Genexpressionstest Mammaprint®

zur Entscheidungs-unterstützung für/gegen adjuvante Chemotherapie bei primärem Brustkrebs

EUnetHTA-Report

eunethta

Ludwig Boltzmann Institut
Health Technology Assessment
LBI-HTA Projektbericht Nr:. 102
ISSN: 1992-0488
ISSN-online: 1992-0496

Genexpressionstest Mammaprint®

zur Entscheidungs-unterstützung für/gegen adjuvante Chemotherapie bei primärem Brustkrebs

EUnetHTA-Report

Ludwig Boltzmann Institut
Health Technology Assessment

Wien, Jänner 2018

Zusammenfassung

Hintergrund:

MammaPrint® ist ein Test zur Ermittlung des Rezidivrisikos nach primärem Brustkrebs. Dieser sog. "Genexpressionstest" basiert auf der Analyse der Aktivität von 70 Genen im Brusttumorgewebe. Der Test soll zusätzlich zur etablierten klinisch-pathologischen Risikoabschätzung – in Abhängigkeit zum Risikobefund (niedrig/ hoch) – als Entscheidungshilfe für oder gegen eine adjuvante systemische Chemotherapie dienen. Der Anspruch dieses Tests ist es, dass einige PatientInnen auf der Basis von MammaPrint® von einer adjuvanten Chemotherapie absehen können und daher die damit verbundenen Nebenwirkungen wie Übelkeit und Erbrechen, Müdigkeit und Haarausfall, aber auch langfristige Nebenwirkungen nicht erleiden müssen.

MammaPrint® = Genexpressionstest zur Ermittlung des Rezidivrisikos und zur Entscheidungshilfe für oder gegen eine adjuvante systemische Chemotherapie

Ziel: Verzicht auf Chemotherapie

MammaPrint® wurde in klinischen Studien an PatientInnen mit Tumoren <5 cm (T1, T2, oder operabel T3) bis max 3 befallenen Lymphknoten (N0-1) unabhängig vom Östrogenrezeptor (ER) oder vom „human epidermal growth factor" 2 (HER2) Rezeptor Status validiert. MammaPrint® wird von Agendia (Niederlande: http://www.agendia.com) vertrieben.

Methode:

Systematische Suche in mehreren Datenbanken nach randomsierten klinischen Studien zu MammaPrint®. Zur Beurteilung der internen Validität wurde das Cochrane Risk of Bias (RoB) Tool verwendet. Die Qualität der Evidenz wurde für jeden Endpunkt mit der GRADE-Methode (Grading of Recommendations, Assessment, Development and Evaluation) durchgeführt.

systematischer Review
RoB: Cochrane
GRADE

Ergebnisse:

Es konnte ein RCT identifiziert werden: die im August 2016 veröffentlichte MINDACT Studie. Diese open-label, randomisiert-kontrollierte Phase 3 Studie wurde mit Frauen mit primärem Brustkrebs (n=6.693, medianes follow-up 5 Jahre) durchgeführt. Ziel von MINDACT war es, den klinischen Zusatznutzen von MammaPrint® zusätzlich zur klinisch-pathologischen Risikoabschätzung (AO!) gegenüber dieser etablierten Methode allein zu belegen. MINDACT fokussierte auf „diskordante Gruppen", also jene Subgruppen mit klinisch-hohem, aber genomisch-niedrigem Risikoprofil (CH/ GL) resp. klinisch-niedrigem und genomisch-hohem Risikoprofil (CL/GH). Der Nutzen von MammaPrint® war definiert als zuverlässiger (sicherer) Verzicht von Chemotherapie in der CH/GL Subgruppe ohne Auswirkungen auf das Fernmetastasen-freie Überleben (DMFS). Die Nicht-Unterlegenheitsgrenze wurde vorweg definiert als ein max 3% Risikounterschied zwischen den beiden Vergleichsgruppen.

1 RCT = MINDACT
6.693 Pts, 5 J FU
primärer Brustkrebs

AO! + MammaPrint® vs. AO!

Fokus: diskordante Subgruppen

Nicht-Unterlegenheitsstudie definiert als ≤ 3% Unterschied

Ergebnisse in der CH/GL Population (n=542 vs. 503) (Überprüfung der non-inferiority in der Per Protocol (PP) Population und superiority bei Lebensqualität (QoL) und Toxizität in intention-to-treat (ITT) Analyse): In der PP-Analyse betrug der Risikounterschied MammaPrint®+ AO! vs. AO! allein

in CH/GL Population (n=542 vs. 503 Pts):

non-inferiority bei DMFS und OS
inferiority bei DFS

- 2,5% bei 5-Jahres DMFS 96,5% (95% CI 94,1-97,9) mit Chemotherapie vs. 94,0% (95% CI 91,4-95,8) ohne Chemotherapie (HR=0,60; 95% CI 0,34-1,06; p=0.080; nicht statistisch signifikant) – zuungunsten von MammaPrint®,
- 4,5% bei 5-Jahres DFS 93,0% (95% CI 90,3-95,4) mit Chemotherapie vs. 88,8% (95% CI 85,7-91,3) ohne Chemotherapie (HR=0,57; 95% CI 0,37-0,87; p=0.01: statistisch signifikant) – zuungunsten von MammaPrint®,
- 1,8% bei 5-Jahres OS 98,8% (95% CI 97,1-99,5) mit Chemotherapie vs. 97,0% (95% CI 94,9-98,2) (HR=0,54; 95% CI 0,23-1,26; p=0.15; nicht statistisch signifikant) – zuungunsten von MammaPrint®.

Die 95% CIs der HR (bei sehr breiten Konfidenzintervallen) von 5-Jahres DMFS und bei 5-Jahres OS zeigten non-inferiority (Schwellwert HR 0.80, non-inferiority von 3% Unterschied), nicht aber die point estimates. Daher kann angenommen werden, dass bei Verzicht auf Chemotherapie basierend auf MammaPrint® ein erhöhtes Risiko auf Fernmetastasen und Tod im Vergleich zur AO!-Risikoabschätzung besteht. Obwohl MINDACT nicht ausreichend gepowered war, um statistisch signifikante Unterschiede zu zeigen, wurden signifikant schlechtere Ergebnisse beim 5-Jahres DFS gemessen: Non-inferiority kann aber nicht ausgeschlossen werden, wenngleich inferiority wahrscheinlicher ist.

sehr breite Konfidenzintervalle

Wahrscheinlichkeit, dass höhere Rate DMFS, DFS, OS

Bedenken zur Qualität der Evidenz und dem Verzerrungsrisiko bestehen wegen hoher Drop-Out Rate und open-label der Studie, der Indirektheit der Endpunkte (Surrogatendpunkte DMFS und DFS) und Impräzision (breite Konfidenzintervalle). Zusätzlich fehlen Ergebnisse zur kurzfristigen wie langfristigen QoL (aufgrund von Toxizitäten) in beiden Gruppen, um Aussagen zur Überlegenheit zu machen.

Bedenken an Qualität und Verzerrungsrisiko

fehlende Daten zu kurz- wie langfristigen QoL

Ergebnisse in der CL/GH Population (n=344 vs. 346): In dieser PatientInnengruppe (ITT Analyse) ist adjuvante Chemotherapie – entsprechend der klinisch-pathologischen Risikoabschätzung (AO!) – nicht indiziert, wohl aber basierend auf MammaPrint®. Der Risikounterschied zwischen MammaPrint® vs. vs. AO! betrug

in CL/GH Population (n= 344 vs. 346)

non-inferiority DMFS, DFS, OS

- 0,8% bei 5-Jahres DMFS 95,8% (95% CI 92,9-97,6) mit Chemotherapie vs. 95,0% (95% CI 91,8-97,0) ohne Chemotherapie (HR 1,17; 95% CI 0,59-2,28; p=0.66; nicht statistisch signifikant) – zuungunsten von MammaPrint®,
- 2% bei 5-Jahres DFS 92,1% (95% CI 88,3-94,6) mit Chemotherapie vs. 90,1% (95% CI 86,1-93,0) ohne Chemotherapie (HR 0,87; 95% CI 0,53-1,45; p=0.60; nicht statistisch signifikant),
- 0,7% bei 5-Jahres OS 97,1% (95% CI 94, 5-98,5) mit Chemotherapie vs. 97,8% (95% CI 95,5-99,0) ohne Chemotherapie (HR 1,28; 95% CI 0,54-3,02; p=0.58; nicht statistisch signifikant) (nicht statistisch signifikant).

Unterschiede nicht signifikant, aber zuungunsten von MammaPrint®

enge Konfidenzintervalle

Auch in dieser PatientInnengruppe konnte kein Zusatznutzen von MammaPrint® nachgewiesen werden, wobei MINDACT aber nicht ausrei-

kein Unterschied

chend gepowered war, um statistisch signifikante Unterschiede zwischen den Gruppen zu zeigen. Demnach kann ein klinischer Nutzen als Entscheidungsinstrument auch nicht ausgeschlossen werden.

Diskussion und Schlussfolgerung:

aus PatientInnen wie Kostenträger-Perspektive eine hochrelevante Fragestellung:

sicherer und verlässlicher Verzicht auf Chemotherapie

Aus Perspektive von Kostenträgern ist ein valider Vergleich zwischen der etablierten Methode der Risikoabschätzung auf ein Rezidivrisiko nach primärem Brustkrebs und einer neuen Methode zusätzlich zur herkömmlichen Methode für Refundierungsentscheidungen unabdingbar. In der MINDACT Studie wurden nur PatientInnen mit diskordanten Ergebnissen aus klinisch-pathologischer und aus genomischer Risikoabschätzung miteinander verglichen. Aus PatientInnen wie Kostenträger-Perspektive ist Frage nach einem möglichen Verzicht auf eine Chemotherapie hoch relevant. Bei Gleichwertigkeit der Risikoabschätungen (unter Akzeptanz der 3% Schwelle) wäre eine nachgewiesene Überlegenheit bei Lebensqualität wünschenswert. Solche Daten liegen allerdings nicht vor.

Ergebnisse von MINDACT : MammaPrint unsicher und der etablieren Risikoeinschätzung nicht überlegen, Verzicht auf Chemotherapie nicht zu rechtfertigen

Die Ergebnisse zeigen, dass der Nutzen von MammaPrint®, gemessen an den Endpunkten Fernmetastasen-freies Überleben (DMFS), krankheitsfreies Überleben (DFS) und Gesamtüberleben (OS) nach fünf Jahren unsicher und der etablieren Risikoeinschätzung nicht überlegen ist und daher der Verzicht auf eine Chemotherapie insb. bei PatientInnen mit hohem Rezidivrisiko nicht zu rechtfertigen ist. In der MINDACT Studie wurden die patientenrelevanten Endpunkte 10-Jahres-Überlebensrate, Lebensqualität und Nebenwirkungen einer Chemotherapie nicht erhoben und es liegt daher dazu keine Evidenz vor.

eunethta

EUROPEAN NETWORK FOR HEALTH TECHNOLOGY ASSESSMENT

EUnetHTA Joint Action 3 WP4

Rapid assessment of other technologies using the HTA Core Model® for Rapid Relative Effectiveness Assessment

MammaPrint®

Added value of using the gene expression signature test MammaPrint® for adjuvant chemotherapy decision-making in early breast cancer

Project ID: OTCA04

Version 1.3, 29 December 2017

1

DOCUMENT HISTORY AND CONTRIBUTORS

Version	Date	Description
V1.0	**18 July 2017**	First draft
V1.1	**29 September 2017**	Input from dedicated reviewers has been processed
V1.2	**23 November 2017**	Input from external review has been processed
V1.2	**23 November 2017**	Input from medical editor has been processed
V1.3	**29 December 2017**	Final version

Disclaimer

The assessment represents a consolidated view of the EUnetHTA network members and is in no case the official opinion of the participating institutions or individuals. EUnetHTA Joint Action 3 is supported by a grant from the European Commission. The sole responsibility for the content of this document lies with the authors and neither the European Commission nor EUnetHTA are responsible for any use that may be made of the information contained therein.

Assessment team

Author	National Health Care Institute (ZIN), The Netherlands
Co-author	Belgian Health Care Knowledge Centre (KCE), Belgium
Dedicated Reviewer	Ludwig Boltzmann Institute for HTA (LBI-HTA), Austria
Dedicated Reviewer	Haute Autorité de Santé (HAS), France

Further contributors

- S. Lange, MD: employee of the Institute for Quality and Efficiency in Health Care (IQWiG);
- Prof. dr. P.M. Bossuyt: Professor of Clinical Epidemiology at Academisch Medisch Centrum Amsterdam and Chairman of the Scientific Advice Committee of Zorginstituut Nederland.

Consultation of the draft Rapid Assessment

Manufacturer [v1.1] (factual accuracy check)	Agendia, The Netherlands
Medical editor [v1.1]	Nextgenediting Ltd, United Kingdom
Patient representatives [v1.1]	Borstkanker Vereniging Nederland, The Netherlands Nederlandse Federatie van Kankerpatiëntenorganisaties, The Netherlands
External experts [v1.1]	Nederlandse Vereniging voor Medische Oncologie (NVMO), The Netherlands

Conflict of interest

All authors, co-authors, reviewers (except 2 external experts), contributors and patient representatives, involved in the production of this assessment have declared they have no conflicts of interest in relation to the technology assessed according to the EUnetHTA Declaration of Interest and Confidentiality Undertaking (DOICU) statement. Two external experts, Carolien H. Smorenburg and Aafke Honkoop, have declared a financial or another

relationship with a Developing and/or Producing and/or Distributing Organisation (DPDO) for the technology or comparators undergoing assessment, and thus have a conflict of interest according to the EUnetHTA guidelines for handling conflict of interest. Carolien H. Smorenburg and Aafke Honkoop declared to have acted as principal investigators in a prospective cohort study regarding MammaPrint® which was funded by Agendia. According to the EUnetHTA guidelines for handling conflict of interest, the involvement of Carolien H. Smorenburg and Aafke Honkoop as external experts was acceptable which does not comply with Haute Autorité de Santé (HAS) procedures for the prevention of conflicts of interest.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

List of abbreviations

AC	Adjuvant doxorubicin and cyclophosphamide
AJCC	American Joint Committee on Cancer
AML	Acute myelogenous leukaemia
AO!	Adjuvant! Online
ASCO	American Society for Clinical Oncology
BCI	Breast Cancer Index
BIG	Breast International Group
CE	Conformité Européene
CH/GL	Clinical high and genomic low risk status
ChT	Chemotherapy
CI	Confidence interval
CL/GH	Clinical low and genomic high risk status
CUR	Health problem and current use of technology domains
DALY	Disability-adjusted life year
DFS	Disease-free survival
DMFS	Distant metastasis-free survival
DNA	Deoxyribonucleic acid
DOICU	Declaration of interest and confidentiality undertaking
EBCTCG	Early Breast Cancer Trialists' Collaborative Group
EFF	Clinical effectiveness
EORTC	European Organisation for Research and Treatment of Cancer
EMA	European Medicines Agency
ER	Oestrogen receptor
ESMO	European Society of Medical Oncology
ET	Endocrine therapy
EUnetHTA	European Network for Health Technology Assessment
FDA	Food and Drug Administration
FFPE	Formalin-fixed, paraffin-embedded tissue
GES	Gene expression signature
GMDN	Global Medical Device Nomenclature
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
HAS	Haute Autorité de Santé
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
HTA	Health technology assessment
ICD	International Classification of Diseases
IHC	Immunohistochemistry
IQWiG	Institute for Quality and Efficiency in Health Care
ITT	Intention to treat
JA3	Joint Action 3
KCE	Belgian Health Care Knowledge Centre
Ki67	Protein encoded by the *MKI67* gene
LBI-HTA	Ludwig Boltzmann Institute for HTA
M0-1	Presence or absence of metastasis (TNM staging system)
MCBS	Magnitude of Clinical Benefit Scale
MDS	Myelodysplastic syndrome
MeSH	Medical Subject Headings
MINDACT	Microarray In Node-negative and 1 to 3 positive lymph node Disease may Avoid ChemoTherapy
MP	MammaPrint®
N0-3	Number of positive lymph nodes (TNM staging system)

NBOCC	National Breast and Ovarian Cancer Centre
NCCN	National Comprehensive Cancer Network
NPI	Nottingham prognostic index
NSABP	National Surgical Adjuvant Breast and Bowel Project
NVMO	Dutch Association of Medical Oncology
OR	Odds ratio
OS	Overall survival
PASKWIL	Palliative, adjuvant, specific side effects, QoL, impact, and level of evidence
PICO	Population, intervention, comparator, and patient-related outcomes
PP	Per protocol
PPS	Per protocol sensitivity
PR	Progesterone receptor
QALYs	Quality-adjusted life years
QoL	Quality of life
RASTER	microarRAy prognoSTics in breast cancER
RCT	Randomised controlled trial
REA	Relative effectiveness assessment
RNA	Ribonucleic acid
SAF	Safety
SEER	Surveillance, Epidemiology, and End Results Program
T1-T3	Tumour size (TNM staging system)
TEC	Description and technical characteristics of the technology
TRANSBIG	Breast International Group to promote translational research
UK	United Kingdom
USA	United States of America
WAR	Dutch Scientific Advice Committee
WP4	Work Package 4
ZIN	National Health Care Institute

1 SUMMARY OF CLINICAL UTILITY OF MAMMAPRINT®

1.1 Scope

The scope of the project can be found here.

1.2 Introduction

Health problem
The target population in this assessment is patients with early breast cancer. Early breast cancer is defined as invasive cancer that is confined to the breast and/or has spread to a limited number of axillary lymph nodes but not metastasised to distant parts of the body **(A0002)**.

Overall, breast cancer has a relatively good prognosis: about 80% of patients with breast cancer are still alive ten years after diagnosis. In women diagnosed with stage I and II breast cancer, overall five-year survival is 87-98% and ten-year survival is 78%-94% **(A0023)**.

The mainstay of the management of early breast cancer is locoregional treatment with surgery alone or combined with radiotherapy. The aim of adjuvant systemic chemotherapy is to treat subclinical metastases already present at diagnosis in order to prevent the subsequent development of distant metastases. However, 60-70% of patients with early breast cancer appear to be free of subclinical metastases at diagnosis so do not develop distant metastases with locoregional management alone [1]. There is, therefore, considerable controversy with regard to the optimal definitions and assessment of risk of relapse in women with early-stage breast cancer. Improved prognostic tools are needed to prevent overtreatment with adjuvant chemotherapy, which carries a risk of late toxicity (especially cardiac toxicity and haematological malignancies) and consequent decreased quality of life (QoL) **(A0007)**. Several gene expression signature (GES) tests have been developed to better select patients for adjuvant chemotherapy.

This EUnetHTA assessment focuses on MammaPrint®, as it is currently the only GES test for which there is direct, peer-reviewed, published evidence of its clinical utility in the entire early breast cancer population.

Description of technology
MammaPrint® is a GES test that measures the expression of 70 genes related to metastatic cascade in breast cancer tissue acquired by biopsy or surgery. MammaPrint® is marketed by Agendia (Amsterdam, The Netherlands; http://www.agendia.com). Using these 70 genetic markers, patients can be divided into low- and high-risk groups, which in turn supports clinical decision-making for adjuvant treatment. MammaPrint® has been prospectively validated for use in early-stage breast cancer patients with tumours <5 cm (T1, T2, or operable T3), and 0-3 positive lymph nodes (N0-1) regardless of oestrogen receptor (ER) or human epidermal growth factor 2 (HER2) receptor status **(B0001)**.

The decision to administer adjuvant chemotherapy is usually based on clinicopathological risk assessment. A number of algorithms are currently used to predict survival and the utility of adjuvant therapy. For example, Adjuvant! Online (AO!) is a widely used web-based algorithm designed to provide estimates of the benefits of adjuvant endocrine therapy and chemotherapy. To provide estimates of risk reduction of breast cancer-related death or relapse at ten years, AO! uses information about the efficacy of different therapeutic regimens from the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) meta-analysis **(B0001)**.

Main research question of the EUnetHTA assessment
In this EUnetHTA assessment the main research question is: does adding MammaPrint® to standard risk assessment with AO! in patients with a high clinical risk profile substantially and

positively affect the health and well-being of women with early breast cancer by limiting the number of patients receiving adjuvant chemotherapy and related adverse events, on the condition that survival is not negatively affected? From a reimbursement perspective, it is necessary that the new and standard approaches are compared.

