, Cinahl, CDSR, HTA, Horizon scanning databases, FDA website, Google; references	?	Unclear language restrictions; grey literature was searched	N		N	No tables provided	N	Not documented	Y		NA		N		N	
MoH Malaysia 2011	?	Unclear if established before conduct of review	?	Not reported	Y	Medline, CDSR, HTA, CENTRAL, National Horizon Scanning, INAHTA, ASERNIP-S, CADTH, FDA website; references	Y		N		N	No tables provided	N	Not documented	Y		NA		N		N	

Appendix 1.3. Evidence tables of included systematic reviews

Arpino 2013	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	Supported by the Breast Journal Club (BJC) Italian Team No conflicts of interest declared
• **Search date**	2012
• **Searched databases**	Medline (1966-2012), Cancerlit (1966-2012), and Embase (1990-2012); ASCO abstracts; references of reviews and articles
• **Included study designs**	All
• **Number of included studies**	Not reported
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Not specifically reported; breast cancer
• **Exclusion criteria**	Not reported
• **Patient & disease characteristics**	Not reported
Interventions	
• **Gene expression profiling**	MammaPrint, Oncotype DX, Theros, MapQuant DX, PAM50, Mammostrat, IHC4
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	Narrative discussion of results Quality partly assessed, but not formally documented No list of included studies; no table with characteristics

Carlson 2013	
Methods	
• **Design**	Systematic review + meta-analysis
• **Source of funding and competing interest**	Carlson JJ is supported by the Agency for Healthcare Research and Quality Mentored Clinical Scientists Comparative Effectiveness Development Program (K12) at University of Washington (HS019482). Roth JA is supported by a National Institute on Aging T32 (AG027677).
• **Search date**	March 2012
• **Searched databases**	Medline, Embase (6/2005 – 3/2012); abstracts of ASCO (6/2005 – 3/2012) and SABCS (1/2008 – 3/2012); references
• **Included study designs**	All
• **Number of included studies**	N=23
• **Statistical analysis**	Pooled analyses using RevMan 5.0, not otherwise specified; random effects model in case of $p<0.10$ for I^2
Patient characteristics	
• **Eligibility criteria**	• Estrogen-receptor-positive, lymph-node-negative, and early-stage breast cancer

Carlson 2013	
	• Reported use of Oncotype DX recurrence score to inform clinical adjuvant chemotherapy decisions • Reported outcomes of interest, including: distribution of Oncotype DX recurrence score; impact of Oncotype DX on adjuvant chemotherapy recommendations; impact of Oncotype DX on adjuvant chemotherapy use; and proportion of patients following the treatment suggested by the Oncotype DX recurrence score
• **Patient & disease characteristics**	• Sample size: range 29 – 924
Interventions	
• **Oncotype DX**	
Results	
• **Mean proportion classified as low, intermediate, and high risk for recurrence (21 studies, N=4 156)**	• Low: 48.8% • Intermediate: 39.0% • High: 12.2%
• **Proportion of patients with changed management (8 studies, N=1 437)**	• 33.4%

Carlson 2013	
• **Proportion of patients receiving adjuvant chemotherapy (14 studies, N=3 104)**	• Low: 5.8% • Intermediate: 37.4% • High: 83.4%
• **Patients following the treatment suggested by Oncotype DX result (low versus high) (14 studies, N=1 900)**	• RR=1.07 (95%CI 1.01-1.14)
Limitations and other comments	
• **Limitations**	No quality assessment of included studies

Kelly 2010	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	One author with honoraria from Pfizer, AstraZeneca, Roche, Novartis
• **Search date**	Not reported
• **Searched databases**	Medline, Embase; abstracts of ASCO and SABCS; references
• **Included study designs**	Prognostic and validation studies
• **Number of included studies**	Not reported
• **Statistical analysis**	Not applicable

Kelly 2010	
Patient characteristics	
• **Eligibility criteria**	Not specifically reported; breast cancer
• **Exclusion criteria**	Not reported
• **Patient & disease characteristics**	Not reported
Interventions	
• **Oncotype DX**	
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	Narrative discussion of results No quality assessment of included studies No list of included studies; no table with characteristics

Kuderer 2009	
Methods	
• **Design**	Systematic review + meta-analysis
• **Source of funding and competing interest**	No conflicts of interest declared
• **Search date**	Not reported
• **Searched databases**	Medline, Embase, Cochrane Library, DARE; references
• **Included study designs**	Validation cohort studies

Kuderer 2009	
• **Number of included studies**	N=11
• **Statistical analysis**	Random effects model (DerSimonian and Laird)
Patient characteristics	
• **Eligibility criteria**	Original studies of patients with early-stage breast cancer
• **Exclusion criteria**	Duplicate studies based on the same group of patients
• **Patient & disease characteristics**	• Sample size: range 20 – 668 • 54% high risk • 23% distant breast recurrence • Recurrence rate: high risk 35% (95%CI 26-42), low risk 9% (7-12)
Interventions	
• **Gene expression profiling**	No further specification
Results for distant recurrence-free survival	
• **Sensitivity and specificity**	• Pooled Se: 82.4% (95%CI 76.1-88.7) • Pooled Sp: 53.3% (95%CI 43.9-62.7) • False negatives (low-risk and recurrence): 17.8% (14.2-22.1) • False positives (high-risk and no recurrence): 46.1% (43.2-49.0)
• **PPV and NPV**	• Pooled PPV: 42.5% (95%CI 32.2-52.7) • Pooled NPV: 89.8% (95%CI 86.7-93.0)
• **Diagnostic odds ratio**	• Pooled LR+: 1.78 (95%CI 1.47-2.16) • Pooled LR-: 0.35 (95%CI 0.26-0.47)