1.3 Assessment method for clinical utility

The clinical utility was at first limited to randomised controlled trials (RCTs) because they provide the highest level of evidence for MammaPrint®'s clinical utility. Evidence on MammaPrint®'s clinical utility is available from one RCT, the Microarray In Node-negative and 1 to 3 positive lymph node Disease may Avoid ChemoTherapy (MINDACT) study. Two prospective studies were also identified, but did not qualify as supportive evidence for clinical utility and were therefore excluded. From a reimbursement perspective, it is necessary that health-outcomes according to the new and standard approaches are compared. Part of the main research question of this EUnetHTA assessment is to prove that MammaPrint® does not negatively affect survival. This is a non-inferiority question, therefore a non-inferiority threshold must be specified. However, there is no international consensus on this non-inferiority threshold, so a 3% difference in ten-year overall survival (OS) or a hazard ratio (HR) <0.8 in case of immature survival data was used based on the European Society for Medical Oncology criteria for the Magnitude of Clinical Benefit Scale (ESMO-MCBS) and a conventional trade-off between toxicity and efficacy [2,3]. Different scientific organisations and countries use different thresholds, so, during the reimbursement decision process, each country will need to decide individually which non-inferiority threshold to use.

We used the Cochrane Risk of Bias Tool to assess internal validity. The quality of the evidence was assessed as part of examining the overall documentation for each outcome using Grading of Recommendations, Assessment, Development and Evaluation (GRADE) methodology.

1.4 Results

Characteristics of the available evidence
One RCT (MINDACT) was identified [4]. MINDACT was an open-label, randomised controlled, phase 3 trial of women with early-stage breast cancer (n=6693, median follow-up five years). The investigators aimed to provide prospective evidence of the clinical utility of MammaPrint® added to standard clinicopathological risk assessment compared to standard clinicopathological risk assessment alone. MINDACT focused on discordant risk groups: a subgroup with a clinical high-risk profile discordant with a genomic low-risk profile (clinical high and genomic low; CH/GL) and a subgroup with a clinical low-risk profile discordant with a genomic high-risk profile (clinical low and genomic high; CL/GH). MINDACT's primary objective was to assess whether chemotherapy could be safely withheld in the CH/GL subgroup without affecting distant metastasis-free survival (DMFS). The authors predefined the cut-off for non-inferiority as the lower boundary of the 95% confidence interval (CI) of the five-year rate of DMFS should be more than 92% in CH/GL patients who did not receive chemotherapy based on the MammaPrint® result.

Results of the clinical high-risk and genomic low-risk population
In case of the CH/GL population, the added value will not be in terms of survival, as MINDACT's primary objective is not superiority but to safely spare chemotherapy in this group (non-inferiority in term of OS; PP analyse). The added value is that patients will potentially experience a better QoL (superiority in terms of QoL and toxicity; ITT analyse).

Direct comparison of overall survival or surrogate outcomes (non-inferiority claim)
Ten-year OS data is not yet available from the MINDACT study, so results of surrogate endpoints five-year DMFS, five-year disease-free survival (DFS) and five-year OS were

reported. In the PPS[1] analysis, the risk differences in the CH/GL patients treated according to MammaPrint® compared to AO! were 2.5% for five-year DMFS (HR= 0.60; 95% CI 0.34-1.06; p=0.080), 4.5% for five-year DFS (HR=0.57;95% CI 0.37-0.87; p=0.009) and 1.8% for five-year OS (HR of 0.54 (95% CI 0.23-1.26; p=0.154) all in advantage of AO! (**Table 1.1**). The differences in five-year DMFS and five-year OS were not significantly different. The 95% CIs of five-year DMFS and five-year OS crossed the non-inferiority threshold of HR 0.80, indicating non-inferiority of MammaPrint® is questionable. However, the point estimates were on the left side of the non-inferiority threshold. It is therefore reasonable to assume that using MammaPrint® may result in an increased risk of death due to distant metastasis, in comparison to AO! risk assignment. Although, the study was not adequately powered to assess statistical differences, MammaPrint® risk assignment led to a significantly worse five-year DFS. The upper 95% CI boundary of 0.87 is above the non-inferiority threshold of 0.80, which means that non-inferiority in terms of OS cannot be ruled out. However inferiority is more likely since the lower boundary of the 95% CI of 0.37 and the point estimate of 0.57 both crosses the ESMO-MCBS threshold of 0.65 (GRADE A)[2], indicating DFS was significantly and clinically relevant worse in patients in whom treatment was based on MammaPrint®. For those HTA-organisations that use risk differences in their assessment the same conclusion can be made. The risk differences are 2.5% for five-year DMFS and 4.5% for five-year DFS and 1.8% for five-year OS. The absolute risk difference of five-year DFS (4.5%) crosses the non-inferiority threshold of 3%. Those of DMFS and OS (2.5% and 1.8% respectively) do not cross the non-inferiority threshold, but since in general the confidence intervals are large, it is reasonable to assume that also for five-year DMFS and five-year OS the required non-inferiority is not shown.

The quality of the evidence for the critical ten-year OS endpoint was low (when five-year DMFS and DFS were used as surrogates) to very low (when direct measurements of OS but only five-year data was used). Low quality was due to concerns about a risk of bias (considerable number of drop-outs and open-label study), indirectness (use of surrogate endpoints), and imprecision and very low (five-year OS) due to additional imprecision (crossing both thresholds). For these reasons, confidence in the OS effect estimate after ten years is at best limited, and the true effect may be substantially different.

Direct comparison of outcomes on quality of life and toxicity (superiority claim)
MINDACT did not evaluate long-term health-related quality of life (QoL). It may be argued that some aspects of QoL are reflected by other outcomes. It is obvious that the QoL of patients receiving adjuvant chemotherapy will be reduced due to chemotherapy side effects during and shortly after treatment compared to patients who do not receive chemotherapy. On the other hand, the MINDACT study shows that refraining from chemotherapy leads to a significant and clinically-relevant worse five-year DFS. Recurrences of all types are stressful to patients even in the case of a curable disease. This distress will have its repercussions on quality of life. Short- and long-term side effects of chemotherapy were measured, but these results have yet to be published. In the absence of data on QoL and toxicity the superiority in terms of QoL cannot be quantified.

Results of the clinical low-risk and genomic high-risk population
In addition to the main EUnetHTA research question, this assessment also examined the added value of MammaPrint® in the other discordant risk group (CL/GH). In this case, adjuvant chemotherapy is not indicated according to standard clinical risk assessment, but is indicated when based on the MammaPrint® result. Therefore, the added value is in terms of OS and hence the ITT population data is presented.

[1] Due to the temporary change in risk as a result of assay problems, all risk groups as enrolled in that particular period are somewhat biased due to incorrect risk assessment. In addition to the prespecified PP analyses, also so-called PPS analysis is presented in the MINDACT publication, in which all patients enrolled during the period of change in risk were excluded. This PPS analysis represents the least biased and therefore most conservative PP analysis.

[2] ESMO describes the highest level of clinical benefit (GRADE A) <0.65 and <0.80 as GRADE B.

CL/GH patients who after randomisation received chemotherapy on the basis of genomic risk had a five-year DMFS rate of 95.8% (95% CI 92.9-97.6), whereas those not receiving chemotherapy (randomly assigned based on clinical risk) had a 0.8% lower five-year DMFS rate of 95.0% (95% CI 91.8-97.0). However, this difference was not statistically significant (p=0.657). The five-year DFS surrogate endpoint was 92.1% (95% CI 88.3-94.6) for patients who received chemotherapy (on the basis of genomic risk) and 90.1% (95% CI 86.1-93.0) for patients who did not receive chemotherapy (based on clinical risk), 2.0% lower than the rate among those who received chemotherapy based on the MammaPrint® result (p=0.603). For five-year OS, 97.1% (95% CI 94.5-98.5) of patients who received chemotherapy (based on genomic risk) and 97.8% (95% CI 95.5-99.0) of patients who did not receive chemotherapy (based on clinical risk) were still alive, 0.7% lower in those who received chemotherapy based on the MammaPrint® result (p=0.578) (**Table 1.1**). None of the surrogate endpoint differences were significant, and added value of MammaPrint® has not been demonstrated at this time. However, the study was not powered to assess significant differences in this discordant group, so a clinical benefit for MammaPrint®-based treatment decision-making cannot be ruled out.

Table 1.1 Summary table of relative effectiveness of the MammaPrint® assay

	Early breast cancer					
	Health benefit				**Harm**	
	DMFS 5y HR (95% CI) p-value	**DFS 5y** HR (95% CI)	**OS 5y** HR (95% CI)	**QoL**	**Short-term side effect of chemotherapy**	**Long-term side effect of chemotherapy**
Clinical high/Genomic low (PPS)						
MammaPrint® + AO! (n=542)	***0.60 (0.34-1.06) p=0.08*** ***Using clinical risk 96.5 (94.1-97.9)*** ***Using genomic risk 94.0 (91.4-95.8)***	***0.57 (0.37-0.87) p=0.01*** ***Using clinical risk 93.3 (90.3-95.4)*** ***Using genomic risk 88.8 (85.7-91.3)***	***0.54 (0.23-1.26) p=0.15*** ***Using clinical risk 98.8 (97.1-99.5)*** ***Using genomic risk 97.0 (94.9-98.2)***	*Not measured#*	*Not reported*	*Not reported*
AO! (n=503)						
Quality of body of evidence+	⊕⊕◯◯ Low (1)	⊕⊕◯◯ Low (1)	⊕◯◯◯ Very low (2)			
Clinical low/ Genomic high (ITT)						
MammaPrint® + AO! (n=344)	***1.17* (0.59-2.28) p=0.66*** ***Using clinical risk 95.0 (91.8-97.0)*** ***Using genomic risk 95.8 (92.9-97.6)***	***0.87*(0.53-1.45) p=0.60*** ***Using clinical risk 90.1 (86.1-93.0)*** ***Using genomic risk 92.1 (88.3-94.6)***	***1.28* (0.54-3.02) p=0.58*** ***Using clinical risk 97.8 (95.5-99.0)*** ***Using genomic risk 97.1 (94.5-98.5)***	*Not measured#*	*Not reported*	*Not reported*
AO! (n=346)						
Quality of body of evidence+	⊕◯◯◯ Very low (3)	⊕◯◯◯ Very low (3)	⊕◯◯◯ Very low (3)			

Abbreviations: DMFS: distant metastasis-free survival; DFS: disease-free survival; OS: overall survival; QoL: quality of life; PPS: per protocol sensitivity population, ITT: intention to treat population

[+] Quality of the body of evidence was rated using GRADE. The interpretation is: high, confident that the true effect lies close to that of the estimated effect; moderate, moderately confident in the effect estimate, the true effect is likely to be close to the estimated effect, but there is a possibility that it is substantially different; low, limited confidence in the effect estimate, the true effect may be substantially different from the estimated effect; very low, very little confidence in the estimated effect, the true effect is likely to be substantially different from the estimated effect.

(1) Downgraded twice due to concerns about risk of bias (open-label study and amount of drop-outs), indirectness (surrogate endpoints), and imprecision;

(2) Downgraded three times due to concerns about risk of bias, indirectness and imprecision;

(3) Downgraded three times due to concerns about indirectness (once) and imprecision (twice).

** The HRs are not concordant with the DFS and OS rates after a median follow-up of five years of the MINDACT.*

[#]QoL was measured in 347(566 enrolled) patients but only six-eight weeks after surgery and not during/directly after chemotherapy and not in the long term. Therefore this study could not be used to determine a difference in QoL in the long term between the patients that received treatment based on the MammaPrint® result and patients that received treatment based on the AO!.

1.5 Discussion

From a reimbursement perspective, a comparison must be made between the new and the standard approach because added value needs to be proven if a test or intervention is added to standard care. MINDACT does not have a formal non-inferiority design, because such a trial would either need to be extremely large or of long duration. In MINDACT's first secondary analysis, outcomes were compared for patients in discordant risk groups according to whether risk assessment based on MammaPrint® (as an add-on) or AO! was used to assign patients to the chemotherapy or non-chemotherapy group. This comparison is of primary importance for reimbursement decisions for MammaPrint®.

Based on the MINDACT it cannot be concluded that it is safe to omit chemotherapy, because the 95% CI's of all surrogate outcomes for ten-year OS (five-year DMFS, five-year DFS and five-year OS) are crossing the non-inferiority threshold (HR 0.80 and 3% risk difference). Even the results of five-year DFS indicate that MammaPrint® is possibly clinically relevant inferior in comparison with treatment based on AO!.

In addition, there are three main observations that underline the uncertainty with respect to the safety of omitting chemotherapy based on MammaPrint®:

First, the surrogate five-year DFS endpoint showed a statistically and clinically relevant worse outcome (p=0.009) for MammaPrint® patients. It is plausible that the investigated group was large enough to reveal an effect on DFS even without the power being calculated for this secondary analysis. Since all outcome measures point in the same direction, it is less likely that this effect is merely due to chance when no true difference between the groups existed.

Second, the study was not powered to assess statistically significant differences, so the absence of a significant DMFS and OS difference should not be interpreted as evidence of absence of an effect.

Third, and related to the second observation, all surrogate outcomes had wide confidence intervals. Expressed as absolute numbers, introducing MammaPrint® can lead to 100 (of 1000) more patients not free of distant metastases after five years compared to AO!-based treatment decisions. Considering the 955 CI of five-year DMFS, it could also lead to six per 1000 less or at worst 287 per 1000 more patients not free of distant metastases after five years. This CI highlights the degree of uncertainty, and there remains a possibility that many patients could be harmed.

1.6 Conclusion on clinical utility

Each HTA-organisation will need to decide individually which non-inferiority threshold and clinical relevant thresholds are considered most appropriate for their assessment considering the local context. Based on the chosen thresholds in this EUnetHTA assessment report, conclusions are as follows:

Taking everything into consideration, it has not yet been demonstrated that patient outcomes (ten-year OS and QoL) are improved by withholding adjuvant chemotherapy based on MammaPrint® testing in the CH/GL risk group. In other words, the clinical utility of the MammaPrint® is not proven. This conclusion is based on the absence of evidence on added value in terms QoL and on the fact that non-inferiority in terms of OS (surrogates five-year DMFS, five-year DFS and five-year OS) is not shown. In addition, there are concerns about the certainty of DMFS because of the imprecision (very wide 95% CI's). Therefore the results do not rule out the possibility of a clinically-relevant increase in distant metastasis and hence risk of death. Also, the significant and clinically-relevant difference in DFS is of importance as QoL data is not available. The quality of the evidence for the critical ten-year OS endpoint was

rated as low to very low. Therefore, the confidence in the OS effect estimate after ten years is limited at best.

Furthermore, a clinical benefit of receiving chemotherapy in the CL/GH risk group according to genomic risk assignment has not been demonstrated. The quality of the evidence for the critical ten-year OS endpoint was rated very low. Therefore, there is very little confidence in the effect estimate, and the true effect is likely to be substantially different from the effect estimate.

Ten years of follow-up will be needed to conclude if an add-on risk assessment approach (AO! combined with MammaPrint®) has superior clinical utility compared with treatment decisions based on AO! alone. At that time, the level of evidence of the data and therefore confidence in the data will be higher, because the use of surrogate endpoints will no longer be necessary. If a revision of AO! becomes available, as it is expected, this could have an impact on the baseline risks of recurrence and hence may potentially limit the clinical applicability of the MINDACT results.

Ultimately, the decision to receive or forgo chemotherapy (or any other treatment) lies with each patient who is properly informed about the potential side effects and the potential benefits of such treatment. For the same risk–benefit scenario, different patients may make different decisions. However, well-informed decision making is only possible if both parameters (OS and QoL) has been quantified.

2 SCOPE

2.1 Background

Gene expression signature tests (GES tests) such as the MammaPrint® 70-gene signature have the potential to provide prognostic information to distinguish early-stage breast cancer patients who are likely to remain free of distant metastases from patients who are likely to develop distant metastasis. These GES tests aim to improve the information on a patient's risk of (distant) recurrence in guiding therapy to patients who will benefit most from adjuvant chemotherapy from those who will have limited benefit.

The scope of this EUnetHTA assessment focuses on the direct evidence of clinical utility of GES tests in treatment decision-making on adjuvant chemotherapy. A scoping search was undertaken to define the scope. First, the background of the current assessment is presented through a short narrative of the scoping search results.

Reviews covering the literature from 1990 to 2014 concluded that, of the evaluated tests, Oncotype DX® and MammaPrint® were the furthest along their validation pathways **(Table A1, Appendix 1).** The majority of the literature provide evidence on test performances, such as the analytic validity (test performances on repeatability and reproducibility) and the clinical validity (prognostic ability in the case of GES, calibration, discrimination, re-/classification) [5-8]. Given that these GES tests were developed as prognostic tests more than as predictive tests and that clinical implementation was already underway, subsequent reviews focused on the clinical utility of Oncotype DX® and MammaPrint® [9,10]. Clinical utility refers to the test's ability to predict or identify the patients who will have more or less benefit of a therapeutic intervention. From 2014 onwards the available evidence of clinical utility evaluated *indirect* outcomes of clinical utility, such as the impact of GES on reclassification of risk in clinical practice, or evaluating the correlations between test score and score on existing measures based on decision-making tools such as the Nottingham Prognostic Index (NPI) and AO!. In contrast to indirect evidence of clinical utility, direct evidence evaluates the overall improvement in patient-relevant treatment outcomes by directly comparing treatment outcomes using the GES test to guide treatment decisions, with treatment outcomes using commonly established risk criteria not incorporating findings from the GES test. Although the scoping search was limited to the English and Dutch languages, one English executive summary was identified which outlined the main conclusions of a systematic review originally written in German by the Institute for Quality and Efficiency in Health Care (IQWiG). This review included one study that may provide direct evidence of clinical utility: a randomised study evaluating MammaPrint® (MINDACT) [11]. The other included studies in this review evaluated indirect evidence of other GES tests. In summary, with the latter review included, it was concluded that there is no proven benefit from using a GES test. As the vast majority of the currently available evidence is indirect, direct evidence of the clinical utility of GES tests is still needed.

Randomised studies provide level 1 evidence on patient benefit, potentially endorsing widespread clinical use. The scoping search identified three ongoing randomised studies: MINDACT (MammaPrint®), TAILORx (Oncotype DX®), and RxPONDER (Oncotype DX®). These studies will provide direct evidence of clinical utility using a randomised controlled study design that follows patients from initial diagnosis through to final health outcomes. The current EUnetHTA assessment focuses on MammaPrint®, as it is currently the only GES test for which results with regard to direct evidence of clinical utility on the entire early-stage breast cancer population have been published in a peer-reviewed journal. In assessing clinical utility, we consider it as important that the GES test should be performed on the entire early-stage breast cancer population in order to identify all patients who would have been treated differently by using the GES test. This will allow the direct comparison of the treatment outcomes using the

GES test and the treatment outcomes using commonly established risk criteria. The TAILORx study and the RxPONDER study were ongoing at the time of the current EUnetHTA assessment. The results of the TAILORx study are due to be published in December 2017, and the results of RxPONDER are due within the next few years.

2.2 Main research question of this EUnetHTA assessment

In this EUnetHTA assessment the main research question is: does adding MammaPrint® to standard risk assessment with Adjuvant! Online in patients with a high clinical risk profile substantially and positively affect the health and well-being of women with early breast cancer by limiting the number of patients receiving adjuvant chemotherapy and related adverse events, on the condition that survival is not negatively affected?

2.3 Scope

Description	Project scope
Population	Early-stage breast cancer patients (pT1-2, operable T3, N0-1,cM0) ICD-10: C50 MeSH: Breast neoplasms
Intervention	MammaPrint® MammaPrint® is a gene expression signature test used to decide on whether to administer adjuvant chemotherapy. MammaPrint® will be assessed as an add-on to standard clinicopathological criteria using the Modified Adjuvant! Online (and as a replacement). Technology: MammaPrint® is a prognostic genomic test that aims to provide a risk assessment of mamma carcinomas by providing a risk profile (i.e., low or high) of the chance of developing distant metastases. MammaPrint® is a gene expression signature test that measures the expression of 70 genes in cancerous breast tissue. MeSH: Gene expression profiling
Comparison	Modified Adjuvant! Online is an online decision tool used to decide on whether to administer adjuvant chemotherapy. Modified Adjuvant! Online was chosen as the comparator. Treatment decision-making on adjuvant chemotherapy is based on clinicopathological risk criteria (modified Adjuvant! Online): a high clinical risk means that adjuvant chemotherapy is indicated, and a low clinical risk means that no adjuvant chemotherapy is indicated.
Outcomes	*Critical endpoints for relative effectiveness/safety* • Ten-year overall survival (OS) • Health-related quality of life (QoL) • Short- and long-term side effects from chemotherapy such as cardiovascular and haemato-oncologic toxicity such as (sub-)clinical cardiac failure or secondary leukaemia, respectively If necessary, surrogate endpoints will be included. The relation between the surrogate endpoint and critical endpoint will be described.
Study design	*Effectiveness/safety* • Randomised controlled trial (RCT) • If evidence from RCTs is limited, prospective observational studies will be considered for inclusion to provide more stable estimates of clinical utility.
Follow-up time	Follow-up time should be at least ten years and, if unavailable, shorter follow-up times where acceptable surrogate endpoints are available will be considered.