Kuderer 2009	
	• DOR: 6.44 (95%CI 3.42-12.08) • AUC: 0.8611
Limitations and other comments	
• **Limitations**	No quality assessment of included studies

Lyman 2006	
Methods	
• **Design**	Systematic review + meta-analysis
• **Source of funding and competing interest**	Not reported
• **Search date**	September 2006
• **Searched databases**	Medline, Embase, Cochrane Library, DARE, Dissertation Abstracts; references
• **Included study designs**	Validation cohort studies
• **Number of included studies**	N=17
• **Statistical analysis**	Random effects model (DerSimonian and Laird)
Patient characteristics	
• **Eligibility criteria**	Original studies of patients with early-stage breast cancer
• **Exclusion criteria**	Duplicate studies based on the same group of patients
• **Patient & disease characteristics**	• Sample size: range 20 – 668 • Median follow-up: 23.5 months

Lyman 2006	
	• 52.6% high risk • 20.5% distant breast recurrence • Recurrence rate: high risk 31.2% (95%CI 29-34), low risk 8.5% (7-10)
Interventions	
• **Gene expression profiling**	MammaPrint (5 studies), Oncotype DX (5 studies), 41-gene (2 studies), 23-gene (1 study), 76-gene (2 studies), 64-gene (1 study)
Results for distant recurrence-free survival	
• **Sensitivity and specificity**	• Pooled Se: 80.6% (95%CI 76-86) • Pooled Sp: 53.6% (95%CI 47-60) • False negatives (low-risk and recurrence): 19.4% (14.5-24.4) • False positives (high-risk and no recurrence): 46.4% (39.9-52.9)
• **PPV and NPV**	• Pooled PPV: 37.7% (95%CI 30-46) • Pooled NPV: 92% (95%CI 89-94)
• **Diagnostic odds ratio**	• Pooled LR+: 1.78 (95%CI 1.5-2.11) • Pooled LR-: 0.38 (95%CI 0.29-0.49) • DOR: 5.53 (95%CI 3.57-8.57) • AUC: 0.775
Limitations and other comments	
• **Limitations**	No quality assessment of included studies

Marchionni 2008	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	Funded under contract no. 290-02-0018 from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services No conflicts of interest declared
• **Search date**	July 2007
• **Searched databases**	Medline, Embase, CDSR, CENTRAL, Cinahl; manufacturers
• **Included study designs**	Prognostic studies
• **Number of included studies**	Oncotype DX: 13 studies; MammaPrint: 8 studies; H/I: 7 studies
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Patients with early-stage breast cancer; English-language study with original data or original data analysis
• **Patient & disease characteristics**	Not reported
Interventions	
• **Gene expression profiling**	MammaPrint, Oncotype DX, H/I
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	Only studies in English

Smartt 2010	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	Funded by the New Zealand Ministry of Health
• **Search date**	December 2009 (update of Marchionni 2008)
• **Searched databases**	Medline, Embase, CDSR, CENTRAL, Cinahl; conference abstracts
• **Included study designs**	Specify the type of study: RCT, CCT, case control, case series
• **Number of included studies**	N=12 in addition to Marchionni 2008
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Patients with early-stage breast cancer; English-language study with original data or original data analysis
• **Patient & disease characteristics**	Not reported
Interventions	
• **Gene expression profiling**	MammaPrint, Oncotype DX, H/I
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	Duplicate selection and data extraction not reported English literature only

Ward 2013	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	Commissioned and funded by the HTA programme on behalf of NICE as project number 10/125/01 No CoI declared
• **Search date**	May 2011
• **Searched databases**	Medline, Embase, CENTRAL, CDSR, DARE, HTA, BIOSIS previews, Web of Science; references; manufacturers; experts
• **Included study designs**	All
• **Number of included studies**	N=32
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	All people diagnosed with early invasive breast cancer being treated in the adjuvant setting
• **Exclusion criteria**	People diagnosed with early invasive breast cancer being treated in the neoadjuvant setting
Interventions	
• **Gene expression profiling**	MammaPrint, Oncotype DX, H/I, BluePrint, PAM50, Randox BCA, Mammostrat, IHC4, NPI+
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	High-quality review