3 METHODS AND EVIDENCE INCLUDED

3.1 Assessment team

The tasks were assigned to the agencies as follows:
As *Author*, ZIN:
- Played a leading role in both scoping and production of the clinical utility assessment;
- Was responsible for the management of the completed scientific work;
- Had ultimate responsibility for quality assurance;
- Answered comments.

For collaborative assessments, the following mode of assessment was pursued:
The 1st authors (AL and YK) were responsible for the production of all domains, including data extraction from clinical trials, finding answers to the questions listed in the Project Plan, and writing the assessment. A third ZIN reviewer (HS) followed and verified every step taken by the 1st authors during the production of the assessment including data extraction, verification of references, risk of bias tables, and adherence to methods.

As *Co-author* of the clinical utility assessment, KCE:
- Was responsible for supporting the author in all project phases;
- Was responsible for reviewing all domains including the clinical utility domain.

As *Dedicated Reviewers*, LBI-HTA and HAS:
- Guaranteed quality assurance by thoroughly reviewing the project plan and the assessment drafts;
- Reviewed methods, results, and conclusions based on the original studies included;
- Provided constructive comments in all project phases.

3.2 Scoping meeting/patient involvement

A scoping meeting was organised at the start of the assessment where representatives of Dutch patients, clinicians, hospitals, and healthcare insurers were present (09.03.2017 at Zorginstituut Nederland at Diemen). During this meeting, the Population, Intervention, Comparator, and patient-related Outcomes (PICO) were discussed. All relevant parties agreed with the PICO as described in this assessment. The experts were asked what threshold for clinical relevance is being used by their scientific societies. During the phase of written consultation of the assessment patient representatives and external experts (clinicians) were consulted. KCE also organised a scoping meeting at the start of the assessment regarding the cost-effectiveness of the MammaPrint® where clinicians and scientist/data-analysts were present. This group of experts could also find themselves in the PICO as described.

3.3 Source of assessment elements

The assessment element selection was based on the HTA Core Model® Application for Rapid REA Assessments. The selected issues (generic questions) were translated into actual research questions (answerable questions).

3.4 Search

Current use of technology (CUR) domain

This domain was developed starting with the information provided by the manufacturer within the Manufacturer's Submission File. A face-to-face meeting with Agendia was organised on 20.12.2016. In addition to the information provided in the Submission File, information was integrated with selected scoping searches, *ad hoc* PubMed and internet searches of the grey

literature using the Google search engine, review of the reference lists and bibliographies of the studies identified through the basic selective search, manufacturers' websites, brochures, and information for use.

Technical characteristics of the technology (TEC) domain

This domain was developed starting with the information provided by the manufacturer within the Manufacturer's Submission File. In addition to the information provided in the Submission File, information was integrated with *ad hoc* PubMed and internet searches of the grey literature using the Google search engine, review of the reference lists and bibliographies of studies identified through the basic selective search, manufacturers' websites, brochures, information for use, and regulatory bodies' databases.

Clinical effectiveness (EFF) and safety (SAF) domains

These domains were developed using a systematic structured literature search. The following databases were searched:

- MEDLINE;
- Embase;
- Cochrane Library.

Search strategy

For the clinical utility outcome, direct evidence from a RCT was necessary. The evidence base covering literature from 1990–2014 established that MammaPrint® (next to Oncotype DX®) was the furthest along the pathway of analytic and clinical validation. Given that these GES tests were developed as prognostic tests and not as predictive tests, and that clinical implementation was already underway, subsequent reviews highlighted the need for direct evidence of the clinical utility of Oncotype DX® and MammaPrint®. Therefore, the current search strategy on the clinical utility of MammaPrint® included all relevant available literature on MammaPrint® from June 2014 (KCE search date). In addition, information provided in the Manufacturer's Submission File and recent literature updates provided by the manufacturer during the assessment period were integrated in the search.

MeSH terms (see **Appendix 1**, **Section 9.2**) were combined with the following terms to perform the searches: "MammaPrint®" (non-MeSH) or "70-gene" or "70 gene" or "MINDACT". As we anticipated that few RCTs had been conducted, we did not include the MeSH term "Randomised Controlled Trial" to identify potential supporting evidence from studies with a lower level of evidence. Therefore, no limits relating to study design were applied.

All searches were limited to English and Dutch language sources published between June 2014 (date of KCE literature search [5] and the time of searches (April 11th 2017).
In addition, the following clinical trials databases were searched to identify on-going trials or studies:

- ClincalTrials.gov;
- Cochrane Register of Controlled Trials;
- https://www.clinicaltrialsregister.eu/.

If possible and of added value, results for the following subgroups (or combinations) were presented:

- Low clinical risk population;
- High clinical risk population;
- ER status;
- HER2 status;
- Lymph node status.

Inclusion and exclusion criteria
Publications on studies with a RCT design were included for the assessment of clinical utility. As described above, to capture all the available literature on direct evidence of clinical utility we did not include limits relating to study design for supportive evidence. However, supportive evidence from studies of a lower level of evidence were only included when prospectively obtained data with sufficient follow up were available of patients with discordant test results to allow for the direct comparison of the treatment outcomes using the MammaPrint® test and the treatment outcomes using commonly established risk criteria.

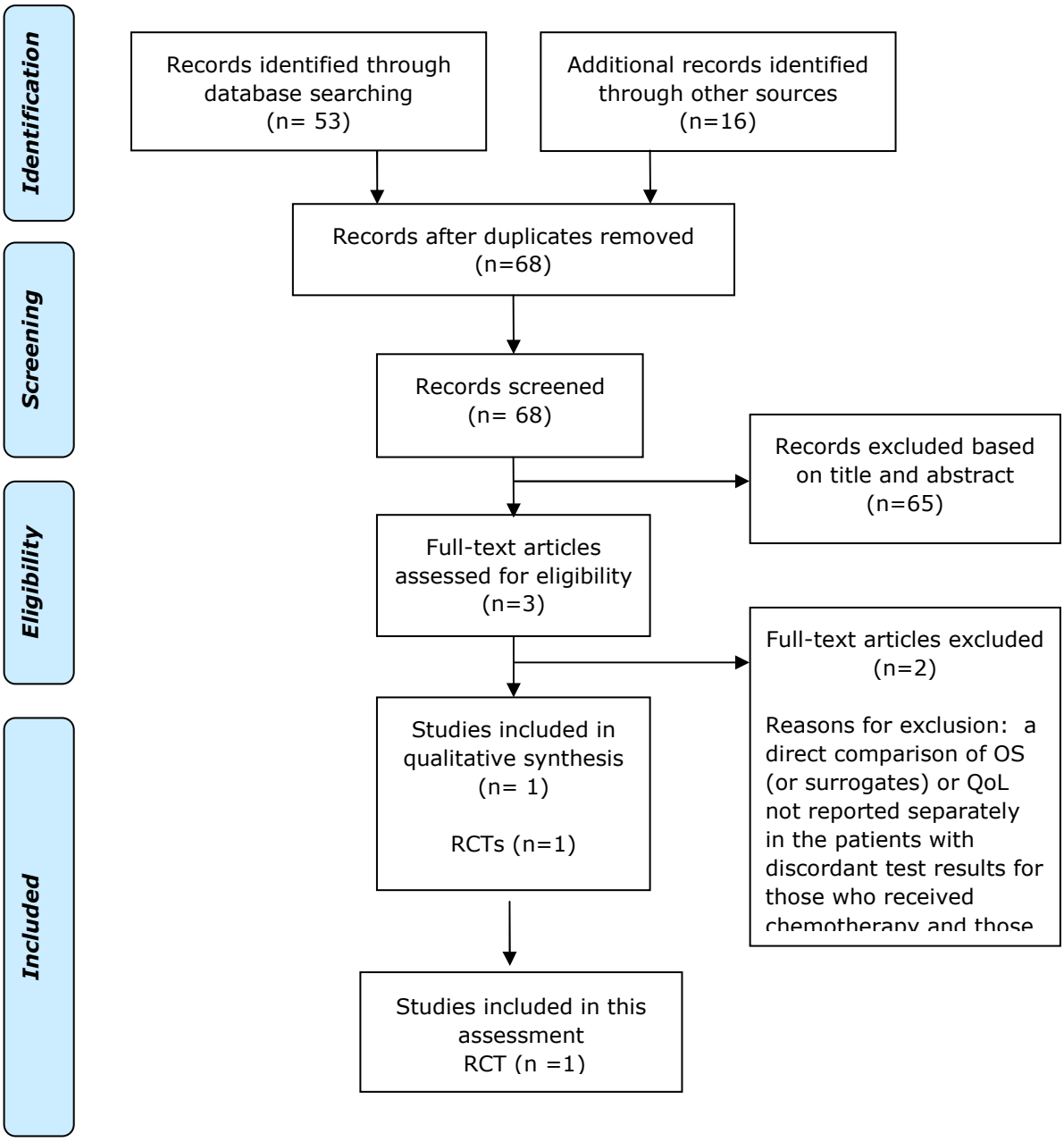
The following publication types were excluded: retrospective studies, technical and/or clinical validation studies; non-early breast cancer patients; studies on neoadjuvant chemotherapy; animal models, preclinical and biological studies; editorials, reviews, guideline publications, expert opinions; histologic types other than ductal carcinoma; abstracts, posters; and non-peer-reviewed publications.

3.5 Data extraction strategy and flow chart of study selection

Two authors screened the records by title and abstract. Disagreements were solved by discussion. The full text of potentially relevant studies were retrieved and reconsidered for actual inclusion in the current evidence review. The two researchers performed data extraction independently. The retrieved data were cross-checked against the submission file received from the manufacturer for completeness. If necessary, additional data was extracted from original papers. Subsequently, a third reviewer verified the references and the information retrieved from the extracted publications. See **Figure 3.1**.

Description of excluded records
Based on title and abstract, records were excluded for the following reasons: guideline publications (n=3); comments, letters and reviews (n=16); miscellaneous reasons (n=26 for reasons of neo-adjuvant setting, other histologic type (not ductal carcinoma), pre-clinical setting, concordance studies, cost-effectiveness studies, other language and not related to GES). Studies on technical and clinical validation were excluded (n=12), including studies with a retrospective-prospective design. These studies retrospectively analyse archived tumour samples derived from prospectively selected cohort(s) of patients with documented information on baseline recurrence risks based on standard clinicopathological criteria and treatment outcomes after re-classification based on MammaPrint®. This design is associated with increased bias compared to prospectively planned studies. These studies were considered as prognostic validation or hypothesis generating studies and therefore cannot be considered as supportive evidence on the clinical utility of the MammaPrint®. In addition, decision impact studies (n=5) were excluded when no data was provided in which health related treatment outcomes, such as OS (or surrogates) or QoL, according to MammaPrint® risk assessment were directly compared with health related treatment outcomes according to standard risk assessment in the patients with discordant test results. These studies were not considered as direct evidence on clinical utility of the MammaPrint® and for this reason excluded as supportive evidence.

Figure 3.1 Flow chart of study selection



3.6 Description of the evidence used

Three potentially relevant studies were retrieved for full text reconsideration of direct evidence of clinical utility. Evidence on the clinical utility of the 70-gene profile (MammaPrint®) was available from one RCT (MINDACT) only. One prospective observational study evaluating clinical utility was also identified (The microarRAy-prognoSTics-in-breast-cancER (RASTER) study), in which survival data (OS or a surrogate) was reported [12]. However, the OS (or surrogate) for the discordant risk groups was not reported separately for those who did and did not receive chemotherapy following either standard risk assessment or the MammaPrint® respectively. In contrast, only the results of the discordant group (CH/GL) as a whole were published in which 44% of the patients received chemotherapy and 56% did not, and not the treatments outcomes of the CH/GL patients that received chemotherapy separately from those who did not receive chemotherapy. For this reason, this prospective study did not qualify as supportive for direct evidence of clinical utility and was therefore excluded. QoL was measured by Retel et al [13] in 347 (566 enrolled) patients of the MINDACT. The primary aims of the study were to evaluate the association between breast cancer patients' well-being and the results of a gene expression profile on to compare different recurrence risk groups, according to their genomic and standard clinical risk assessment. Different questionnaires were taken to assess the QoL. The QoL was assessed 6-8 weeks after surgery. This study does not compare QoL in the long term between CH/GL patients receiving treatment based on the MammaPrint® result and receiving treatment based on the AO!. Therefore, this study was not qualified as supportive evidence on clinical utility.

All relevant information about the included study is described in **Table A5** of **Appendix 1**.In **Table A6** **of** **Appendix 1** the excluded studies and reasons for exclusion are presented.

MINDACT

MINDACT is an international, open-label, randomised, controlled, phase 3 trial in women with early-stage breast cancer (n=6693, five-year median follow-up period) comparing standard clinicopathological risk profiling with MammaPrint® 70-gene risk profiling added to standard clinicopathological criteria based on a modified version of AO! for selecting patients for adjuvant chemotherapy. In MINDACT, 6693 patients were enrolled in 112 institutions in nine European countries between 2007 and 2011. Eligible patients were women between aged between 18 and 70 years with histologically confirmed primary invasive breast cancer (stage T1, T2, or operable T3). The primary objective of MINDACT was to assess whether, in patients with high clinical risk and low genomic risk (CH/GL) in whom chemotherapy was withheld, the lower boundary of the 95% CI for the five-year distant metastasis-free survival (DMFS) rate would be 92% (i.e., the non-inferiority boundary) or higher at a one-sided significance of 0.025.

The study consisted of three consecutive randomisations. In the treatment randomisation (R-T randomisation), patients with discordant risk scores were randomly assigned to the chemotherapy group or the no-chemotherapy group on the basis of either the clinical or the genomic result.

The treatment randomisation (R-T randomisation) used a minimisation technique that was stratified according to institution, risk group, HR status, nodal involvement, age (<50 years vs. ≥50 years), HER2 status, axillary treatment (sentinel node only vs. dissection), and type of surgery (mastectomy vs. breast conservation). Two additional (optional) randomisations were conducted, in which patients who were assigned to receive adjuvant chemotherapy (R-C randomisation; either randomly because of discordant results or due to high-risk concordance in both tests) could be randomly assigned to receive an anthracycline-containing regimen or a docetaxel-plus-capecitabine regimen. Similarly, patients with hormone receptor-positive breast cancer could undergo further randomisation to a tamoxifen-letrozole regimen or a letrozole-only regimen (R-E randomisation). The current publication of Cardoso et al. [4] focuses on the results of the R-T randomisation only.

In the RT-randomization, three pre-specified secondary analyses were conducted. The first pre-specified secondary analysis evaluated patient outcomes in the discordant risk groups according to whether the patients were assigned to the chemotherapy group or the no-chemotherapy group. In the second pre-specified secondary analysis, the outcomes of all patients according to whether chemotherapy had been recommended by either clinical or genomic risk alone were evaluated. In the third pre-specified secondary analysis, the percentage of all enrolled patients who would be assigned to chemotherapy on the basis of either clinical risk or genomic risk was evaluated. The main characteristics of MINDACT are described in **Table 3.1** and **Table A5** in **Appendix 1.**

Table 3.1 Main characteristics of studies included

Author and year or study name	Study type	Number of patients	Intervention(s)	Main endpoints	Sponsoring
Cardoso et al. 2016 MINDACT [4]	RCT	6693	Experimental: MammaPrint® added to Adjuvant! Online to make "yes-no" decision on adjuvant chemotherapy. Comparator: Treatment decision based on Adjuvant! Online alone. Additional (optional) randomisations: • Chemotherapy randomisation: anthracycline or docetaxel-plus-capecitabine regimen. • Hormone randomisation: tamoxifen-letrozole or letrozole-only regimen.	Primary: five-year DMFS[$] Secondary: Proportion of patients that received chemotherapy according to the clinical risk compared with the genomic risk. Five-year OS* Five-year DFS[#]	The MINDACT trial was supported by grants from the European Commission Sixth Framework Program, the Breast Cancer Research Foundation, Novartis, F. Hoffmann-La Roche, Sanofi-Aventis, Eli Lilly, Veridex, the U.S. National Cancer Institute, the European Breast Cancer Council-Breast Cancer Working Group, Jacqueline Seroussi Memorial Foundation for Cancer Research, Prix Mois du Cancer du Sein, Susan G. Komen for the Cure, Fondation Belge contre le Cancer, Dutch Cancer Society, Genomics Initiative-Cancer Genomics Centre, Association le Cancer du Sein, Parlons-en!, the Brussels Breast Cancer Walk-Run and the American Women's Club of Brussels, NIF trust, German Cancer Aid, the Grant Simpson Trust and Cancer Research UK, Ligue Nationale contre le Cancer, and the EORTC Cancer Research Fund. Whole-genome analysis was provided by Agendia without cost.

Abbreviations: DFS: disease-free survival; DMFS: distant metastasis-free survival; OS: overall survival.

Source: MINDACT

[$] Distant metastasis-free survival: survival without distant metastasis was defined as the time until the first distant metastatic recurrence or death from any cause.

[#] Disease-free survival was defined as the time until first disease progression (loco-regional, distant relapse, ipsilateral or contralateral invasive breast cancer, ductal carcinoma *in situ* or an invasive second primary cancer, or death from any cause.

* Overall survival was defined as the time until death from any cause.

3.7 Assessment method for clinical utility

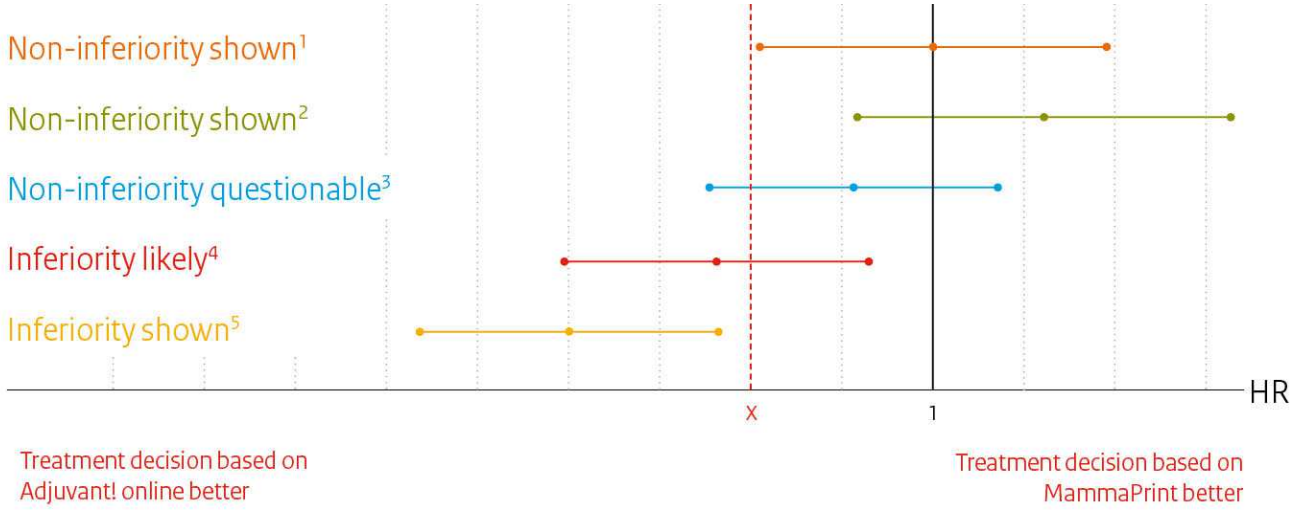
From a reimbursement perspective, the added value of a new intervention or diagnostic test must be proven compared to standard care. In this EUnetHTA assessment, the main research question was to assess whether MammaPrint®, when added to standard risk assessment, leads to a substantial positive effect on the health and well-being of women by limiting the number of patients receiving adjuvant chemotherapy and related adverse events, on the condition that survival is not negatively affected. The first randomisation in which the treatment decision is randomised, that is the R-T randomisation, was the analysis of interest for the assessment of clinical utility. In fact, MINDACT's first pre-specified secondary analysis addressed the main EUnetHTA research question: to answer the question of whether it is safe to withhold chemotherapy based on genomic risk assignment, a direct comparison is required between patients in whom chemotherapy was withheld following genomic risk assignment vs. patients who received chemotherapy following standard risk assignment. By contrast, MINDACT's primary analysis did not compare the two arms, instead focusing only on one arm. Hence, the primary analysis of the study did not address the current research question.

3.7.1 When is clinical utility proven in the CH/GL risk group?

If MammaPrint® is added to standard clinical risk assessment, MammaPrint® must add value in terms of health-related outcomes. In the case of the population with high clinical risk and low genomic risk (CH/GL), the added value will not be in terms of survival, as MINDACT's primary objective is not superiority but to safely spare chemotherapy in this group (non-inferiority assessment). The added value (superiority assessment) is that patients will potentially experience a better QoL and fewer side effects from chemotherapy like cardiovascular toxicity and secondary haematological cancers. To determine clinical utility of MammaPrint® the weighing of positive and negative effects must lead to a benefit in terms of patient health related outcomes.

The claimed benefit of the MammaPrint® is that treatment decision based on MammaPrint® does not lead to worse survival in comparison with AO!. This is a non-inferiority claim. In case of non-inferiority significant differences are irrelevant. Non-inferiority is shown when the 95% CI is entirely above the non-inferiority threshold. **Figure 3.2** shows when non-inferiority is shown (orange and green line), questionable (blue line) and when there is (possible) inferiority (red and yellow line). The per protocol (PP) analysis of the OS data is reported. The PP analysis represents a conservative estimate of the possible harm of MammaPrint® as opposed to the ITT analysis, which tends to find no effect. In this case, one wants to be certain that a treatment decision to forgo adjuvant chemotherapy based on MammaPrint® does no harm in terms of OS. For this reason it was decided that the PP analysis should be used as by convention [14]. For the sake of completeness the ITT analyses are reported in full in **Table 6.1** and **Table A14**.

Figure 3.2 Concept of non-inferiority of OS in the CH/GL group



X = the non-inferiority threshold

In theory, this means that when non-inferiority of OS is shown (see orange[1] and green line[2] in **Figure 3.2**) and there is evidence of reduced toxicity and improved QoL, MammaPrint®'s clinical utility would be proven. When non-inferiority of OS is questionable (see the blue line[3] in **Figure 3.2**), it cannot be ruled out that MammaPrint® use results in a clinically-relevant deterioration. In this case, positive effects on toxicity and QoL must outweigh a possible inferior effect on OS. When there is a statistically significant difference (red[4] and yellow line[5] in **Figure 3.2**), inferiority is likely but non-inferiority cannot be ruled out either (red line[4]) or inferiority is shown (yellow line[5]). HTA-organisations or countries possible will have different opinions if (possible) loss in OS can be outweighed by benefits in QoL and toxicity or not.

3.7.2 When is clinical utility proven in the CL/GH risk group?
In the case of adding MammaPrint® to the low clinical risk and a high genomic risk population (CL/GH), superiority in OS must be demonstrated, because chemotherapy is being given to a population that will not normally receive chemotherapy according to standard clinicopathological criteria.

3.8 Method of quality rating the studies

One RCT (MINDACT) was included in the assessment of clinical utility of MammaPrint®. The quality of this study was analysed using the Cochrane risk of bias checklist and GRADE. GRADE is used to qualitatively summarise results from the relative effectiveness and safety domains. No quality assessment tool was used for the TEC and CUR domains. Multiple sources were used to validate individual, potentially biased, sources. Descriptive analysis and synthesis were performed using different information sources.

GRADE quality ratings reflect the certainty of evidence. It is therefore desirable to pre-specify the thresholds or ranges used to rate the certainty of evidence for individual outcomes. The available clinical thresholds are described below in **Section 3.8.1**.

3.8.1 Clinical relevance thresholds
To identify widely accepted clinical relevance thresholds, the literature was searched and experts were asked for clinically relevant differences in OS (or surrogates) for oncological interventions.

Different scientific organizations and countries use different thresholds (see **Table 3.2**). From

a reimbursement perspective, the added value of a new intervention or diagnostic test must be proven compared to standard care. In this EUnetHTA assessment, the main research question was to assess whether MammaPrint®, when added to standard risk assessment, leads to a substantial positive effect on the health and well-being of women by limiting the number of patients receiving adjuvant chemotherapy and related adverse events, on the condition that survival is not negatively affected. To answer this question, a direct comparison between the two arms is required. By contrast, MINDACT's primary analysis did not compare the two arms, instead focusing only on one arm. 92% was used as non-inferiority threshold in the MINDACT in case of a one-arm analysis, but cannot be used as a non-inferiority threshold in case of a risk difference between two arms. Therefore this MINDACT threshold of 92% cannot be used in this EUnetHTA assessment. Consequently, other literature on clinical relevance and non-inferiority thresholds were searched (see **Table 3.2**). This list is not an exhaustive overview of thresholds used. Each country will need to decide individually which threshold to use. Even though MammaPrint® is a diagnostic test, and thresholds for clinical relevance are primarily developed to assess the efficacy of anticancer therapies, the thresholds can reasonably be used to determine the clinical utility of a diagnostic test because in case of clinical utility the added value of the 'test plus treatment' on health related outcomes needs to be proven. In the case of a diagnostic test which enables reclassification of prognostic risk, such as MammaPrint®, these thresholds can be applied to determine whether a difference in OS (or surrogate outcomes) is clinically relevant. Since the current case is to assess whether a risk-difference in OS is clinically relevant to prove that MammaPrint®, is non-inferior, a threshold for non-inferiority is needed. However, there is no international consensus on non-inferiority thresholds (both in case of efficacy and in case of diagnostic tests). Tanaka et al. [14] concluded in a systematic review on cancer studies evaluating non-inferiority that selection of non-inferiority margins in early breast cancer studies is frequently selected from efficacy trials in which a trade-off between toxicity and efficacy is evaluated. According to this method (referred to as the conventional method) the non-inferiority margin is determined by the size of effects that are considered to be of no clinical relevance or to be outweighed by other benefits of the experimental treatment (conventional method) [16].

Table 3.2 Summary of possible thresholds of clinical relevance

Reference	Threshold of clinical relevance
MINDACT	Lower boundary of 95% CI of five-year DMFS of 92%
Experts scoping meeting (NVMO)	Ten-year OS of 92% (point estimate)
Experts scoping meeting (NVMO)	3% risk difference in ten-year OS benefit
IQWiG [12]	(One-sided) lower 95% CI for DFS <95%
PASKWIL criteria (NVMO) [15,16]	Five-year OS risk difference >5% or HR<0.7 (point estimate)
ESMO-MCBS	3% risk difference in case of mature survival data or HR <0.80 (lower boundary of 95% CI) in studies without mature survival data
Tanaka et al. 2012 [14]	non-inferiority margins in early breast cancer trials varied between HR 1.25 and HR 1.40 (or 0.60 and 0.75 in case of omitting therapy)

Abbreviations: DFS: disease-free survival; ESMO-MCBS: European Society for Medical Oncology-Magnitude of Clinical Benefit Scale; HR: hazard ratio; IQWiG: Institute for Quality and Efficiency in Health Care; NMVO: Dutch Association of Medical Oncology; OS: overall survival; PASKWIL: Palliative, adjuvant, specific side effects, QoL, impact, and level of evidence; ZINL: National Health Care Institute.

Since this is a European assessment, the European Society for Medical Oncology-Magnitude of Clinical Benefit Scale (ESMO-MCBS) clinical relevance thresholds were used [2].
ESMO has developed ESMO-MCBS, a validated and reproducible tool to assess the magnitude of clinical benefit for anti-cancer therapies. The ESMO-MCBS is intended to assist oncologists in explaining the likely benefits of a particular treatment to their patients and to help public

health decision-makers prioritise therapies for reimbursement. The ESMO-MCBS can be applied to comparative outcome studies evaluating the relative benefit of treatments in solid cancers using outcomes of survival, treatment toxicity, QoL, or surrogate survival, QoL, or treatment outcomes. The ESMO-MCBS has separate evaluation forms: an evaluation form for therapies with curative or adjuvant intent, an evaluation form for therapies that are not likely to be curative and with primary endpoints other than OS or PFS, and an evaluation form for equivalence studies.

Although MINDACT did not adopt a standard non-inferiority design, it relates to an equivalence question. For that reason, the evaluation form for equivalence studies (form 2C) is applicable to the evaluation of the MammaPrint® [2]. According to ESMO-MCBS criteria for equivalence studies, a high level of clinical benefit is proven when reduced toxicity or improved QoL (using validated scales) is shown with evidence of statistical non-inferiority or superiority in PFS or OS. To determine non-inferiority, a non-inferiority threshold is needed, but there is no international consensus on non-inferiority thresholds. For that reason, the GRADE B threshold of the ESMO-MCBS (form 1) is used. ESMO describes the highest level of clinical benefit (GRADE A) as a survival improvement of >5% at ≥3 years follow-up and a GRADE B survival improvement as between 3% and 5%. In the case that mature survival data is not available, GRADE A clinical benefit is assigned when DFS improvements are found in which the HR is <0.65 and GRADE B clinical benefit is assigned when the HR is between 0.65 and 0.80. In both cases, the HR threshold refers to the lower extreme of the 95% CI to take into account estimate variability. Since the threshold will be used as a non-inferiority threshold, the 3% difference in OS or an HR <0.8 is used here to assess if MammaPrint® is non-inferior in the case of ten-year OS (or a surrogate). This 3% difference in ten-year OS was also mentioned in the scoping meeting as a minimal important difference. Dutch oncologists (NVMO) consider chemotherapy if at least 3% more patients (with similar characteristics) are alive after ten years. A smaller benefit does not outweigh the risk of serious toxicity of around 2%. In the publication of De Boer et al. [3] it is explained how this consensus of 3% difference was achieved [3]. This method in which a trade-off between toxicity and efficacy is one of the conventional methods for selecting non-inferiority threshold as mentioned by Tanaka et al [14]. Tanaka et al. [14] evaluated non-inferiority trials in early breast cancer where non-inferiority margins in HR varied between 1.25 and 1.40 (or 0.60 and 0.75 in case of omitting therapy). This range approaches the ESMO-MCBS range of 0.65 and 0.80.

3.9 Deviations from project plan

The Scientific Advisory Committee of Zorginstituut Nederland was not included in the current EUnetHTA assessment because the Scientific Advice Committee of Zorginstituut Nederland is specifically entrusted with assessments for Dutch reimbursement decisions. We instead added Prof. dr. P.M.M. Bossuyt as a contributor. Prof. dr. P.M.M. Bossuyt is Professor of Clinical Epidemiology with a special interest in methods for evaluating medical tests and markers. Prof. dr. P.M.M. Bossuyt is also chairman of the Scientific Advice Committee of Zorginstituut Nederland. Dr. S. Lange, MD, employee of IQWiG, was also added as a contributor.

4 HEALTH PROBLEM AND CURRENT USE OF TECHNOLOGY

4.1 Research questions

Element ID	Research question
A0002	What is the disease in the scope of this assessment?
A0007	What is the target population in this assessment?
A0023	How many people belong to the target population?
A0005	What are the symptoms and the burden of disease for the patient?
A0004	What is the natural course of the disease?
A0006	What are the consequences of the disease for society?
A0024	How is breast cancer currently diagnosed and staged?
A0025	How is the breast cancer treatment and risk assessment for adjuvant systemic management according to published guidelines and in practice?

4.2 Results

4.2.1 Breast cancer: overview

Breast cancer is the most frequently diagnosed malignancy in women worldwide. In 2012, the estimated age-adjusted annual incidence of breast cancer in 40 European countries was 94.2/100,000 and mortality was 23.1/100,000 [17]. Breast cancer mortality rates in Europe have decreased over the last few decades due to early detection by breast cancer screening, improved treatment, and better coordinated clinical pathways [18-20]. Overall, breast cancer has a relatively good prognosis, and about 80% of patients with breast cancer are still alive ten years after diagnosis. Relative survival depends on tumour stage at the time of diagnosis. In women diagnosed with stage I and II breast cancer, overall five-year survival is 87-98% and ten-year survival is 78%-94%. In women diagnosed with advanced stage III and IV disease, survival rates are 65-85% and 46-76%, respectively [21] **(A0002)**.

Breast cancer may not cause any signs or symptoms in its early stages. Signs and symptoms often appear when the tumour grows large enough to be felt as a lump in the breast or when the cancer spreads to surrounding tissues or distant parts of the body. However, the disease burden is significant and can be quantified in terms of quality-adjusted life years (QALYs) or disability-adjusted life years (DALYs), both of which quantify the number of years lost due to the disease. In women, breast cancer is sixth in the top ten of diseases with the highest burden. Causes of years of healthy life lost to disability include side effects during and after treatment (for example, after radiotherapy, chemotherapy, or hormonal therapies), potential changes in the menopause, the effects of lymphoedema, and the psychosocial differences in "life after therapy". Most of the disease burden (70%) of breast cancer is caused by premature death from breast cancer. Distant metastases account for the majority of breast cancer deaths [17] **(A0005)**.

4.2.2 Target population in this assessment

The target population in this assessment is patients with early breast cancer. Early breast cancer is an invasive cancer that is confined to the breast and/or has spread to a limited number of axillary lymph nodes but there are no detectable metastases in distant parts of the body. Breast cancer tends to be a disease of the older age with about a quarter of breast cancers occurring before age 50 and <5% before age 35. The estimated five-year prevalence (percentage of patients still alive five years after diagnosis) in Europe in 2012 was 1.8 million cases [22]. Ninety per cent were older than 45 years, with the majority (70%) being older than 60 years.

Two thirds of breast cancer patients are diagnosed with hormone receptor-positive disease. Early-stage breast cancer patients have a good ten-year OS as described above [23]. Hence,

preserving QoL and surviving with cancer have become important aims of disease management as more patients live long enough to be at risk of long-term side effects from treatment and for developing competing causes of morbidity and death such as cardiovascular disease **(A0007+A0023)**.

4.2.3 Natural course of early breast cancer

Despite the good survival rates in the general early breast cancer population, clinicians have long recognised heterogeneity in the natural course of the disease and responsiveness to cancer therapies [24]. In contrast to locoregional recurrence, distant metastases represent virtually incurable disease, as the death rate is close to 100%. There is, however, a complex relation between the onset of distant recurrence and the hazard for death due to distant recurrence. Late-onset distant metastases, occurring five years or more after primary surgery, are often associated with a better prognosis than early-onset distant metastases occurring within the first five years after primary surgery [25]. The annual hazard of distant recurrence peaks in the second year following primary surgery but remains at 2-5% in years five to 20. About a half of all patients with hormone receptor-positive and HER2-negative cancers who develop distant recurrences do so after five years [23]. Patients with node-positive disease tend to have higher annual hazards of recurrence than patients with node-negative tumours. In the first few years following primary treatment, the risk of recurrence is higher in patients with ER-negative cancers, but five to eight years after diagnosis, their annual hazard of recurrence drops below the level of ER-positive tumours. Breast cancer relapse may occur as late as over 20 years after initial diagnosis, particularly in patients with hormone receptor-positive disease. **(A0004)**.

4.2.4 Diagnosis and staging

A definitive diagnosis of early breast cancer is provided by histopathological assessment of tumour tissue following surgery. Overall, breast cancer diagnosis is based on triple assessment: clinical assessment, imaging, and tissue sampling. Histopathological assessment of tumour tissue provides information on prognostic factors that might indicate the presence of occult metastasis. Independent prognostic factors are tumour size, nodal status, histological grade, and lymphatic invasion [24,26]. In addition, immunohistochemical (IHC) detection of hormone receptor expression (oestrogen (ER) and progesterone (PR) receptors) and human epidermal growth factor receptor 2 (HER2) gene amplification status and protein expression are evaluated. As well as being prognostic, these factors predict responses to endocrine and HER2-targeting therapies. Proliferation markers such as Ki67 may supply additional useful prognostic information [27,28].

In early-stage breast cancer, routine staging evaluations are particularly directed at locoregional disease. As distant metastases are very rare in early-stage disease, patients do not benefit from comprehensive laboratory (including tumour markers) and radiological staging at diagnosis [29]. Current breast cancer staging is based on the TNM (tumour-nodes-metastasis) system. For breast cancer, the different TNM combinations correspond to five stages that indicate tumour burden and disease extent: stage 0 corresponds to carcinoma *in situ*, stage I to III indicate invasive disease with tumour size (T stage) up to 5 cm and/or spread of the cancer limited to nearby lymph nodes (N stage); and stage IV corresponds to cancer spread to distant tissues and organs (M1 stage: distant metastasis) [30]. Stage I, stage IIA, stage IIB, and stage IIIA describe early breast cancer: T1-2 (tumour size 2–5 cm) and operable T3 (in case of tumours larger than 5 cm) and N0-1 (disease has spread up to lymph nodes) and M0 (no distant metastasis) **(A0024)**.

4.2.5 Treatment of early breast cancer

The mainstay of early breast cancer management is locoregional treatment of the breast through surgery alone or combined with radiotherapy. Optimal management also includes the treatment of occult or subclinical metastases (i.e., metastases not detectable at diagnosis) with adjuvant systemic therapy. In general, adjuvant systemic therapy will be advised if the

estimated absolute ten-year survival benefit is at least 3%. This 3% survival benefit threshold is selected since it is considered to be needed to outweigh the incidence of severe toxicity of approximately 2% [3]. Adjuvant chemotherapy primarily prevents early metastasis [31-33]. However, there is an on-going risk of distant recurrence after five years. This is especially the case in ER-positive/HER2-negative breast cancer, where recurrences occur after five years in approximately one half of all distant recurrence cases [23]. It is unclear whether these late recurrences are prevented by adjuvant chemotherapy. The 2005 Early Breast Cancer Trialists' Collaborative Group (EBCTCG) overview [31] stated that the relative benefit of adjuvant chemotherapy is similar in all subgroups independent of age, stage, histopathological grade, and ER status. However, data from the 2012 EBCTCG overview suggested that the absolute benefit is greater for those at high-risk of recurrent disease [33]. As the current assessment focuses on adjuvant chemotherapy, neoadjuvant chemotherapy is not discussed further **(A0025)**.

Risk assessment and adjuvant chemotherapy planning

Therapeutic decisions on adjuvant disease management are based on risk estimates. With current treatment strategies, up to 30% of women diagnosed with early-stage breast cancer will eventually develop metastatic breast cancer, at which point it becomes virtually incurable. The death rate from metastatic disease, which is close to 100%, has provided the impetus for adjuvant systemic therapy use in early-stage breast cancer. However, the chance of occult metastases at the time of diagnosis is not equal in each patient. Approximately 60-70% of patients with early breast cancer appear to be free of subclinical metastases at diagnosis and do not develop distant metastasis when managed with locoregional treatment alone [1,5]. Risk profiling is necessary in order to assign adjuvant chemotherapy to patients most likely to benefit from systemic therapy without unnecessarily placing the patient at increased risk of treatment side effects **(A0025)**.

Historically, young age, tumour size, and axillary nodal status have been important clinical prognostic factors for disease relapse. Larger tumours and a multiple number of positive lymph nodes are considered poor prognosis tumours. However, there seems to exist a subgroup of patients with small tumours who are at high risk of distant recurrence [24,26]. Likewise, patients with a limited number of positive lymph nodes (up to three) may represent a subgroup of lymph node-positive women with a low (similar to node-negative) risk of distant recurrence [1,31]. The so-called "triple-negative" tumours (ER-, PR-, and HER2-negative) form a subgroup with early metastatic capacity. ER-positive tumours are considered to have a more indolent prognosis, as they are responsive to hormonal therapy. In addition to being clinically heterogeneous, microarray-based gene expression studies have also revealed the significant molecular heterogeneity of breast cancer [34-38] (**Table A3** in **Appendix 1**). As a result, distinct "intrinsic" molecular breast cancer subtypes can now be recognised within the breast cancer spectrum. These subtypes range from tumours with favourable long-term survival (luminal A-type breast cancer) to tumours with significantly worse long-term survival (HER2-positive tumours and basal-like (luminal B often triple-negative) tumours). Each subtype also has been shown to be associated with a differential response to treatment with adjuvant therapy. It has now become routine practice to tailor treatment regimens to a tumour's molecular signature (i.e. ER and PR status and HER2 status) while also considering traditional clinicopathologic characteristics (age, tumour size and tumour grade). Guidelines from international scientific societies such as the St. Gallen Breast Cancer Conference, ESMO, and the American Society of Clinical Oncology (ASCO) recommend that adjuvant chemotherapy should be considered for all patients with early invasive breast cancer after surgery. The guidelines underline that the absolute benefit to the individual patient varies substantially according to the disease burden and intrinsic tumour biology [39-46]. (**Figure A1** in **Appendix 1**).