Zhao 2014	
Methods	
• **Design**	Systematic review + meta-analysis
• **Source of funding and competing interest**	Specify the source of funding: public research funds, government, not governmental organization, healthcare industry or other (give name of organization or corporation) and the presence of declaration of interest (stated/not stated and specify if any)
• **Search date**	October 2013
• **Searched databases**	Medline, Embase, China National Knowledge Infrastructure database; references
• **Included study designs**	Cohort or case-control studies
• **Number of included studies**	N=10
• **Statistical analysis**	Fixed and random effects models (Review Manager v5.2)
Patient characteristics	
• **Eligibility criteria**	Patients with breast cancer treated with tamoxifen; no other systematic adjuvant therapy except for tamoxifen
• **Patient & disease characteristics**	• Node positive: range 0-100%
Interventions	
• **Gene expression profiling**	H/I
Results	
• **Recurrence-free survival**	Women with higher HOXB13-to-IL17BR expression ratio had significantly worse RFS: HR=1.47; 95%CI 1.17-1.84; p=0.001

Zhao 2014	
	Node-negative subgroup: HR=1.66; 95%CI 1.38-2.01; p<0.00001
• **Overall survival**	Women with higher HOXB13-to-IL17BR expression ratio had non-significantly worse OS: HR=1.32; 95%CI, 0.97-1.80; p=0.08 Node-negative subgroup: HR=1.93; 95%CI 1.15-3.23; p=0.01
Limitations and other comments	
• **Limitations**	No quality appraisal of included studies

MAS 2010	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	No conflicts of interest to declare
• **Search date**	March 2010 (update of Marchionni 2008)
• **Searched databases**	Medline, Embase, Cinahl, Cochrane Library, INAHTA; references; contact with experts
• **Included study designs**	All
• **Number of included studies**	N=26
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Women with newly diagnosed early stage (stage I–IIIa according to The American Joint Committee on Cancer staging system) invasive breast cancer that is estrogen-receptor (ER) positive and/or progesterone-receptor (PR) positive
Interventions	
• **Gene expression profiling**	Oncotype DX
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	English literature only

MoH Malaysia 2008	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	Ministry of Health Malaysia
• **Search date**	2008
• **Searched databases**	Medline, Cinahl, CDSR, HTA, Horizon scanning databases, FDA website, Google; references
• **Included study designs**	All
• **Number of included studies**	Unclear
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Women with breast cancer
Interventions	
• **Gene expression profiling**	MammaPrint
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	Poorly described methodology No quality assessment of included studies

MoH Malaysia 2011	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	Ministry of Health Malaysia
• **Search date**	2011
• **Searched databases**	Medline, CDSR, HTA, CENTRAL, National Horizon Scanning, INAHTA, ASERNIP-S, CADTH, FDA website; references
• **Included study designs**	All
• **Number of included studies**	Unclear
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Women with breast cancer
Interventions	
• **Gene expression profiling**	MammaPrint
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	Poorly described methodology No quality assessment of included studies

EUNetHTA 2013	
Methods	
• **Design**	Systematic review
• **Source of funding and competing interest**	EUNetHTA
• **Search date**	October 2011 (update of search in December 2011)
• **Searched databases**	Medline, Embase, Cochrane Library, Cinahl, CRD databases; trial registries; reference lists
• **Included study designs**	All, except studies limited to the prognostic/predictive accuracy of the test
• **Number of included studies**	N=15 (on effectiveness)
• **Statistical analysis**	Not applicable
Patient characteristics	
• **Eligibility criteria**	Women with early invasive breast cancer
Interventions	
• **Gene expression profiling**	MammaPrint, Oncotype DX, uPA/PAI-1
Results	
Narratively presented:	See evidence report
Limitations and other comments	
• **Limitations**	English literature only No quality appraisal of included studies

APPENDIX 2. ECONOMIC REVIEW

Appendix 2.1. Search strategies

Appendix 2.1.1. MEDLINE

---

1 exp Breast Neoplasms/ (222893)
2 exp Gene Expression Profiling/ (84678)
3 exp Immunohistochemistry/ (522929)
4 2 or 3 (601978)
5 1 and 4 (17831)
6 mammaprint.mp. (93)
7 oncotype.mp. (203)
8 (21-gene or 21 gene).mp. (443)
9 (70-gene or 70 gene).mp. (418)
10 randox.mp. (155)
11 blueprint.mp. (2156)
12 (80-gene or 80 gene).mp. (51)
13 (PAM50 or PAM 50).mp. (62)
14 ("breast cancer index" or BCI).mp. (1773)
15 (7-gene or 7 gene).mp. (746)
16 (2-gene or 2 gene).mp. (13443)
17 (5-gene or 5 gene).mp. (1548)
18 mammostrat.mp. (8)
19 IHC4.mp. (20)
20 ("nottingham prognostic index plus" or NPI+).mp. (1377)
21 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 (22218)
22 1 and 21 (1212)
23 5 or 22 (18591)
24 exp Economics/ (499668)

25 exp Health Care Costs/ (47749)

26 exp Economics, Medical/ (13664)

27 (cost and cost analysis).mp. (44519)

28 exp Economics, Pharmaceutical/ or exp Economics, Medical/ or Economics/ or exp Economics, Hospital/ or exp Economics, Nursing/ (66087)

29 exp "Value of Life"/ (5930)

30 ("cost effectiveness" or cost-effectiveness).mp. (33416)

31 exp Quality-Adjusted Life Years/ (7314)