To improve decision-making for adjuvant chemotherapy, the use of prognostic risk classification systems such as AO!, and the Nottingham Prognostic Index (NPI) is recommended. AO! uses individual patient and pathological data combined with Surveillance,

Epidemiology, and End Results Program (SEER) population data to assess baseline risk. Clinical trial efficacy data from the EBCTCG is also incorporated into the risk tool to produce age- and tumour-specific relative risk estimates and to predict individual treatment benefit from chemotherapy [63]. Although these algorithms have been globally investigated, concerns have been raised regarding their applicability to populations other than those used in the validation studies [46]. Hence, the strength of clinical guideline recommendations and the use of AO and the NPI vary across countries **(A0025)**. More recent guidelines from the St Gallen, ESMO and ASCO societies recommend on the use of GES tests to guide decision-making for adjuvant chemotherapy. These guidelines have been adjusted very recently based on the publication of the MINDACT study results (**Table A4** in **Appendix 1**). The strength of the guideline recommendations for using GES tests such as MammaPrint® varies in different countries, increased use in clinical practice has been observed over time **(A0025)**.

4.2.6 Consequences of the disease for the patient and society

The consequences of breast cancer and breast cancer treatment include the impact on patient well-being during treatment and after completion of treatment. The consequences of morbidity and disability associated with breast cancer and its treatment also include the impact on the ability of the patient to continue working and include costs associated with treatment and care due to morbidity and disability.

Side effects during and shortly after chemotherapy and targeted therapy

Anthracycline- or taxane-based chemotherapy regimens are current standard adjuvant systemic therapy. In addition, HER2-targeted therapy with agents such as trastuzumab may be indicated for HER2-positive disease. Chemotherapy and targeted therapy are associated with side effects during and shortly after treatment. Common toxic effects of current regimens include nausea, vomiting, anorexia, diarrhoea, myelosuppression, febrile neutropenia, haemorrhagic cystitis (in schemes using cyclophosphamide), fever and infection, alopecia, mucositis, neurotoxicity (taxane-based chemotherapy), amenorrhea (in premenopausal women), and serum biochemistry abnormalities. More general symptoms such as anorexia, fatigue, and cognitive changes are experienced by patients and have long been observed following adjuvant chemotherapy [47]. However, defining standardised and uniform outcome measures for these general symptoms has been hampered by methodological limitations.

Late side effects of chemotherapy and targeted therapy

Chemotherapy is associated with late toxicity, in particular cardiac and haematological toxicity. Although the frequency of these toxicities is low, these specific toxicities are serious and potentially life threatening [47] **(A0006)**.

Anthracycline cardiotoxicity is well established, particularly at higher cumulative doses. Trastuzumab for HER2-positive breast cancer is also an established cardiotoxic agent. The frequency of cardiac adverse events is low (<1%), but these events are potentially life threatening and the reported cumulative cardiac mortality is 0.6% [48,49]. Sequential and non-anthracycline/trastuzumab-based chemotherapy regimens reduce subsequent risk [50]. "Late-onset cardiotoxicity" develops following a prolonged asymptomatic period, with heart failure presenting one year to decades following chemotherapy. Reassuringly, long-term data from these trials suggest that the cumulative cardiac event rate after seven years of follow-up is small at 1.7% [43,51]. Still, long-term knowledge on the late effects from anthracyclines in otherwise asymptomatic patients is limited. Evidence from cardiology studies in non-cancer patients suggests that the presence of asymptomatic left ventricular dysfunction alone increases the risk of death and symptomatic congestive heart failure [52]. Despite limited data from clinical trials, population-based data from SEER-Medicare showed that older patients have a higher risk of cardiac events when treated with anthracycline-based adjuvant chemotherapy. Older age, low baseline left ventricular failure, and a history of hypertension are now recognised risk factors for cardiotoxicity from anthracycline and/or trastuzumab [53,54] **(A0006)**.

The risk of developing secondary acute leukaemia was 1-2% at five to ten years follow-up for anthracycline, in particular epirubicin-containing regimens [55]. This risk was directly related to dose. The National Surgical Adjuvant Breast and Bowel Project (NSABP) reported an eight-year cumulative incidence of 0.27% for myelodysplastic syndrome (MDS) and/or acute myelogenous leukaemia (AML) following treatment with adjuvant doxorubicin and cyclophosphamide (AC regimen) [56]. The risk continued to increase beyond five years [55]. However, there is little information on the risk of bone marrow neoplasms with non-anthracycline regimens like docetaxel/cyclophosphamide or docetaxel/capecitabine **(A0006)**.

Health-related quality of life

Health-related QoL is affected by a complex set of trade-offs between preferences on survival benefits and the side effects of adjuvant chemotherapy. It is obvious that the QoL of patients receiving adjuvant chemotherapy will be reduced due to chemotherapy side effects during and shortly after treatment compared to patients who do not receive chemotherapy. However, treatment in the adjuvant setting may provide benefits to the patient other than survival benefits. Important drivers behind negative effects on QoL are anxiety, fear of recurrent disease, and expectations of the treatment outcome [57-59]. Thus, in addition to the trade-offs between expected survival benefits and side effects, patient preference for adjuvant chemotherapy may have an important impact on health-related QoL. To evaluate the overall effect of a medical intervention such as adjuvant chemotherapy, and likewise omitting chemotherapy, an outcome measure that combines survival and QoL is necessary. Quality-adjusted life-years (QALYs), where the quantity of life-years gained is multiplied by a weight reflecting the QoL is such an outcome measure. However, relatively few studies have estimated health-related QoL for breast cancer patients using preference-based measures in the adjuvant setting [60-62].

5 DESCRIPTION AND TECHNICAL CHARACTERISTICS OF TECHNOLOGY

5.1 Research questions

Element ID	Research question
B0001	What is Mammaprint® and what is Adjuvant! Online?
B0002	What is the claimed benefit of MammaPrint® in relation to Adjuvant! Online?
B0004	Who is involved (prescriber, assessor) in applying the technology?
A0020	For which indications has MammaPrint® received market authorisation or CE marking?

5.2 Results

5.2.1 MammaPrint®

In an attempt to better understand breast cancer biology and to find new, clinically-relevant molecular markers, the Netherlands Cancer Institute (NKI) examined global gene expression levels of 78 tumour samples from untreated node-negative breast cancer patients. They used a homogenous population of node-negative patients, all under 55 years of age, all treated with loco-regional therapy alone, and with a median follow-up of 8.7 years to derive their prognostic gene signature. They established a 70-gene expression profile that predicted early distant metastases (recurrence within five years) [63,64]. This gene expression profile is called MammaPrint®, and it is marketed by Agendia (Amsterdam, the Netherlands; http://www.agendia.com). MammaPrint® measures the expression of 70 genes in cancerous breast tissue across seven genomic pathways related to metastatic cascade. Each gene is weighted equally without influence from clinicopathological factors. The measured expression profile is then used in a proprietary algorithm to categorise patients as being at either high of low risk of breast cancer recurrence. As its output, MammaPrint® provides a numerical MammaPrint® index within the range of -1 to +1. This numerical score is overlaid with a binary Low Risk / High Risk clinical classification system. A breast cancer tumour with a MammaPrint® index below or equal to zero is classified as High Risk and a tumour with a MammaPrint® index higher than zero is classified as Low Risk."

MammaPrint® is easy to use for the physician. The procedure is as follows: the MammaPrint® request can be submitted by a physician via the manufacturer's online portal. Core needle biopsy or surgical specimen samples can be submitted.

Specimen Requirements

- Formalin-fixed, paraffin-embedded tissue (FFPE):
  - Specimen block with invasive tumour; OR
  - Ten unstained slides with a five-micron section on each slide.
- Core needle biopsies in RNARetain®. If a 14-gauge needle is used, it is preferable that at least three cores are submitted to increase the probability of tumour-positive (30% tumour cells) biopsies. To minimise sampling failures, one of the cores selected for MammaPrint® should be the first or second core obtained.
- At least 30% invasive tumour in the specimen.

Analysis is performed in one of two central laboratories (the Netherlands and the USA). Physicians receive the test results within ten working days, which reports the classification of the patient's cancer as either high risk or low risk.

MammaPrint® validation

Systematic reviews covering the literature on genetic tests in early breast cancer from 1990 to 2014 concluded that, of the evaluated tests, Oncotype DX® and MammaPrint® were the furthest along their validation pathways **(Table A1, Appendix 1).** In summary, it was concluded that there is no proven benefit from using a GES test. As the vast majority of the

currently available evidence is indirect, direct evidence of the clinical utility of GES tests was still needed. As the comments of Agendia show there are many publications available that provide evidence on test performance, such as analytical validity and the clinical validity.

Claimed benefit of the technology in relation to Adjuvant! Online ***(B0002)***

The primary claim of MammaPrint® is that it can more reliably identify patients who are at low or high risk of distant recurrence, and as such MammaPrint® intends to limit the number of patients with side effects by limiting the number of patients receiving chemotherapy. Agendia claims that patients defined as low risk by MammaPrint® but who are at high risk according to current clinicopathological criteria can safely forgo chemotherapy without deterioration of the clinical outcome. By using MammaPrint®, overtreatment can be prevented. Therefore, fewer women will be unnecessarily exposed to potential toxicity and chemotherapy side effects, thereby improving the overall health and QoL of breast cancer patients.

5.2.2 Adjuvant! Online

Different countries use different risk stratification tools in standard practice. For example, the modified AO!, the NPI, or international clinical guidelines such as those produced by the National Comprehensive Cancer Network (NCCN) or St Gallen are used in Europe. In the MINDACT study, the modified AO! (which includes HER2 status) was used as the comparator risk stratification tool. AO! is a software program (www.adjuvantonline.com) that calculates a ten-year survival probability based on the patient's age, co-morbidities, tumour size, grade, ER status and, recently, also HER2 status. The prognostic model is constructed using risk estimates based on the observed OS of thousands of breast cancer patients recorded in the SEER database, and AO! was validated in over 4000 breast cancer patients from British Colombia [65]. In addition, AO! calculates the absolute survival benefit of any proposed adjuvant therapy by using treatment effect estimates from meta-analyses and RCTs to proportionately adjust its mortality and recurrence rates. Access to AO! is currently disabled because it is being updated to reflect the most recent data. Physicians currently depend on using other guidelines with varying recommendations such as PREDICT. In November 2016, AO! and PREDICT risk models were both endorsed by the American Joint Committee on Cancer for their use in routine practice (AJCC Cancer Staging Manual, 8th edition of TNM classification) [40,66].

5.2.3 Features of the technology and comparators

The features of the technology and comparators are summarised in **Table 5.1**.

Table 5.1 Features of the intervention and comparators

	Technology	**Comparator**
Name	MammaPrint®	Modified Adjuvant! Online
Proprietary name	MammaPrint®	Modified Adjuvant! Online
Manufacturer	Agendia	Adjuvant Inc.
Names in other countries	Not applicable	Not applicable
Class / GMDN code	*In vitro* Diagnostic Directive GMDN code: 60943	Not applicable

5.2.4 Marketing authorisation or CE-marking (A0020)

CE marking

In May 2005, Agendia received CE marking for MammaPrint® in accordance with the CE guidelines for a medical device for *in vitro* diagnostics (directive 98/79/EG) [67]. CE marking was given for both fresh frozen tissue and FFPE tissue. The intended use for which the CE marking was given is: MammaPrint® is a qualitative *in vitro* diagnostic test, performed in a central Agendia laboratory, using the gene expression profile from FFPE breast cancer tissue

samples to assess a patient's risk of distant metastasis. The test is performed on tissue from breast cancer patients with stage I or stage II disease, tumour size ≤5.0 cm, lymph node negative or with up to three positive lymph nodes.

FDA status

The US Food and Drug Administration (FDA) gave *In Vitro* Diagnostic Multivariate Index Assay (IVDMIA) clearance for MammaPrint® in 2007. It was the first cleared molecular test to profile genetic activity. The MammaPrint® test is a laboratory-developed test which falls into the class of IVDMIA. MammaPrint® FFPE received a Predicate Device 510(k) clearance in 2015. The FDA label indicates that MammaPrint® and MammaPrint® FFPE can be used as a prognostic risk stratification tool for early-stage breast cancer patients. The FDA cleared MammaPrint® for marketing with the intended use: MammaPrint® (FFPE and fresh-frozen tissue) is a qualitative in vitro diagnostic test, performed in a central laboratory, using the gene expression profile obtained from FFPE breast cancer tissue samples or fresh breast cancer tissue to assess a patient's risk for distant metastasis within 5 years (FFPE) or up to ten years for patients less than 61 years old, up to 5 years for patients ≥61 years (fresh breast cancer tissue). The genes and scoring algorithm for MammaPrint® FFPE are the same as those used for MammaPrint®, performed with fresh and fresh-frozen tissues. The test is performed for breast cancer patients, with Stage I or Stage II disease, with tumour size ≤ 5.0 cm and lymph node negative. The MammaPrint® FFPE result is indicated for use by physicians as a prognostic marker only, along with other clinicopathological factors. FDA formulated the following special conditions for use statement(s): MammaPrint® FFPE is not indicated as a standalone test to determine the outcome of disease, nor to suggest or infer an individual patient's likely response to therapy. Results should be taken in the context of other relevant clinicopathological factors and standard practice of medicine [68].

6 CLINICAL UTILITY

6.1 Research questions

Element ID	Research question
D0001	What is the expected effect of the intervention on mortality?
D0032	How does the test-treatment intervention modify the magnitude and frequency of morbidity?
D0011	What is the effect of MammaPrint® and treatment on patients' body functions?
D0016	How does MammaPrint® use and treatment affect activities of daily living?
D0012	What is the effect of MammaPrint® and treatment on generic health-related quality of life?
D0013	What is the effect of MammaPrint® and treatment on disease-specific quality of life?
C0008	How safe is the technology in relation to the comparator(s)?
C0004	What are the advantages of not receiving chemotherapy (what harms were prevented)?
C0005	What are the susceptible patient groups that are more likely to be harmed through the use of MammaPrint®?
C0006	What are the consequences of false-positive, false-negative, and incidental findings generated by using the technology from the patient safety viewpoint?
D0017	How many patients follow the treatment advice based on the MammaPrint® result?

6.2 Results

In collaborative EUnetHTA reports, clinical effectiveness and safety outcomes are usually reported in two separate chapters. However, because we are primarily interested in the clinical utility of MammaPrint® all critical endpoints (OS, QoL, and short- and long-term side effects of chemotherapy) are reported in the same section. The critical endpoint of OS *is* in fact the safety endpoint in this assessment.

6.2.1 Included studies

Direct evidence of the clinical utility of the 70-gene signature test MammaPrint® was available from one RCT (MINDACT) [4]. A description of the evidence used is provided in **Section 3.6**.

6.2.2 Study characteristics

6.2.2.1 Study design

The MINDACT study is an international, open-label, randomised controlled, phase 3 trial in women with early-stage breast cancer (n=6693, median follow-up five years). MINDACT compares the MammaPrint® 70-gene signature test added to standard clinicopathological criteria with current clinicopathological risk assignment based on AO!,(using a modified version) in selecting patients for adjuvant chemotherapy. MINDACT was supported by several grants as described in **Table 3.1.** Between 2007 and 2011, 6693 patients were enrolled in 112 institutions in nine European countries. Eligible patients were women between 18 and 70 years of age with histologically-confirmed primary invasive breast cancer (stage T1, T2, or operable T3). The primary objective of MINDACT was to assess whether, in patients with a high clinical risk and a low genomic risk (CH/GL) in whom chemotherapy was omitted, the lower boundary of the 95% CI for the rate of five-year DMFS would be 92% (i.e., the non-inferiority boundary)

or higher at a one-sided significance of 0.025.

As a starting point for determining the non-inferiority threshold in MINDACT, there was consensus in the TRANSBIG Consortium 2014 (a 21-country network associated with the Breast International Group (BIG)) that for every hundred women spared chemotherapy, eight patients will still be at risk and develop metastatic disease. The non-inferiority threshold was derived from a ten-year breast cancer survival probability using AO! (version 7) of 92% for ER-negative tumours to account for the estimated four percentage point absolute benefit from adjuvant endocrine therapy in ER-positive tumours [55,69]. The 4% gain in survival is the benefit that is needed to counterbalance the 2% risk of severe adverse events associated with adjuvant chemotherapy [50,51,55,56]. The TRANSBIG consortium subsequently reached a consensus on a threshold for non-inferiority of 92% at five years (DMFS).

MINDACT's main characteristics are described in **Table 3.1** and **Table A.5** (**Appendix 1**).

Protocol revisions

- In the first version (January 2006) of the protocol, the primary objective of MINDACT was to demonstrate the superiority of the molecular profiling approach over usual clinical assessment in assigning adequate risk categories (and the need to receive adjuvant chemotherapy or not) to node-negative breast cancer patients [63]. In July 2006, the EORTC published a revision of the primary objective: the primary objective of MINDACT was to confirm that patients with a "low risk" molecular prognosis and "high risk" clinical prognosis can be safely spared chemotherapy without affecting DMFS.
- In August 2009, the protocol was revised to allow enrolment of women with up to three positive axillary nodes instead of only lymph node-negative patients.
- A change in the RNA extraction solution used in the assay caused a temporary shift in the risk calculation. This shift resulted in 162 patients originally identified as being at high genomic risk subsequently being reclassified as low genomic risk with the use of the correct solution. The clinical effect of this risk revision was that an additional 28 patients received chemotherapy before the results were corrected, although no patient was undertreated.
- The sample size was modified during the trial from 6000 to 6600 patients, because the proportion of patients designated as low clinical and genomic risk was higher than initially projected because of the need to compensate for the change in RNA extraction solution.

6.2.2.2 Study patients

A total of 11288 patients underwent screening, and 6693 patients were enrolled in the study. Of the 4595 patients (40.7%) who were not enrolled, the main reasons were the unsuitability of tumour material for testing (n=1182; 26%) and a decision by the patient or an investigator not to participate in the study (n=899; 20%). Other reasons for non-enrolment are presented in **Table A7** in **Appendix 1**. There were four main groups (corresponding to corrected risk): low clinical and low genomic risk (CL/GL; n=2745; 41%); low clinical and high genomic risk (CL/GH; n=592; 8.8%); high clinical and low genomic risk (CH/GL; n=1550; 23.2%); and high clinical and high genomic risk (CH/GH; n=1806; 27%). The median age of the patients was 55 years (range 23-71); 79% had node-negative disease, and 20.9% had one to three positive nodes. 88.4% of tumours expressed ER, PR, or both, and 9.5% were HER2-positive. Baseline characteristics are presented in **Table A2** in **Appendix 1**.

6.2.2.3 Study adherence

In the discordant risk groups in the MINDACT trial, overall adherence to chemotherapy assignment was 86%. In CH/GL patients, the adherence rate was 85% in those in the chemotherapy group and 89% in those in the no-chemotherapy group. In CL/GH patients, adherence rates were 80% and 88%, respectively **(D0017)**.

6.2.3 Quality of the study

Different scientific societies and individual countries use different clinical relevance thresholds. As described in **Section 3.8.1**, the ESMO-MCBS threshold of HR 0.80 to determine imprecision in GRADE is used. This HR threshold refers to the lower extreme of the 95% CI.

6.2.3.1 Risk of bias

The risk of bias was rated as high at the study level for the CH/GL population due to the considerable number of patients who were not treated as assigned by randomisation. Of the patients randomised to chemotherapy (n=749), 174 (23%) were not included in the PP analysis (of whom 128 did not receive chemotherapy, 26 had a change in risk, 11 were ineligible, and nine had an unknown chemotherapy status). Of those assigned to no chemotherapy (n=748), 119 (16%) were not included in the PP analysis (of whom 85 received chemotherapy, 21 had a change in risk, 12 were ineligible, and one had an unknown chemotherapy status). However, the baseline patient characteristics of the randomised groups in the PP population were not presented. Therefore, no information could be obtained to assess comparability of baseline characteristics. Additionally, no lost-to-follow-up data was reported and it was an open-label study. The risk of bias for the CL/GH population was low. Although MINDACT was an open-label study, we found downgrading only for this too strict, because the OS, DMFS, and DFS are objective outcome measures. The risk of bias of the study and the different outcome measures are reported in **Table A9** and **Table A10** in **Appendix 1**.

6.2.3.2 GRADE

The complete GRADE summary of finding tables of both discordant groups are included as **Table A11** and **Table A12** in **Appendix 1**.

Clinical high and genomic low (CH/GL)

The quality of the evidence for the CH/GL population for the critical endpoint of ten-year OS was rated as low (with DMFS and DFS used as surrogates) to very low (when direct measurements of OS but only five-year data was used), due to the risk of bias (see **Table A9** and **Table A10** in **Appendix 1**), indirectness (use of a surrogate outcome/follow-up), and imprecision (CIs crossing one or two thresholds of clinical relevance). Therefore, confidence in the OS effect estimate after ten years is limited at best, and the true effect may be substantially different from the effect estimate.

Clinical low and genomic high (CL/GH)

The quality of the evidence for the CL/GH population for the critical endpoint of ten-year OS was rated very low (for both DMFS and DFS surrogate endpoints and also the direct measurement of OS, but only five-year follow-up data was used) due to the risk of bias (see **Table A9** and **Table A10** in **Appendix 1**), indirectness (use of a surrogate outcome), and imprecision (CIs crossing both thresholds of clinical relevance). Therefore, there is very little confidence in the effect estimate, and the true effect is likely to be substantially different from the effect estimate.

6.2.4 Results of the primary EUnetHTA assessment: clinical utility in the CH/GL risk group

From a reimbursement perspective, a diagnostic test added to standard care must have proven added value in health-related outcomes compared to standard care alone. MINDACT's first pre-specified secondary analysis specifically addressed the main EUnetHTA research question. In the case of the CH/GL population, the added value will not be in terms of survival, as MINDACT's primary objective is not superiority but to safely spare chemotherapy in this group (non-inferiority in term of OS). The added value is that patients will potentially experience a better QoL(superiority in terms of QoL and toxicity).

6.2.4.1 Overall survival

Surrogate endpoints for ten-year OS

Data on ten-year OS is not yet available, so surrogate endpoints were considered. First the relation between the surrogate endpoints and the critical endpoint is described. In MINDACT,

the primary endpoint was five-year DMFS as assessed in the time-to-event analyses. Other reported endpoints were five-year OS and five-year DFS. The DMFS definition was the time until the first distant metastatic recurrence or death from any cause. The DFS definition was the time until first disease progression (locoregional, distant relapse, ipsilateral or contralateral invasive breast cancer, ductal carcinoma *in situ*, or a second invasive primary cancer) or death from any cause. DFS also takes into account locoregional recurrences, which still can be treated with curative intent. Regarding the relation between five-year DMFS and ten-year OS, it is known that if distant metastases occur, two out of three women will die of cancer within ten years [11]. For this reason, DMFS after five years predicts the threat to patient survival and so is a surrogate for OS after ten years [70]. However, the relation between ten-year OS and five-year DMFS is not one-on-one. The benefit of adjuvant chemotherapy is primarily limited to reducing recurrences within the first five years. However, there is an on-going risk of distant recurrence after five years. Another frequently used surrogate for ten-year OS is five-year DFS. DFS is a widely accepted surrogate for OS in adjuvant cancer treatments, for example in the European Medicines Agency (EMA) and FDA scientific guidelines. Furthermore, it is a relevant outcome measure for patients, because any recurrence discovered during follow-up (regardless of whether it is curable or not) is a relevant event for a patient [59]. Also for DFS applies that the relation between five-year DFS and ten-year OS is not one-on-one. In this report the results of five-year DMFS, five-year DFS and five-year OS are reported, because it is possible that different countries will choose different surrogate endpoints for ten-year OS (**Table 6.1** and **Figure 6.1**).

PPS versus PP population
Due to the temporary change in risk assessment between May 24, 2009 and January 30, 2010 (as a result of assay problems), all risk groups as enrolled in that particular period are somewhat biased due to incorrect risk assessment. Next to the prespecified PP analysis, also so-called PPS analysis is presented in the MINDACT publication, in which all patients enrolled during the period of change in risk were excluded. This PPS analysis is less biased as the legend of table S5 of the MINDACT publication confirms. This PPS analysis is presented as sensitivity analysis, but in fact this PPS analysis represents the least biased and therefore most conservative PP analysis. Therefore, the results of this PPS population are described in this section. For the sake of completeness the results of the PPS, PP and the ITT analyses are presented in **Table 6.1**.

Results of surrogate endpoints for ten-year OS
The CH/GL patients in the PPS analysis who received chemotherapy after randomisation on the basis of clinical risk had a five-year DMFS rate of 96.5% (95% CI 94.1-97.9), whereas those who did not receive chemotherapy (randomly assigned on the basis of genomic risk) had a rate of 94.0% (95% CI 91.4-95.8), a 2.5% lower rate (HR 0.60; 95% CI 0.34-1.06; p=0.080). Expressed in absolute numbers, omitting chemotherapy after following MammaPrint® risk assessment and not AO! can lead to 100 (of 1000) more patients that will not be free of distant metastases after five years. Looking at the 95% CI of five-year DMFS, it could lead to six less or at worst 287 more patients who are not free of distant metastases after five years (see **Table A11**, **Appendix 1**). Five-year DFS of those who received chemotherapy based on clinical risk was 93.3% (95% CI 90.3-95.4), whereas those who did not receive chemotherapy based on genomic risk had a 4.5% lower rate of 88.8% (95% CI 85.7-91.3). Although the study was not powered to assess statistical differences, the adjusted HR of 0.57 (95% CI 0.37-0.87; p=0.009) was statistical significant. The five-year OS was 1.8% higher in the AO! treated patients with a HR of 0.54 (95% CI 0.23-1.26), but this difference was not significantly different (p=0.154). The 95% CIs of five-year DMFS and five-year OS crossed the non-inferiority threshold of HR 0.80, indicating non-inferiority of MammaPrint® is questionable (like the blue line in **Figure 3.2**). However, the point estimates are on the left side of the non-inferiority threshold (like the red line in **Figure 3.2**)**.** It is therefore reasonable to assume that using MammaPrint® may result in an increased risk of death in comparison with AO! risk assignment (see green line and orange line in **Table 6.1**). Because the difference in five-year DFS is statistically significant different, inferiority is likely but non-inferiority cannot be ruled

out either (like the red line in **Figure 3.2** and see the blue line in **Table 6.1**). For those HTA-organisations that use risk differences in their assessment the same conclusion can be made. The risk differences are 2.5% for five-year DMFS and 4.5% for five-year DFS and 1.8% for five-year OS. The risk difference of five-year DFS (4.5%) crosses the non-inferiority threshold of 3%. Those of DMFS and OS do not cross the non-inferiority threshold, but since in general the confidence intervals are large, it is reasonable to assume that also for five-year DMFS and five-year OS the required non-inferiority is not shown.

As described above, the five-year DFS for treatment decisions based on clinical risk (chemotherapy) and treatment decisions based on MammaPrint® (no chemotherapy) was significantly worse (p=0.009) for MammaPrint® patients. The lower limit of the 95% CI of 0.37 was <0.65, making this difference GRADE A according to ESMO-MCBS criteria and indicating substantial inferiority for those patients in whom the treatment decision was based on MammaPrint®.

Table 6.1 Summary of the results of the surrogate OS endpoints for CH/GL patients

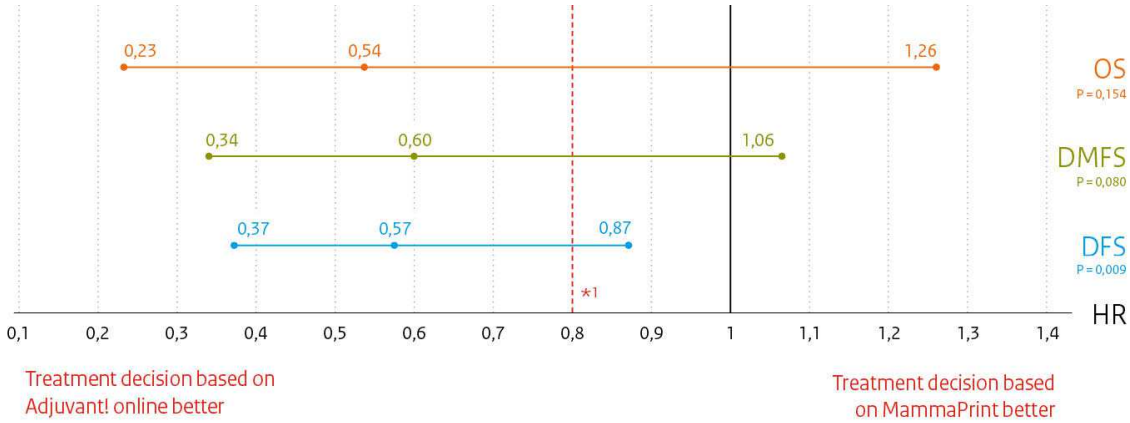
	Patients (n)	HR (95% CI, p-value)	Absolute benefit between treatment based on AO! (receiving chemotherapy) and treatment based on MammaPrint® (omitting chemotherapy)**
PPS*	1045	OS: 0.54 (0.23-1.26, p=0.154) DFS: 0.57 (0.37-0.87, **p=0.009**) DMFS: 0.60 (0.34-1.06, p=0.080)	OS: -1.8% DFS: -4.5% DMFS: -2.5%
PP	1228	OS: 0.63 (0.29-1.37, p=0.25) DFS: 0.64 (0.43-0.95, **p=0.03**) DMFS: 0.65 (0.38-1.10, p=0.11)	OS: -1.5% DFS: -3.0% DMFS: -1.9%
ITT	1497	OS: 0.69 (0.35-1.35, p=0.278) DFS: 0.71 (0.50-1.01, p=0.055) DMFS: 0.78 (0.50-1.21, p=0.267)	OS: -1.5% DFS: -1.9% DMFS: -1.4%

Abbreviations: AO!: Adjuvant! Online; DFS: disease-free survival; DMFS: distant metastasis-free survival; HR: hazard ratio; ITT: intention to treat population; OS: overall survival; PP: per protocol; PPS: per protocol sensitivity.

* The PPS population is the PP population but with all patients enrolled between May 24 2009 and January 30 2010 excluded.

** A minus refers to a benefit of risk assignment based on the AO risk score.

Figure 6.1 Results in the CH/GL risk patients in relation the non-inferiority threshold



Abbreviations: DFS: disease-free survival (including CI's); DMFS: distant metastasis-free survival (including CI's); HR: hazard ratio; OS: overall survival (including CI's).

*1 = non-inferiority threshold of HR 0.80.

6.2.4.2 Quality of life

A better QoL was the expected advantage for patients not receiving chemotherapy based on the MammaPrint® result. However, in MINDACT, long-term modifications in morbidity such as

improved QoL, fatigue, or physical functioning were not measured. However, long-term QoL is not directly measured in the MINDACT trial, it may be argued that some aspects of QoL are reflected by other outcomes. As mentioned in **Section 4.2.6** it is obvious that the QoL of patients receiving adjuvant chemotherapy will be reduced due to chemotherapy side effects during and shortly after treatment compared to patients who do not receive chemotherapy. This benefit in quality of life during the administration period chemotherapy is indirectly known from empirical evidence. In addition, the MINDACT study shows that refraining from chemotherapy leads to a significant and clinically relevant worse five-year DFS. Nearly all types of recurrences are stressful to patients even in the case of a curable disease. This distress will have its repercussions on quality of life. Because QoL is not available in the MINDACT, the added value of the MammaPrint® in terms of QoL in the long term cannot be quantified.

6.2.4.3 Short- and long-term side effects of chemotherapy
Fewer short- and long-term side effects were expected as an advantage of not receiving chemotherapy for patients not receiving chemotherapy based on the MammaPrint® result. Short- and long-term side effects were measured in the MINDACT trial but not published in Cardoso et al [4]. However, it is known that chemotherapy complications include leukaemia, cardiovascular disease, and other side effects as described in **Section 4.2.6**.

Supportive evidence on clinical utility
As described in **Section 3.6**, a description of prospective studies relevant to clinical utility question was permitted when only one RCT was found. However, no controlled observational studies that provided evidence on clinical utility were found.

6.2.5 Results of the secondary analyses

Clinical low risk and genomic high risk (CL/GH)
Ten-year OS (surrogates: five-year DMFS, five-year DFS, and five-year OS)
Ten-year OS data is not yet available, so here five-year DMFS, five-year DFS, and five-year OS as surrogates are reported (**Table 6.2** and **Figure 6.2**). In this case, adjuvant chemotherapy is given based on the MammaPrint® result to patients that would not usually receive chemotherapy based on their clinical risk, so an OS effect would represent the added value. Therefore, ITT population data is presented. CL/GH patients who after randomisation received chemotherapy on the basis of genomic risk had a five-year DMFS rate of 95.8% (95% CI 92.9-97.6), whereas those not receiving chemotherapy (randomly assigned based on clinical risk) had a 0.8% lower five-year DMFS rate of 95.0% (95% CI 91.8-97.0). However, this difference was not statistically significant (p=0.657).

The five-year DFS surrogate endpoint was 92.1% (95% CI 88.3-94.6) for patients who received chemotherapy (on the basis of genomic risk) and 90.1% (95% CI 86.1-93.0) for patients who did not receive chemotherapy (based on clinical risk), 2.0% lower than the rate among those who received chemotherapy based on the MammaPrint® result (p=0.603). For five-year OS, 97.1% (95% CI 94.5-98.5) of patients who received chemotherapy (based on genomic risk) and 97.8% (95% CI 95.5-99.0) of patients who did not receive chemotherapy (based on clinical risk) were still alive, 0.7% lower in those who received chemotherapy based on the MammaPrint® result (p=0.578). None of the surrogate endpoint differences were significant, and added value of MammaPrint® has not been demonstrated at this time.[c] However, the study was not powered to assess significant differences in this discordant group, so a clinical benefit for MammaPrint®-based treatment decision-making cannot be ruled out.

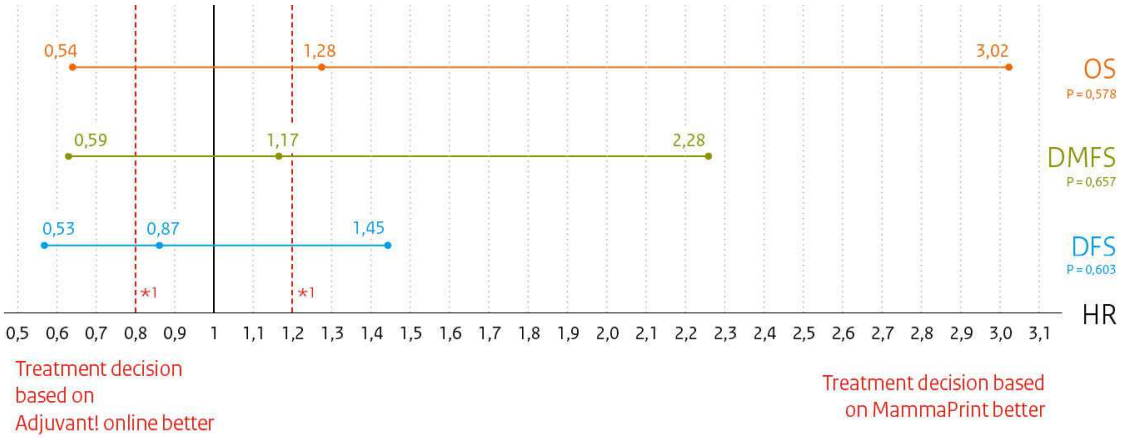
[c] We did not report the HRs of OS and DFS because the HRs do not correspond with the % at five-year DFS and OS.

Table 6.2 Summary of the results of the surrogate endpoints for the CL/GH patients

	Patients (n)	**HR (95% CI, p-value)**	**Absolute benefit between treatment based on MammaPrint® (receiving chemotherapy) and treatment based on AO! (omitting chemotherapy)****
ITT	1497	OS: 1.28 (0.54-3.02, p=0.578) DFS: 0.87 (0.53-1.45, p=0.603) DMFS: 1.17 (0.59-2.28, p=0.657)	OS: -0.7% DFS: 2.0% DMFS: 0.8%

Abbreviations: AO!: Adjuvant! Online; DFS: disease-free survival; DMFS: distant metastasis-free survival; HR: hazard ratio; ITT: intention to treat population; OS: overall survival.
** A minus refers to a benefit of risk assignment based on the AO! risk score, where above zero refers to that risk assignment based on MammaPrint® leading to a benefit.

Figure 6.2 Results in the CL/GH risk patients in relation to threshold for clinical relevance



Abbreviations: DFS: disease-free survival; DMFS: distant metastasis-free survival; HR: hazard ratio; OS: overall survival.
*1 = HR 0.80 and HR 1.20 are the thresholds for clinical relevance.

MammaPrint® used instead of Adjuvant! Online
Outcomes were also assessed in all patients if chemotherapy use had been guided by either clinical risk or genomic risk assignment alone. This is especially relevant for countries where prognostic models such as AO! are not widely accepted, for example if external validation is not available.

50.1% of all patients had a high clinical risk (3356/6693), of whom 1550 had a low genomic risk and 1806 had a high genomic risk. 35.8% (2398/6693) of patients were categorised as being high genomic risk (592 with a low clinical risk and 1806 with a high clinical risk). Thus, the difference between the two strategies for chemotherapy administration would be 14.3% (958 patients). Among all patients at high clinical risk, the use of MammaPrint® to guide chemotherapy treatment would lead to a reduction in the use of adjuvant chemotherapy in 1550 of 3356 patients (46.2%). MINDACT also estimated the outcomes in all patients if chemotherapy use had been recommended by either clinical risk or genomic risk alone. To obtain an unbiased estimate, the discordant patients were doubly weighted because they were underrepresented by a factor of two. Comparison by means of classical statistical inference would be incorrect, so MINDACT only showed the estimated five-year DMFS. At five years, the DMFS would have been 95.0% with the clinical risk strategy alone and 94.7% with the genomic risk strategy alone. The results are presented in **Table A16** in **Appendix 1**.

6.2.6 Results of the pre-specified subgroup analyses
MINDACT also reports exploratory the results of pre-specified subgroups (nodal status, T status, HR+/HER2-LN0). The results of these subgroups are described below. The study authors reported the DMFS after five years only for the ITT population. These subgroup analyses are underpowered, so no definitive conclusions can be made.

Nodal status (CH/GL, ITT population)
For the subgroup of CH/GL patients with lymph node-negative (LN0) status, DMFS was 95.7% (95% CI 93.0-97.4) when the clinical risk profile was followed (chemotherapy) (n=395) and 93.2% (95% CI 90.1-95.4) when genomic risk was followed (no chemotherapy) (n=392). The HR was 0.69 (95% CI 0.39-1.21) (p=0.193).

For the lymph node-positive subgroup, the DMFS was 96.3% (95% CI 93.1–98.1) when clinical

risk was used (chemotherapy) (n=353) and 95.6% (95% CI 92.7-97.4) when genomic risk was used (no chemotherapy) (n=356). The HR was 0.88 (95% CI 0.42-1.82) (p=0.724).

Nodal status (CL/GH, ITT population)
For the subgroup of CL/GH patients with lymph node-negative disease who did not receive (using clinical risk) chemotherapy (n=333), DMFS was 95.1% (95% CI 91.9-97.1) and, for those who received chemotherapy (based on their genomic risk) (n=333), the DMFS was 96.0% (95% CI 93.1-97.7). This difference of 0.9% was not statistically significant (p=0.815). The HR was 1.09 (95% CI 0.54-2.19).

The lymph node-positive group (CL/GH) was too small for analysis.

Tumour size (CH/GL, ITT population)
For the subgroup of CH/GL patients with a tumour larger than 2 cm but no larger than 5 cm (T2), DMFS was 94.5% (95% CI 91.4-96.6) when the clinical risk profile was followed (chemotherapy) (n=402) and 93.7% (95% CI 90.6-95.8) when the genomic risk was followed (no chemotherapy) (n=406). This difference of 0.8% was not statistically significant (p=0.706). The HR was 0.90 (95% CI 0.53-1.54).

For the subgroup of CH/GL patients with a tumour smaller than 2 cm (T1), DMFS was 97.6% (95% CI 95.1–98.9) when clinical risk was used (chemotherapy) (n=322) and 94.8% (95% CI 91.5-96.8) when genomic risk was used (no chemotherapy) (n=314). This difference of 2.8% was not statistically significant (p=0.201). The HR was 0.59 (95% CI 0.26-1.33).