32 (cost-utility or "cost utility").mp. (2339)

33 exp Health Expenditures/ (15887)

34 budget*.mp. (23201)

35 (price or prices or pricing).mp. (22605)

36 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 (533228)

37 23 and 36 (125)

Appendix 2.1.2. EMBASE

(((((('breast tumor'/exp and [embase]/lim) and (('gene expression profiling'/exp and [embase]/lim)

or ('immunohistochemistry'/exp and [embase]/lim))) or (('breast tumor'/exp and [embase]/lim)

and (('dna microarray'/exp and [embase]/lim) or (mammaprint:ab,ti and [embase]/lim)

or (oncotype:ab,ti and [embase]/lim) or ('21 gene':ab,ti and [embase]/lim)

or ('70 gene':ab,ti and [embase]/lim) or (randox:ab,ti and [embase]/lim)

or (blueprint:ab,ti and [embase]/lim) or ('80 gene':ab,ti and [embase]/lim)

or (pam50:ab,ti and [embase]/lim) or ('breast'/exp and 'cancer'/exp and index:ab,ti

and [embase]/lim) or (bci:ab,ti and [embase]/lim) or ('7 gene':ab,ti and [embase]/lim)

or ('2 gene':ab,ti and [embase]/lim) or ('5 gene':ab,ti and [embase]/lim)

or (mammostrat:ab,ti and [embase]/lim) or (ihc4:ab,ti and [embase]/lim) or

(npiplus:ab,ti and [embase]/lim) or (npi+:ab,ti and [embase]/lim) or

(nottingham and prognostic and index and plus:ab,ti and [embase]/lim)))) and

(('economics'/exp and [embase]/lim) or ('health care cost'/exp and [embase]/lim) or

('health economics'/exp and [embase]/lim) or ('health care financing'/exp and [embase]/lim)

or ('cost benefit analysis'/exp and [embase]/lim) or ('cost effectiveness analysis'/exp and

[embase]/lim) or ('cost of illness'/exp and [embase]/lim)

or ('cost control'/exp and [embase]/lim) or ('hospital cost'/exp and [embase]/lim)

or ('cost utility analysis'/exp and [embase]/lim) or ('cost minimization analysis'/exp

and [embase]/lim) or (price or prices or pricing and [embase]/lim) or (budget* and [embase]/lim)))

and ('article'/it or 'review'/it))

Appendix 2.1.3. EconLit

1 (breast adj neoplasm*).mp. [mp=heading words, abstract, title, country as subject] (0)

2 (breast adj cancer).mp. [mp=heading words, abstract, title, country as subject] (214)

3 (breast and neoplasm).mp. [mp=heading words, abstract, title, country as subject] (1)

4 (gene adj expression adj profiling).mp. [mp=heading words, abstract, title, country as subject] (1)

5 (gene adj expression).mp. [mp=heading words, abstract, title, country as subject] (69)

6 test.mp. [mp=heading words, abstract, title, country as subject] (39713)

7 (inmunohistochemistry or IHC).mp. [mp=heading words, abstract, title, country as subject] (3)

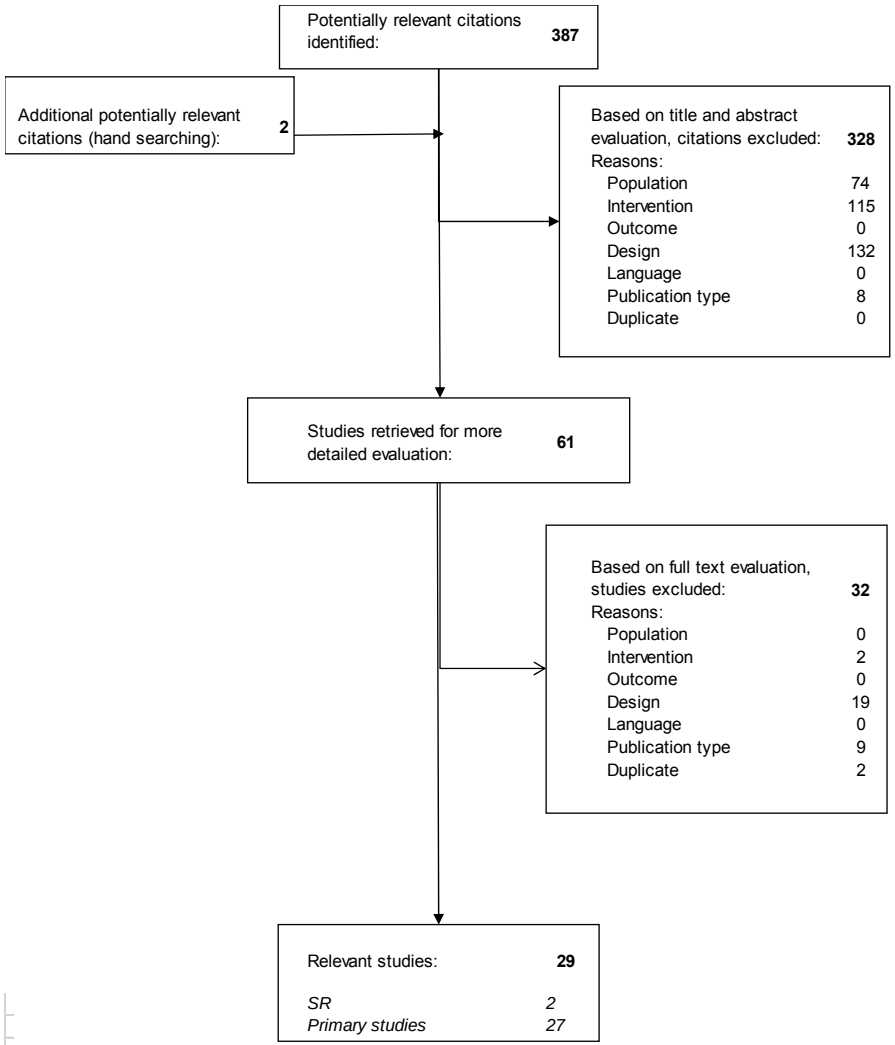
8 5 and 6 (11)
9 1 or 2 or 3 (214)
10 4 or 7 or 8 (15)
11 9 and 10 (2)