Tumour size (CL/GH, ITT population)
For the subgroup of CL/GH patients with a tumour smaller than 2 cm (T1), DMFS was 94.9% (95% CI 91.6-96.9) when the clinical risk profile was followed (no chemotherapy) (n=338) and 95.7% (95% CI 92.6-97.5) when genomic risk was followed (chemotherapy) (n=333). This difference of 0.8% was not statistically significant (p=0.851). The HR was 1.07 (95% CI 0.54-2.12).

The subgroup of CL/GH patients with T2 tumours was too small to be analysed.

HR+/HER2-/LN0 (CH/GL, ITT population)
For the subgroup of CH/GL patients with hormone receptor-positive (HR+), human epidermal growth factor receptor 2 negative (HER2-) and lymph node-negative (LN0) status, the DMFS was 95.5% (95% CI 92.5-97.3) when the clinical risk profile was followed (chemotherapy) (n=349) and 93.9% (95% CI 90.6-96.1) when genomic risk was followed (no chemotherapy) (n=350). This difference of 1.6% was not statistically significant (p=0.456). The HR was 0.80 (95% CI 0.44-1.45).

HR-/HER2+/LN0 (CL/GH, ITT population)
For the subgroup of CL/GH patients with hormone receptor-negative (HR-), human epidermal growth factor receptor 2 positive (HER2+) and lymph node-negative (LN0) status, the DMFS was 95.5% (95% CI 91.6-97.6) when the clinical risk profile was followed (no chemotherapy) (n=262) and 95.1% (95% CI 91.5-97.2) when genomic risk was followed (chemotherapy) (n=272). This difference of 0.4% was not statistically significant (p=0.333). The HR was 1.45 (95% CI 0.68-3.08).

7 DISCUSSION

From a reimbursement perspective, a diagnostic test added to standard care must have proven added value in health-related outcomes compared to standard care alone. It has been claimed that MammaPrint® has a substantial, positive effect on the health and well-being of women with early breast cancer by limiting the number of patients receiving adjuvant chemotherapy and, as a consequence, related adverse events, without negatively affecting overall survival. For this purpose, the MINDACT authors predefined a non-inferiority threshold as the cut-off for the benefit of using MammaPrint® for the decision of administering chemotherapy: among CH/GL patients who did not receive chemotherapy, the lower boundary of the CI of the five-year DMFS should be 92% or higher.

According to the MINDACT study authors, five-year DMFS was not negatively affected in the primary test population of women with the CH/GL risk profile when the genomic profile (MammaPrint®) was followed. However, our assessment concludes that the current data is insufficient to determine that it is safe to omit chemotherapy and insufficient to determine that there is added value in the CH/GL population of early breast cancer patients. Our reasoning is as follows.

7.1 Non-inferiority threshold and magnitude of clinical relevance

From a reimbursement perspective, it is necessary that new and standard approaches are compared because added value must be proven when a test or intervention is added to standard care. Also, the IQWiG concluded that the risk of distant metastases in women in the groups with and without chemotherapy should have been compared instead of evaluating only one arm. Bogaerts et al. [69] mentioned this issue as a major criticism of the primary analysis of MINDACT. Despite the aim to prove non-inferiority for the five-year DMFS endpoint, MINDACT does not have a formal non-inferiority design, and a trial with a non-inferiority design would need to be very large or of extremely long duration. Bogaerts et al. [69] proposed that if the primary test is significant and the gene signature selects fewer patients to be treated with chemotherapy while not adversely affecting five-year DMFS, then this can be taken as equivalent to proving that the signature has very good sensitivity and a specificity that is better than the clinicopathological method.

This assumption would be acceptable in the situation that the five-year DMFS of the CH/GL subgroup who received chemotherapy was reliably known. However, this subgroup could not be selected from the SEER database. Since this information was not available at the start of the MINDACT study and a threshold had to be prespecified, the TRANSBIG consortium members decided on a non-inferiority threshold of 92% derived from a ten-year breast cancer survival probability using Adjuvant! Online. This choice of non-inferiority threshold is rational. However, different thresholds could have been chosen as decisions with respect to the use of adjuvant chemotherapy are subjective and highly variable among patients, physicians, scientific and HTA-organisations. In their comment on MINDACT, Thewes et al. [71] suggested that most patients with breast cancer are willing to accept adjuvant chemotherapy for very small survival gains (≤1%). Hamelinck et al. [72] also concluded that most patients judged small to moderate benefits sufficient to consider adjuvant systemic therapy worthwhile, but individual preferences varied widely. Taken together, the 92% five-year DMFS non-inferiority boundary of the MINDACT trial can be regarded as unconventional. Furthermore, the lower boundary of the 95% CI of the PPS population was 91.4%, which is below the non-inferiority threshold of 92% defined by the MINDACT authors. The PPS population provides the least biased and most conservative estimate of OS.

The question remains which threshold should instead be used to determine if omitting chemotherapy is safe. Because the five-year DMFS of the CH/GL subgroup who received chemotherapy was not reliably known, a direct comparison between the two arms is needed. Overall, there is no consensus what non-inferiority threshold should be used to determine

when a difference between the two arms is unacceptable. Since this is a European assessment, the ESMO-MCBS clinical relevance thresholds were used [2]. Based on the ESMO-MCBS, a non-inferiority threshold of HR 0.80 or a risk difference of 3% is used. It can be argued that these criteria were primarily intended to stratify the magnitude of clinical benefit that can be anticipated from anticancer therapies. However, this conventional method of selecting a non-inferiority threshold is often used as concluded by Tanaka et al [14] and in accordance with non-inferiority thresholds used in non-inferiority trials in early breast cancer. In addition, even though MammaPrint® is a diagnostic test, the added value of the 'test plus treatment' on health related outcomes needs to be proven.

In the first prespecified secondary MINDACT analyses, outcomes were compared in patients in the discordant risk groups according to whether they were assigned to the chemotherapy group or the non-chemotherapy group. This is the direct comparison that is of primary importance for reimbursement decisions. The MINDACT authors stated that the five-year DMFS (the primary study endpoint) was not significantly different. Because the study was not sufficiently powered to assess these differences, this finding should not be interpreted as evidence of absence of a therapeutic effect [76].

Furthermore, since the PPS analysis of five-year DFS was significant (p=0.009), it is assumed that the investigated group was large enough to reveal an effect in DFS even without the power calculation targeting this secondary analysis. While the possibility of a chance finding always exists, since all outcome measures point in the same direction, it is doubtful that this effect arose by chance. Instead, we believe that this effect reveals a true difference between the two groups. This treatment effect may become more pronounced over the next five years because more events ((distant) recurrences and deaths) will occur.

Based on the MINDACT it cannot be concluded that it is safe to omit chemotherapy, because the 95% CI's of all surrogate outcomes (five-year DMFS, five-year DFS and five-year OS) for ten-year OS are crossing the non-inferiority threshold (HR 0.80 and 3% risk difference). However, the point estimates are on the left side of the non-inferiority threshold. It is therefore reasonable to assume that using MammaPrint® may result in an (clinically-relevant) increased risk of death in comparison with AO! risk assignment.

The ESMO-MCBS criteria have been criticised [73-75]. ESMO uses the lower boundary of the 95% CI to determine if a statistical significant difference is clinically relevant, so small trials qualify more easily for efficacy than large trials of identical efficacy, since small trials generally have wide 95% CIs. It has been suggested that the point estimate should be used instead [74,75]. However, even the point estimate of five-year DFS of 0.57 crosses the ESMO-MCBS threshold of 0.80 and even the HR 0.65 threshold, suggesting a substantial level of clinical benefit for DFS from chemotherapy following clinical risk assignment. Another critique is that, in the ESMO-MCBS, surrogates as primary endpoints have equal weight to the patient-relevant endpoints of OS and QoL. The effect can still be considered clinically relevant, but the ten-year survival data is mandatory to definitively confirm that the effect is clinically relevant. Furthermore, no rational arguments are provided for the threshold values [75]. Another criticism is that the credibility of the ESMO-MCBS would be augmented by external validation, for example comparing its results with those of HTAs carried out in Europe [73]. Cherney at al. [76,77] reported that this form of correlative validation has been carried out and shows a very high degree of concordance with the published evaluation of the same agents using the ESMO-MCBS. They suggested that the methodology incorporated into the ESMO-MCBS is remarkably consistent with EUnetHTA guidelines for outcome measures, surrogate endpoints, health-related QoL, and the application of these outcomes in relative effectiveness evaluations. An adapted version of the ESMO-MCBS criteria suggested by Wild et al. [75] is already used in HTA assessments.

Despite the lack of consensus for using the ESMO-MCBS criteria, using no loss in OS as a non-inferiority threshold or even other thresholds for non-inferiority (**Table 3.2**) will lead to the

same conclusion: the 95% CI's of all surrogate outcomes cross these thresholds or it is reasonable to expect that these thresholds will be crossed in the next five years. The IQWiG also came to this conclusion [11]. Therefore, non-inferiority in terms of survival cannot be concluded. However, each HTA-organisation, needs to decide individually which non-inferiority threshold and clinical relevant thresholds are considered best for their assessment.

7.2 PP or PPS analysis

Due to the temporary change in risk as a result of assay problems, all risk groups as enrolled in that particular period are somewhat biased due to incorrect risk assessment. In addition to the prespecified PP analyses, so-called PPS analysis is also presented in the MINDACT publication, in which all patients enrolled during the period of change in risk were excluded. This PPS analysis is the least biased as the legend of table S5 of the MINDACT publication confirms. This PPS analysis is presented as sensitivity analysis, but in fact this PPS analysis represents the least biased and therefore most conservative PP analysis. Because the supplement of the MINDACT in which the PPS analysis was presented may not have undergone peer review, it could be argued that it should not be used as primary analysis in this assessment. But even if it is used in the way it is presented, i.e. as a sensitivity analysis, it is of complementary and confirmative information in order to assess the robustness of the findings and herewith an important way to assess the final impact of the study results for clinical practice. Unfortunately, the PPS analysis points in the direction in which the MammaPrint® group scores worse, thereby casting doubt on the robustness of the prespecified PP analysis. In addition, when the results of the PP analysis were used, instead of the PPS, the conclusion on clinical utility of the MammaPrint® would have been the same.

7.3 Surrogate endpoints for ten-year OS

In general there is no consensus on the use of surrogate endpoints to assess (added) clinical benefit of a health technology, because the relationship between a patient-relevant clinical endpoint and its various surrogates has rarely been investigated in such depth that one particular surrogate is universally accepted as a replacement [78]. Each country/HTA organisation, needs to decide individually which surrogate endpoint is considered best for their assessment. Therefore, the relevance of each endpoint is described, based on its biological plausibility and empirical evidence, providing specific information relevant for each endpoint. In summary: DMFS has a biological rationale and the MammaPrint® is developed as predictor for five-year DMFS, whereas DFS is also a relevant patient related outcome, as it includes all types of recurrences. Therefore, DFS reflects for the patient all stressful events, even when the disease is still curable, potentially impacting quality of life [59].

MINDACT's primary endpoint is five-year DMFS in the CH/GL discordant risk group. According to the study authors, five-year DMFS is the primary endpoint as distant metastasis from breast cancer represents a virtually incurable disease with almost 100% mortality. The MINDACT authors stated that five-year DMFS was not significantly different and that the five-year results can be considered as mature data. However, as noted above, the DMFS results have a wide CI, as do the other endpoints. Expressed in absolute numbers, omitting chemotherapy after following MammaPrint® risk assessment and not AO! can lead to 100 (of 1000) more patients that will not be free of distant metastases after five years. Looking at the 95% CI of five-year DMFS, it could also lead to six less or at worst 287 more patients who are not free of distant metastases after five years (see **Table A11** in **Appendix 1**). The 95% CI shows that there is a lot of uncertainty and a possibility that many patients could be harmed. Furthermore, there is an on-going risk of distant recurrence after five years. This is especially the case in ER-positive/HER2-negative (luminal-type) breast cancer, where recurrences occur after five years in approximately one half of all distant recurrence cases [23]. Since ER-positive/HER2-negative tumours comprise the majority of the MINDACT study population, many distant metastasis cases might be expected. The benefit of adjuvant chemotherapy is primarily limited to

reducing recurrences within the first five years, however it is unclear whether these late recurrences are prevented by adjuvant chemotherapy [31-33]. Therefore, ten-year follow-up data is necessary.

7.4 Quality of life/short- and long-term side effects of chemotherapy

It is generally recognised that OS is the least ambiguous and most clinically relevant endpoint in clinical trials for cancer therapy. Beyond OS, the QoL endpoint is also very relevant. Unfortunately, QoL was not included in the MINDACT trial. According to the investigators, adding QoL questionnaires would be too burdensome for patients as they had to comprehend the complexities of the trial, including information on genomic testing. In addition, according to the investigators, no validated instrument was available at the time of study. Therefore, the TRANSBIG consortium, which involved patients and advocates, decided not to include a QoL evaluation in the overall study population.

Although long-term QoL is not directly measured in the MINDACT trial, it may be argued that some aspects of QoL are reflected by other outcomes. As mentioned in **Section 4.2.6** it is known that the QoL of patients receiving adjuvant chemotherapy will be reduced due to chemotherapy side effects during and shortly after treatment compared to patients who do not receive chemotherapy. In addition, the MINDACT study shows that refraining from chemotherapy leads to a significant and clinically-relevant worse five-year DFS. Nearly all types of recurrences are stressful to patients even in the case of a curable disease. This distress will have its repercussions on quality of life. Retel et al. [13] conclude in their QoL assessment 6-8 weeks after their decision regarding adjuvant chemotherapy that patients were generally satisfied with the information they received about recurrence risk based on the MammaPrint®, but clinicians should be aware that genomic test results may be associated with greater distress levels, especially for patients with high recurrence risk or discordant test results. Because long-term QoL is not available in the MINDACT, the added value of the MammaPrint® in terms of QoL in the long term cannot be quantified.

Toxicity data is measured in the MINDACT trial but are not yet published. When considering the protocol, future analyses will be limited to a comparison between side effects of the two regimens of chemotherapies and endocrine therapy. At this time, it is only known from previous publications that chemotherapy has an absolute risk of heart failure or leukaemia of approximately 2% [79,80]. The absence of data on toxicity will not be critical when data on OS and QoL would be available, as toxicity will have its repercussions on QoL and/or OS.

7.5 Clinical utility of MammaPrint®

Each HTA-organisation will need to decide individually which non-inferiority threshold and clinical relevant thresholds are considered most appropriate for their assessment considering the local context. Based on the chosen thresholds in this EUnetHTA assessment report, conclusions are as follows:

Taking everything into consideration, it has not yet been demonstrated that patient outcomes (ten-year OS and QoL) are improved by withholding adjuvant chemotherapy based on MammaPrint® testing in the CH/GL risk group (**Table 7.1**). In other words, the clinical utility of the MammaPrint® is not proven. This conclusion is based on the absence of evidence on added value in terms QoL and on the fact that non-inferiority in terms of OS (surrogates five-year DMFS, five-year DFS and five-year OS) is not shown. Next to that there are concerns about the certainty of five-year DMFS because of the imprecision (very wide 95% CI's). Therefore the results do not rule out the possibility of a clinically-relevant increase in distant metastasis and hence risk of death. Also, the significant and clinically-relevant difference in DFS is of importance as QoL data is not available. The quality of the evidence for the critical ten-year OS endpoint was rated as low to very low. Therefore, the confidence in the OS effect estimate after ten years is limited at best.

Furthermore, a clinical benefit of receiving chemotherapy in the CL/GH risk group according to genomic risk assignment has not been demonstrated. The quality of the evidence for the critical ten-year OS endpoint was rated very low. Therefore, there is very little confidence in the effect estimate, and the true effect is likely to be substantially different from the effect estimate.

Ten years of follow-up will be needed to conclude if an add-on risk assessment approach (AO! combined with MammaPrint®) has superior clinical utility compared with treatment decisions based on AO! alone. At that time, the level of evidence of the data and therefore confidence in the data will be higher, because the use of surrogate endpoints will no longer be necessary. If a revision of AO! becomes available, as is expected, this could have an impact on the baseline risks of recurrence and hence may potentially limit the clinical applicability of the MINDACT results.

Ultimately, the decision to receive or forgo chemotherapy (or any other treatment) lies with each patient who is properly informed about the potential side effects and the potential benefits of such treatment. For the same risk–benefit scenario, different patients may make different decisions. However, well-informed decision making is only possible if both parameters (OS and QoL) had been quantified.

Table 7.1 Clinical utility of MammaPrint® in the different risk groups

Risk group	Add-on test (MammaPrint®) results in MINDACT	Clinical utility (five-year follow-up)
CH/GL	Following MammaPrint® potentially leads to significant and clinically relevant decreased survival (five-year DFS absolute difference of 4.5% HR=0.57 (95% CI 0.37-0.87; p=0.009)	Not proven, results indicate potentially inferiority of treatment decision based on MammaPrint®
CH/GH	Concordant risk group, treatment decision is chemotherapy	Clinical utility not relevant
CL/GH	Clinical benefit of genomic risk assignment not demonstrated but a clinical benefit of MammaPrint® based treatment decision cannot be ruled out	Not proven
CL/GL	Concordant risk group, treatment decision is no chemotherapy	Clinical utility not relevant

Abbreviations: CH: clinical high; CL: clinical low; DFS: disease-free survival; GH: genomic high; GL: genomic low; HR: hazard ratio.

8 REFERENCES

1. Joensuu H, Pylkkanen L, Toikkanen S. Late mortality from pT1N0M0 breast carcinoma. *Cancer* 1999;85:2183-9.

2. Cherny NI, Sullivan R, Dafni U, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann Oncol* 2015;26:1547-73.

3. de Boer M, Voogd AC, Tjan-Heijnen VCG. Pleidooi voor selectief gebruik van de MammaPrint. *Ned Tijdschr Geneeskd* 2017;161:D1160.

4. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med* 2016;375:717-29.

5. KCE. Gene expression profiling and immunohistochemistry tests for personalised management of adjuvant chemotherapy decisions in early breast cancer – a Rapid Assessment. 2015. [Cited June 2017]. https://kce.fgov.be/sites/default/files/page_documents/KCE_237_Gene%20expression%20profiling_Report.pdf.

6. Marchionni L, Wilson RF, Marinopoulos SS, et al. Impact of gene expression profiling tests on breast cancer outcomes. *Evid Rep Technol Assess* (Full Rep) 2007;1-105.

7. Smartt P. A Comparison of Gene Expression Profiling Tests for Breast Cancer. 2010;

8. Ward S, Scope A, Rafia R, et al. Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2013;17:1-302.

9. Medical Advisory Secretariat / Health Quality Ontario. Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast cancer: an evidence-based and economic analysis. *Ont Health Technol Assess Ser* 2010;10:1-57.

10. Marrone M, Stewart A, Dotson WD. Clinical utility of gene-expression profiling in women with early breast cancer: an overview of systematic reviews. *Genet Med* 2015;17:519-32.

11. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Biomarker-based tests to support the decision for or against adjuvant systemic chemotherapy in primary breast cancer. 2017;

12. Drukker CA, Bueno-de-Mesquita JM, Retel VP, et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. *Int J Cancer* 2013;133:929-36.

13. Retel VP, Groothuis-Oudshoorn CGM, Aaronson NK, et al. Association between genomic recurrence risk and well-being among breast cancer patients. *BMC Cancer* 2013;13:295.

14. Tanaka S, Kinjo Y, Kataoka Y, et al. Statistical issues and recommendations for noninferiority trials in oncology: a systematic review. *Clin Cancer Res* 2012;18:1837-47.

15. Over de adviezen. PASKWIL criteria. 2017;

16. Systematische beoordeling van oncologische middelen. *Medische Oncologie*

2012;15(1):26-9.

17. European Commission. Cancer Statistics. 2017. [Cited November 2017]. http://ec.europa.eu/eurostat/statistics-explained/index.php/Cancer_statistics.

18. Jatoi I, Miller AB. Why is breast-cancer mortality declining? *Lancet Oncol* 2003;4:251-4.

19. Smigal C, Jemal A, Ward E, et al. Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin* 2006;56:168-83.

20. American Cancer Society. Cancer facts & Figures 2010. 2010;

21. SEER Cancer Statistics Review (CSR) 1975-2014. Bethesda, MD: National Cancer Institute, 2017. [Cited November 2017]. https://seer.cancer.gov/csr/1975_2014/.

22. Cancer Factsheets: Breast Cancer. International Agency for Research on Cancer, 2017. http://eco.iarc.fr/EUCAN/CancerOne.aspx?Cancer=46&Gender=2#block-table-f.