Appendix 2.1.4. CRD NHS databases

Search	Hits	
1	MeSH DESCRIPTOR Gene Expression Profiling EXPLODE ALL TREES	48
2	MeSH DESCRIPTOR Breast Neoplasms EXPLODE ALL TREES	1582
3	#1 AND #2	32

Limit to NHS EED & NHS HTA – 28 (exclusion of DARE)

Appendix 2.2. Flow chart selection of Economic Evaluations



Appendix 2.3. Template table for data extraction – Economic evaluations

1	Title
2	Reference (including all authors)
3	Conflict of interest and/or study funding
4	Country
5	Study question – clear and complete including statement of problem
6	Need for modelling – justified
7	Type of analysis (analytic technique)
8	Specific model design –complete description
9	Population – full description
10	Intervention
11	Comparator
12	Time horizon – appropriate and justified
13	Discount rate – inclusion and justification of rates used
14	Perspective
15	Costs
	• Cost items included
	• Measurement of resource use
	• Valuation of resource use
	• Data sources and references
	• Currency and cost year
16	Outcomes
	• Endpoints taken into account and/or health states
	• Valuation of health states
	• Treatment effect and Extrapolation
	• Utility assessment (Quality of Life)
	• Data sources for outcomes and references – values used in base case scenario and justification
17	Uncertainty
	• Scenario analysis
	• Sensitivity analysis – univariate and or multidimensional – ranges of values used and justification
18	Assumptions and discussion regarding their impact on the results
19	Results
	• Cost-effectiveness and/or cost-utility (base case)
	• Scenario analysis
	• Sensitivity analysis
20	Conclusions and applicability
21	Remarks – ongoing research which could affect results

APPENDIX 3. POINTS RAISED BY DR FATIMA CARDOSO

Point 1. The validator did not agree with the choice of methodology chosen for the review of the clinical evidence (i.e. review of reviews) and thought such a decision could have had an impact on the overall conclusions of our report.

Answer to Point 1: The rationale behind the pragmatic approach chosen by doing a review of reviews is that from a pre-assessment of the literature it was clear that on the one hand high-quality recent reviews were available and on the other hand the evidence is currently limited to observational studies. Important consequences of our approach are that interpretation bias (i.e. interpretation of data by other reviewers) is introduced and that the most recent observational studies are not included. However, our estimation is that these studies would only have a small impact on the present conclusions.

Point 2. The validator expressed that some European countries have performed health economic evaluations on MammaPrint and Oncotype, although these may not have been published. She added that there are already real-life data available from countries publicly reimbursing these tests, such as Spain which she would have liked to see included in our review.

Answer to Point 2: The authors of this report believe that, although such data would have been a nice-to-have, its inclusion in our report would have required contacting governmental agencies to obtain such data which would have extended our timelines significantly. Given the rapid nature of our review and the fact that most often governmental data focuses primarily on the cost side (budget impact) and that it relies on standards of practice, which differ from one country to another, we decided not to engage in such an exercise and limit the scope of our economic review (as explained in page 29 of the report), to published literature, and more specifically, to published full economic evaluations (looking at both costs and outcomes). The grey literature is thus excluded from our review and this is recognised as a limitation of the study in our general discussion.

Point 3. The validator shared with us a presentation by the Dirección General de Farmacia y Productos Sanitarios of the region of Valencia in Spain as an example of what could be implemented in terms of a pilot study in Belgium. On that basis, she made a specific suggestion to restrict the target patient population for the tests, during the pilot study, by following a similar approach to that used in the Spanish region. She mentioned all LN-, ER+/HER2- patients represent about two thirds of the breast cancer patients and in many of these cases, clinicians know how to decide (less likely to benefit from the tests).

Answer to Point 3: Although indeed restricting the use of these tests as much as possible would diminish the potential budgetary impact they would have on our system, we do not have at present any data on additional patient/tumour characteristics, that would allow us to objectively select those most likely to benefit from the test. In the Valencian study certain parameters were used, e.g. PgR negative <1%; ER + but only low levels (10-60%); or Ki67 intermediate (13-30%), but mimicking those for Belgium, without first having a good understanding of the characteristics of the patients for which chemotherapy decisions are not clear in our country, the current risk stratification tools/processes as well as current chemotherapy prescription patters would give us a very limited view which may not be of great help at the time of performing a full cost-effectiveness evaluation on these tests in Belgium once the RCTs become available. Thus, although the variables used in the Valencian study could be captured and explored during the pilot study, the authors of this review preferred to suggest a less restrictive approach for the pilot in Belgium in order to encourage a greater understanding of current risk stratification practices and chemotherapy treatment decisions.

Point 4. The validator expressed a major concern for the use of OncotypeDX in the budgetary estimations for Belgium since there is at present no good data on what should be done with the intermediate risk patients (question likely to be clarified in the ongoing TAILORX trial), and thus, assumptions were required in this regard. She suggested to add calculations also for either MammaPrint or Prosigna (PAM-50), which have a dichotomous results (high/low) and are also the only two tests with FDA clearance up to date.

Answers to Point 4: There were no full economic evaluations identified in our review for PAM50 and thus, there was no data to be able to replicate similar calculations for this test to those performed for OncotypeDX.

The budget calculations for OncotypeDX are clearly presented as an illustrative example and the areas where assumptions were required were tested by means of a simple, univariate sensitivity analysis. The choice for limiting the estimations to Oncotype DX was made, primarily because of two reasons:

First, according to the published economic studies on OncotypeDX, there appears to be some consensus regarding the best data source to use as clinical input (Paik et al. 2006), for the risk classification of patients post-test. For MammaPrint this is not the case, with high variation in the sources used in the different economic evaluations reviewed.

Second, the sources used in the cost effectiveness studies published up to date for MammaPrint show an important proportion of patients being classified as “high risk” (e.g. Bueno de Mesquita – 51,07%; Van de Vijver – 61%). Although the data shared with us from the Spanish example shows a lower proportion of patients being classified as “high risk” with MammaPrint, these still reach 33,14% of patients. Departing from the low proportion of adjuvant chemotherapy rates given to the target population here in Belgium pre-test (15% according to 2008 data from the cancer registry), and modelling the budgetary impact for MammaPrint of treating all high risk patients would have resulted in a much higher financial impact than the one presented in our estimations, even when limiting testing to 3,4< NPI<5,4 patients as per the OncotypeDX example. Given the lack of consensus on the best evidence source for MammaPrint on patient reclassification and the fact that the ongoing MINDACT trial should offer some valuable answers on this regard, the authors of this review preferred not to present a budget calculation for MammaPrint at this point in time.

A further reason to choose Oncotype DX for our example (clearly stated in the text) is that Oncotype is the most expensive test and as a consequence, the use of other cheaper tests, (assuming similar outcomes and predictive accuracy) is likely to result in a lower budget impact in general. Our current estimations offer a conservative approach, as recommended in economic evaluations or budgetary impact analyses guidelines.

Nevertheless the intention of our calculations was not to recommend to limit use/funding to Oncotype DX, given the limited evidence comparing tests between themselves, and the opinion from the Belgian experts consulted through our project, who believed there is not much difference between one gene expression profiling test or another. Our recommendation is instead to leave the choice of the test open for the purpose of the pilot study in order to capture as much data as possible and encourage competition.

The specific proportions of high risk and low risk patients with MammaPrint as well as the intention to treat with chemo pre-test and post-test should be clarified via the recommended pilot study, which would in turn facilitate the development of a full HTA (with a budgetary impact and cost-effectiveness model presenting real Belgian data), once the results from the ongoing trials are published. The completion of such full HTA combining data from the RCTs (once available) with the data captured via our pilot Belgian study data is another recommendation coming out of our report.

This report is intended to offer a first view of the area in order to highlight the data gaps still existing at present in Belgium, some of which could be appropriately addressed via the recommended pilot study, while awaiting the RCT results. Ignoring these gaps would mean that even when the results from the RCTs are published, modelling the economic situation in Belgium would be difficult and too many assumptions would need to be made, or more time would need to be spent in registering the relevant data at that point in time.

REFERENCES

1. Belgian Cancer Registry. Cancer Incidence in Belgium 2008. 2011. Available from: http://www.kankerregister.org/media/docs/StK_publicatie.pdf
2. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thurlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013;24(9):2206-23.
3. National Cancer Institute. Stage Information for Breast Cancer [Web page]. Available from: http://www.cancer.gov/cancertopics/pdq/treatment/breast/healthprofessional/page3
4. National Cancer Institute. Cancer Staging [Web page]. Available from: http://www.cancer.gov/cancertopics/factsheet/detection/staging
5. Ward S, Scope A, Rafia R, Pandor A, Harnan S, Evans P, et al. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: A systematic review and cost-effectiveness analysis. Health Technology Assessment. 2013;17(44):V-302.
6. Belgian Cancer Registry. Cancer Survival in Belgium. 2012. Available from: http://www.kankerregister.org/media/docs/publications/CancerSurvivalinBelgium.PDF
7. FOD Economie - Statbel. Doodsoorzaken [Web page].2009. Available from: http://statbel.fgov.be/nl/modules/publications/statistiques/bevolking/downloads/bevolking_-_doodsoorzaken.jsp
8. 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.
9. The Nottingham Prognostic Index Biomedical Informatics Research Group. The Nottingham Prognostic Index Calculator [Web page].2014. Available from: http://primed.info/finprog/npi_expl.asp

10. Adjuvant!Online. Adjuvant!Online. Decision making tools for health care professionals [Web page].2009. Available from: www.adjuvantonline.com/index.jsp
11. Wildiers H, Stordeur S, Vlayen J, Scholten R, van de Wetering F, Bourgain C, et al. Breast cancer in women: diagnosis, treatment and follow-up – KCE Reports 143C 3rd EDITION. D/2013/10.273/37. 2013. Available from: https://kce.fgov.be/sites/default/files/page_documents/KCE_143Cs_Breast_cancer.pdf
12. Schleidgen S, Klingler C, Bertram T, Rogowski WH, Marckmann G. What is personalized medicine: sharpening a vague term based on a systematic literature review. BMC Med Ethics. 2013;14(1):55.
13. National Institute for Health and Care Excellence. Gene expresion profiling and expanded immunohistochemistry test for guiding adjuvant chemotherapy decision in early breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat. 2013. NICE diagnostics guidance 10 Available from: http://www.nice.org.uk/nicemedia/live/14279/65265/65265.pdf
14. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
15. Genomic Health Inc. Oncotype DX [Web page].2014. Available from: http://www.oncotypedx.com
16. Jefferson T, Nicola Vicari N, Raatz H, (eds.). Prognostic tests for breast cancer recurrence (uPA/PAI-1 [FEMTELLE], MammaPrint, Oncotype DX) Agenzia nationale per i servizi sanitari regionali (age.na.s), Italy 2013. EUnetHTA core HTA Available from: https://meka.thl.fi/htacore/ViewCover.aspx?id=113
17. Isola JS, M; Kataja, V; Lundin, J; Hytönen M; Isojärvi, J; Mäkinen E. Geeniprofilointitestien merkitys rintasyövän hoidon valinnassa. Suomen Lääkäri lehti. 2013;50-52:3321-7.
18. Marchionni L, Wilson RF, Wolff AC, Marinopoulos S, Parmigiani G, Bass EB, et al. Systematic review: gene expression profiling assays in early-stage breast cancer. Ann Intern Med. 2008;148(5):358-69.
19. Smartt P. A comparison of gene expression profiling tests for breast cancer. Christchurch, New Zealand: Health Services Assessment Collaboration; 2010.
20. Medical Advisory Secretariat. Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer. 2010. Available from: www.health.gov.on.ca/english/providers/program/mas/tech/reviews/pdf/gep_20101213.pdf
21. 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.
22. Ab Rahman PM. Oncotype DX: breast cancer assay. Putrajaya: MaHTAS, Ministry of Health Malaysia; 2011.
23. Arpino G, Generali D, Sapino A, Del Matro L, Frassoldati A, de Laurentis M, et al. Gene expression profiling in breast cancer: a clinical perspective. BREAST. 2013;22(2):109-20.
24. Kuderer NM, Lyman GH. Gene expression profile assays as predictors of distant recurrence-free survival in early-stage breast cancer. Cancer Invest. 2009;27(9):885-90.
25. Lyman GH, Kuderer NM. Gene expression profile assays as predictors of recurrence-free survival in early-stage breast cancer: a metaanalysis. Clin Breast Cancer. 2006;7(5):372-9.
26. Kelly CM, Warner E, Tsoi DT, Verma S, Pritchard KI. Review of the clinical studies using the 21-gene assay. Oncologist. 2010;15(5):447-56.
27. 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.
28. 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.

29. Tang G, Shak S, Paik S, Anderson SJ, Costantino JP, Geyer CE, Jr., et al. Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20. Breast Cancer Res Treat. 2011;127(1):133-42.
30. Zhao L, Zhu S, Gao Y, Wang Y. Two-gene expression ratio as predictor for breast cancer treated with tamoxifen: evidence from meta-analysis. Tumour Biol. 2014;35(4):3113-7.
31. Krishnasamy M. MammaPrint. Putrajaya: MaHTAS, Ministry of Health Malaysia; 2008.
32. T. DMJ. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. . BMJ. 1996;313:275–83.
33. Lamond NW, Skedgel C, Younis T. Is the 21-gene recurrence score a cost-effective assay in endocrine-sensitive node-negative breast cancer? Expert rev. pharmacoecon. outcomes res. 2013;13(2):243-50.
34. Rouzier R, Pronzato P, Chereau E, Carlson J, Hunt B, Valentine WJ. Multigene assays and molecular markers in breast cancer: Systematic review of health economic analyses. Breast Cancer Research and Treatment. 2013;139(3):621-37.
35. 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. Journal of Medical Economics. 2013;16(1):30-40.
36. Hall PS, McCabe C, Stein RC, Cameron D. Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. Journal of the National Cancer Institute. 2012;104(1):56-66.
37. 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.
38. 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? Genetics in Medicine. 2005;7(6):380-9.
39. 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. European Journal of Cancer. 2013;49(18):3773-9.
40. 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.
41. 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 Research and Treatment. 2012;131(2):627-36.
42. Vataire AL, Laas E, Aballea S, Gligorov J, Rouzier R, Chereau E. Cost-effectiveness of a chemotherapy predictive test. Bulletin du Cancer. 2012;99(10):907-14.
43. 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.
44. 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.
45. Hornberger J, Chien R, Krebs K, Hochheiser L. US insurance program's experience with a multigene assay for early-stage breast cancer. Journal of Oncology Practice. 2011;7(3 SUPPL.):e38s-e45s.
46. 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. American Journal of Managed Care. 2005;11(5):313-24.
47. Mislick K, Schonfeld W, Bodnar C, Tong KB. Cost-Effectiveness analysis of mammostrat(registered trademark)compared with oncotype DX(registered trademark)to inform the treatment of breast

cancer. ClinicoEconomics and Outcomes Research. 2014;6(1):37-47.

48. 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. Genetics in Medicine. 2013;15(3):203-11.
49. 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.
50. 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.
51. 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(registered trademark) assay in oestrogen receptor positive node negative breast cancer. European Journal of Cancer. 2013;49(11):2469-75.
52. 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.
53. 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.
54. 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.
55. 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 in Health. 2013;16(5):729-39.
56. 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.
57. Kondo M, Hoshi SL, Ishiguro H, Toi M. Economic evaluation of the 70-gene prognosis-signature (MammaPrint(registered trademark)) in hormone receptor-positive, lymph node-negative, human epidermal growth factor receptor type 2-negative early stage breast cancer in Japan. Breast Cancer Research and Treatment. 2012;133(2):759-68.
58. 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 Research and Treatment. 2008;112(1):175-87.
59. Kondo M, Hoshi SL, Yamanaka T, Ishiguro H, Toi M. Economic evaluation of the 21-gene signature (Oncotype DX(registered trademark)) in lymph node-negative/positive, hormone receptor-positive early-stage breast cancer based on Japanese validation study (JBCRG-TR03). Breast Cancer Research and Treatment. 2011;127(3):739-49.
60. 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 in Health. 2010;13(4):381-7.
61. 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.
62. 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.
63. 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.

64. Earle CC, Chapman RH, Baker CS, Bell CM, Stone PW, Sandberg EA, et al. Systematic overview of cost-utility assessments in oncology. J Clin Oncol. 2000;18(18):3302-17.
65. Vataire AL, Laas E, Aballea S, Gligorov J, Rouzier R, Chereau E. Analyse cout-efficacite d'un test predictif de la chimiotherapie dans le cancer du sein (Oncotype DX()) en France. Bull Cancer. 2012;99(10):907-14.
66. Barcenas CH, Niu J, Zhang N, Zhang Y, Buchholz TA, Elting LS, et al. Risk of hospitalization according to chemotherapy regimen in early-stage breast cancer. J Clin Oncol. 2014;32(19):2010-7.
67. Wildiers HD, H; Fontaine, C; Awada, A. Adjuvant and neoadjuvant chemotherapy regimens in breast cancer: summary from the BSMO breast cancer task force meeting Belg J Med Oncol. 2014;8(4):116-24.
68. Bordoni REH, S T; Gilmore , J W; et al., 1Georgia Cancer Specialists A, Georgia; 2Amgen Inc, Thousand Oaks, California. Estimation of the incidence of febrile neutropenia in women receiving docetaxel plus cyclophosphamide as adjuvant therapy for early-stage breast cancer: a large community-based retrospective study Community Oncology. 2012.
69. Aapro MS, Cameron DA, Pettengell R, Bohlius J, Crawford J, Ellis M, et al. EORTC guidelines for the use of granulocyte-colony stimulating factor to reduce the incidence of chemotherapy-induced febrile neutropenia in adult patients with lymphomas and solid tumours. Eur J Cancer. 2006;42(15):2433-53.
70. 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.
71. Laas E, Vataire AL, Aballea S, Valentine W, Gligorov J, Chereau E, et al. Evaluation of the costs and resource use associated with adjuvant chemotherapy for breast cancer in France. J Med Econ. 2012;15(6):1167-75.
72. 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.
73. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst. 2009;101(21):1446-52.
74. Johnsson A, Fornander T, Rutqvist LE, Vaez M, Alexanderson K, Olsson M. Predictors of return to work ten months after primary breast cancer surgery. Acta Oncol. 2009;48(1):93-8.
75. Peugniez C, Fantoni S, Leroyer A, Skrzypczak J, Duprey M, Bonneterre J. Return to work after treatment for breast cancer: single center experience in a cohort of 273 patients. Bull Cancer. 2011;98(7):E69-79.
76. Broekx S, Den Hond E, Torfs R, Remacle A, Mertens R, D'Hooghe T, et al. The costs of breast cancer prior to and following diagnosis. Eur J Health Econ. 2011;12(4):311-7.