23. Sparano JA, Zhao F, Martino S, et al. Long-Term Follow-Up of the E1199 Phase III Trial Evaluating the Role of Taxane and Schedule in Operable Breast Cancer. *J Clin Oncol* 2015;33:2353-60.

24. Chia SK, Speers CH, Bryce CJ, et al. Ten-year outcomes in a population-based cohort of node-negative, lymphatic, and vascular invasion-negative early breast cancers without adjuvant systemic therapies. *J Clin Oncol* 2004;22:1630-7.

25. Demicheli R, Abbattista A, Miceli R, et al. Time distribution of the recurrence risk for breast cancer patients undergoing mastectomy: further support about the concept of tumor dormancy. *Breast Cancer Res Treat* 1996;41:177-85.

26. Leitner SP, Swern AS, Weinberger D, et al. Predictors of recurrence for patients with small (one centimeter or less) localized breast cancer (T1a,b N0 M0). *Cancer* 1995;76:2266-74.

27. Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst* 2011;103:1656-64.

28. Guiu S, Michiels S, Andre F, et al. Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement. *Ann Oncol* 2012;23:2997-3006.

29. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol* 2007;25:5287-312.

30. Sobin LH, et al., (eds) . TNM Classification of Malignant Tumours, 7th Edition. 7. Wiley-Blackwell, 2011.

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

32. Clarke M, Coates AS, Darby SC, et al. Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. *Lancet* 2008;371:29-40.

33. Peto R, Davies C, Godwin J, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women

in 123 randomised trials. *Lancet* 2012;379:432-44.

34. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012;486:346-52.

35. Dvinge H, Git A, Graf S, et al. The shaping and functional consequences of the microRNA landscape in breast cancer. *Nature* 2013;497:378-82.

36. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747-52.

37. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci* U S A 2001;98:10869-74.

38. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.

39. Curigliano G, Burstein HJ, Winer P, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol* 2017;28:1700-12.

40. Amin M ea, (ed.). AJCC Cancer Staging Manual, 8th edition. American Joint Committee on Cancer 2017, 2017.

41. Krop I, Ismaila N, Andre F, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Focused Update. *J Clin Oncol* 2017;35:2838-47.

42. Senkus E, Kyriakides S, Penault-Llorca F, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 Suppl 6:vi7-23.

43. Regan MM, Pagani O, Walley B, et al. Premenopausal endocrine-responsive early breast cancer: who receives chemotherapy*? Ann Oncol* 2008;19:1231-41.

44. Coates AS, Millar EKA, O'Toole SA, et al. Prognostic interaction between expression of p53 and estrogen receptor in patients with node-negative breast cancer: results from IBCSG Trials VIII and IX. *Breast Cancer Res* 2012;14:R143.

45. Coates AS, Winer EP, Goldhirsch A, et al. Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol* 2015;26:1533-46.

46. de Glas NA, van de Water W, Engelhardt EG, et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. *Lancet Oncol* 2014;15:722-9.

47. Azim HAJ, de Azambuja E, Colozza M, et al. Long-term toxic effects of adjuvant chemotherapy in breast cancer. *Ann Oncol* 2011;22:1939-47.

48. Zambetti M, Moliterni A, Materazzo C, et al. Long-term cardiac sequelae in operable breast cancer patients given adjuvant chemotherapy with or without doxorubicin and breast irradiation. *J Clin Oncol* 2001;19:37-43.

49. Bonneterre J, Roche H, Kerbrat P, et al. Epirubicin increases long-term survival in adjuvant chemotherapy of patients with poor-prognosis, node-positive, early breast cancer: 10-year follow-up results of the French Adjuvant Study Group 05 randomized trial. *J Clin Oncol*

2005;23:2686-93.

50. Perez EA, Suman VJ, Davidson NE, et al. Cardiac safety analysis of doxorubicin and cyclophosphamide followed by paclitaxel with or without trastuzumab in the North Central Cancer Treatment Group N9831 adjuvant breast cancer trial. *J Clin Oncol* 2008;26:1231-8.

51. Romond EH, Jeong JH, Rastogi P, et al. Seven-year follow-up assessment of cardiac function in NSABP B-31, a randomized trial comparing doxorubicin and cyclophosphamide followed by paclitaxel (ACP) with ACP plus trastuzumab as adjuvant therapy for patients with node-positive, human epidermal growth factor receptor 2-positive breast cancer. *J Clin Oncol* 2012;30:3792-9.

52. Yusuf S, Pitt B, Davis CE, et al. Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. *N Engl J Med* 1992;327:685-91.

53. Doyle JJ, Neugut AI, Jacobson JS, et al. Chemotherapy and cardiotoxicity in older breast cancer patients: a population-based study. *J Clin Oncol* 2005;23:8597-605.

54. Pinder MC, Duan Z, Goodwin JS, et al. Congestive heart failure in older women treated with adjuvant anthracycline chemotherapy for breast cancer. *J Clin Oncol* 2007;25:3808-15.

55. Wolff AC, Blackford AL, Visvanathan K, et al. Risk of marrow neoplasms after adjuvant breast cancer therapy: the national comprehensive cancer network experience. *J Clin Oncol* 2015;33:340-8.

56. Smith RE. Risk for the development of treatment-related acute myelocytic leukemia and myelodysplastic syndrome among patients with breast cancer: review of the literature and the National Surgical Adjuvant Breast and Bowel Project experience. *Clin Breast Cancer* 2003;4:273-9.

57. Lidgren M, Wilking N, Jonsson B, et al. Health related quality of life in different states of breast cancer. *Qual Life Res* 2007;16:1073-81.

58. Jansen SJ, Kievit J, Nooij MA, et al. Patients' preferences for adjuvant chemotherapy in early-stage breast cancer: is treatment worthwhile? *Br J Cancer* 2001;84:1577-85.

59. Warren M. Uncertainty, lack of control and emotional functioning in women with metastatic breast cancer: a review and secondary analysis of the literature using the critical appraisal technique. *Eur J Cancer Care* (Engl) 2010;19:564-74.

60. Haas JS, Phillips KA, Liang SY, et al. Genomic testing and therapies for breast cancer in clinical practice. *J Oncol Pract* 2011;7:e1s-e7s.

61. Conner-Spady BL, Cumming C, Nabholtz JM, et al. A longitudinal prospective study of health-related quality of life in breast cancer patients following high-dose chemotherapy with autologous blood stem cell transplantation. *Bone Marrow Transplant* 2005;36:251-9.

62. Earle CC, Chapman RH, Baker CS, et al. Systematic overview of cost-utility assessments in oncology. *J Clin Oncol* 2000;18:3302-17.

63. EORTC 10041 BIG 3-04 Intergroup Study. 2005. http://www.nejm.org/doi/suppl/10.1056/NEJMoa1602253/suppl_file/nejmoa1602253_protocol.pdf.

64. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical

outcome of breast cancer. *Nature* 2002;415:530-6.

65. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001;19:980-91.

66. PREDICT: What's New. Public Health England and Cambridge University, 2017. [Cited November 2017]. http://www.predict.nhs.uk/whatsnew.html.

67. Declaration of Conformity. Agendia, 2015. http://www.agendia.com/media/2-MP-DOC-06May2015.pdf.

68. Agendia's FDA clearances for MammaPrint (k062694, k070675, k080252, k081092, k101454, k141142). 2017;

69. Bogaerts J, Cardoso F, Buyse M, et al. Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. *Nat Clin Pract Oncol* 2006;3:540-51.

70. Hudis CA, Barlow WE, Costantino JP, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol* 2007;25:2127-32.

71. Thewes B, Prins J, Friedlander M. 70-Gene Signature in Early-Stage Breast Cancer. *N Engl J Med* 2016;375:2199-200.

72. Hamelinck VC, Bastiaannet E, Pieterse AH, et al. Patients' preferences for surgical and adjuvant systemic treatment in early breast cancer: a systematic review. *Cancer Treat Rev* 2014;40:1005-18.

73. Hartmann M. The ESMO magnitude of clinical benefit scaling tool: from theory to practice. *Ann Oncol* 2015;26:2357-8.

74. Muhonen T, Joensuu H, Pfeiffer P. Comment on ESMO Magnitude of Clinical Benefit Scale. *Ann Oncol* 2015;26:2504.

75. Wild C, Grossmann N, Bonanno PV, et al. Utilisation of the ESMO-MCBS in practice of HTA. *Ann Oncol* 2016;27:2134-6.

76. Cherny NI, Sullivan R, Dafni U, et al. Reply to the letter to the editor 'The ESMO Magnitude of Clinical Benefit Scaling Tool: from theory to practice' by Hartmann and the letter 'Comment on ESMO Magnitude of Clinical Benefit Scale' by Muhonen et al. *Ann Oncol* 2016;27:203-5.

77. Cherny NI, Sullivan R, Dafni U, et al. Reply to the letter to the editor 'Utilisation of the ESMO-MCBS in practice of HTA' by Wild et al. *Ann Oncol* 2016;27:2136-7.

78. EUnetHTA methodological guideline for rapid relative effectiveness assessment (REA) of Pharmaceuticals – surrogate endpoints. 2013. [Cited November 2017]. http://www.eunethta.eu/outputs/methodological-guideline-rea-pharmaceuticals-surrogate-endpoints.

79. Colozza M, de Azambuja E, Cardoso F, et al. Breast cancer: achievements in adjuvant systemic therapies in the pre-genomic era. *Oncologist* 2006;11:111-25.

80. Park BW, Lee S, Lee AR, et al. Quality of Life Differences between Younger and Older Breast Cancer Patients. *J Breast Cancer* 2011;14:112-8.

81. Ontario Health Technology Advisory Committee. OHTAC Recommendation: Multi-gene expression profiling for guiding adjuvant chemotherapy decisions in women with early breast

cancer . 2010;

82. Ignatiadis M, Bedard P, Haibe-Kains B, et al. A Meta-Analysis of Gene Expression Profiling Studies Identifies Clinically Relevant Oncogenic Pathways in Basal-Like Breast Cancer. *Cancer Research* 2009;69:Abstract nr 106.

83. National Comprehensive Cancer Network. Breast cancer. 2017. [Cited June 2017]. https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf.

84. AETSA. Pruebas genómicas para el pronóstico de pacientes con cáncer de mama. MammaPrint® y Oncotype DX®. Prognostic genomic tests in early breast: MammaPrint® and Oncotype DX®. 2014. [Cited June 2017]. http://www.aetsa.org/download/publicaciones/AETSA_2013_3_Mammaprint.pdf.

85. Cancer Care Ontario. Clinical utility of multigene profiling assays in early-stage breast cancer . 2016. https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=362126.

86. AHRQ. Technology Assessment of Molecular Pathology Testing for the Estimation of Prognosis for Common Cancers . 2014. [Cited June 2017]. https://www.cms.gov/Medicare/Coverage/DeterminationProcess/Downloads/id94TA.pdf.

87. Institute Nationale du Cancer. Tests prédictifs dans le cancer du sein. 2012. [Cited November 2017]. http://www.e-cancer.fr/Professionnels-de-sante/Recommandations-et-outils-d-aide-a-la-pratique/Cancers-du-sein/Tests-predictifs-dans-le-cancer-du-sein.

88. NICE. Gene expression profifiling and expanded immunohistochemistry tests for guiding adjuvant chemotherapy decisions in early breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat. 2013. [Cited June 2017]. http://tinyurl.com/y8fhpx6o.

89. DKG. Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms. 2012. [Cited June 2017]. http://www.awmf.org/uploads/tx_szleitlinien/032-045OL_l_S3__Brustkrebs_Mammakarzinom_Diagnostik_Therapie_Nachsorge_2012-07.pdf.

90. NABON. Mammacarcinoom. 2012. [Cited November 2017]. http://www.oncoline.nl/mammacarcinoom.

91. Unidad de Evaluacion de Tecnologias Sanitarias (UETS). Evaluación económica de las pruebas genéticas en el tratamiento del cáncer de mama. 2012. [Cited June 2017]. http://tinyurl.com/y6wj7rkg.

92. Duffy MJ, Harbeck N, Nap M, et al. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). *Eur J Cancer 2017*;75:284-98.

93. Makama M, Drukker CA, Rutgers EJT, et al. An association study of established breast cancer reproductive and lifestyle risk factors with tumour subtype defined by the prognostic 70-gene expression signature (MammaPrint(R)). *Eur J Cancer* 2017;75:5-13.

94. Whitworth P, Beitsch P, Mislowsky A, et al. Chemosensitivity and Endocrine Sensitivity in Clinical Luminal Breast Cancer Patients in the Prospective Neoadjuvant Breast Registry Symphony Trial (NBRST) Predicted by Molecular Subtyping. Ann Surg Oncol 2017;24:669-75.
95. Xu W, Jia G, Cai N, et al. A 16 Yin Yang gene expression ratio signature for ER+/node- breast cancer. *Int J Cancer* 2017;140:1413-24.

96. MammaPrint Reduces Breast Cancer Overtreatment. *Cancer Discov* 2016;6:OF4.

97. Araki K, Ito Y. [A Review Multigene Assays for Clinical Utility in Breast Cancer]. *Gan To Kagaku Ryoho* 2016;43:1332-40.

98. Baron P, Beitsch P, Boselli D, et al. Impact of Tumor Size on Probability of Pathologic Complete Response After Neoadjuvant Chemotherapy. *Ann Surg Oncol* 2016;23:1522-9.

99. Bartlett JMS, Bayani J, Marshall A, et al. Comparing Breast Cancer Multiparameter Tests in the OPTIMA Prelim Trial: No Test Is More Equal Than the Others. *J Natl Cancer Inst* 2016;108:

100. Beitsch P, Whitworth P, Baron P, et al. Genomic Impact of Neoadjuvant Therapy on Breast Cancer: Incomplete Response is Associated with Altered Diagnostic Gene Signatures. *Ann Surg Oncol* 2016;23:3317-23.

101. Beumer I, Witteveen A, Delahaye L, et al. Equivalence of MammaPrint array types in clinical trials and diagnostics. *Breast Cancer Res Treat* 2016;156:279-87.

102. Beumer IJ, Persoon M, Witteveen A, et al. Prognostic Value of MammaPrint(R) in Invasive Lobular Breast Cancer. *Biomark Insights* 2016;11:139-46.

103. Blok EJ, van de Velde CJ, Smit VT. 70-Gene Signature in Early-Stage Breast Cancer. *N Engl J Med* 2016;375:2199.

104. Falato C, Tobin NP, Lorent J, et al. Intrinsic subtypes and genomic signatures of primary breast cancer and prognosis after systemic relapse. *Mol Oncol* 2016;10:517-25.

105. Gregoire JM, Fleury L, Salazar-Cardozo C, et al. Identification of epigenetic factors regulating the mesenchyme to epithelium transition by RNA interference screening in breast cancer cells. *BMC Cancer* 2016;16:700.

106. Gyanchandani R, Lin Y, Lin HM, et al. Intratumor Heterogeneity Affects Gene Expression Profile Test Prognostic Risk Stratification in Early Breast Cancer. *Clin Cancer Res* 2016;22:5362-9.

107. Kuijer A, Drukker CA, Elias SG, et al. Changes over time in the impact of gene-expression profiles on the administration of adjuvant chemotherapy in estrogen receptor positive early stage breast cancer patients: A nationwide study. *Int J Cancer* 2016;139:769-75.

108. Kuijer A, van Bommel ACM, Drukker CA, et al. Using a gene expression signature when controversy exists regarding the indication for adjuvant systemic treatment reduces the proportion of patients receiving adjuvant chemotherapy: a nationwide study. *Genet Med* 2016;18:720-6.

109. Li H, Zhu Y, Burnside ES, et al. MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. *Radiology* 2016;281:382-91.

110. Lopez Ruiz JA, Zabalza Estevez I, Mieza Arana JA. Breast cancer: determining the genetic profile from ultrasound-guided percutaneous biopsy specimens obtained during the diagnostic workups
Cancer de mama: determinacion del perfil genetico a partir de la biopsia percutanea ecoguiada diagnostica. *Radiologia* 2016;58:214-20.

111. Ma CX, Bose R, Ellis MJ. Prognostic and Predictive Biomarkers of Endocrine Responsiveness for Estrogen Receptor Positive Breast Cancer. *Adv Exp Med Biol* 2016;882:125-54.

112. Markopoulos C, van de Velde C, Zarca D, et al. Clinical evidence supporting genomic tests in early breast cancer: Do all genomic tests provide the same information? *Eur J Surg Oncol* 2017;43:909-20.

113. Myers MB. Targeted therapies with companion diagnostics in the management of breast cancer: current perspectives. *Pharmgenomics Pers Med* 2016;9:7-16.

114. Nagarajan R, Upreti M. An approach for deciphering patient-specific variations with application to breast cancer molecular expression profiles. *J Biomed Inform* 2016;63:120-30.

115. Nunes RA, Wray L, Mete M, et al. Genomic profiling of breast cancer in African-American women using MammaPrint. *Breast Cancer Res Treat* 2016;159:481-8.

116. Patil P, Colantuoni E, Leek JT eal. Genomic and clinical predictors for improving estimator precision in randomized trials of breast cancer treatments. *Contemp Clin Trials Comm* 2016;3:48-54.

117. Pohl H, Kotze MJ, Grant KA, et al. Impact of MammaPrint on Clinical Decision-Making in South African Patients with Early-Stage Breast Cancer. *Breast J* 2016;22:442-6.

118. Ribnikar D, Cardoso F. Tailoring Chemotherapy in Early-Stage Breast Cancer: Based on Tumor Biology or Tumor Burden? Am Soc Clin Oncol Educ Book 2016;35:e31-e38.

119. Schilsky R, Davies W. Highlights from the 2016 WIN Symposium, 27-29 June 2016, Paris: personalised therapy beyond next-generation sequencing. Ecancermedicalscience 2016;10:669.
120. Schmidt C. Mammaprint Reveals Who Can Skip Chemotherapy for Breast Cancer. *J Natl Cancer Inst* 2016;108:
121. Schmidt M, Thomssen C, Untch M. Intrinsic Subtypes of Primary Breast Cancer--Gene Expression Analysis. *Oncol Res Treat* 2016;39:102-10.

122. Stein RC, Dunn JA, Bartlett JMS, et al. OPTIMA prelim: a randomised feasibility study of personalised care in the treatment of women with early breast cancer. *Health Technol Assess* 2016;20:xxiii-201.

123. van't Veer LJ, Rutgers EJT, Piccart M. 70-Gene Signature in Early-Stage Breast Cancer. *N Engl J Med* 2016;375:2200-1.

124. Viale G, Slaets L, de Snoo FA, et al. Discordant assessment of tumor biomarkers by histopathological and molecular assays in the EORTC randomized controlled 10041/BIG 03-04 MINDACT trial breast cancer : Intratumoral heterogeneity and DCIS or normal tissue components are unlikely to be the cause of discordance. *Breast Cancer Res Treat* 2016;155:463-9.

125. Yerlikaya A, Erdogan E, Okur E, et al. A novel combination treatment for breast cancer cells involving BAPTA-AM and proteasome inhibitor bortezomib. *Oncol Lett* 2016;12:323-30.

126. Gyorffy B, Hatzis C, Sanft T, et al. Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res* 2015;17:11.

127. Hadi NI, Jamal Q. "OMIC" tumor markers for breast cancer: A review. *Pak J Med Sci* 2015;31:1256-62.

128. Rahilly-Tierney C, Walton SM. Cost-effectiveness of the 70-gene signature versus adjuvant! Online and systematic chemotherapy for risk stratification of patients with node-negative breast cancer: does accuracy matter? *J Clin Oncol* 2015;33:1628-9.

129. Segui MA, Crespo C, Cortes J, et al. In response: Genomic profile of breast cancer. *Expert Rev Pharmacoecon Outcomes Res* 2015;15:395-7.

130. Shimizu H, Horimoto Y, Arakawa A, et al. Application of a 70-Gene Expression Profile to Japanese Breast Cancer Patients. *Breast Care* (Basel) 2015;10:118-22.

131. Suo C, Hrydziuszko O, Lee D, et al. Integration of somatic mutation, expression and functional data reveals potential driver genes predictive of breast cancer survival. *Bioinformatics* 2015;31:2607-13.

132. Zemmour C, Bertucci F, Finetti P, et al. Prediction of early breast cancer metastasis from DNA microarray data using high-dimensional cox regression models. *Cancer Inform* 2015;14:129-38.

133. Bayraktar S, Royce M, Stork-Sloots L, et al. Molecular subtyping predicts pathologic tumor response in early-stage breast cancer treated with neoadjuvant docetaxel plus capecitabine with or without trastuzumab chemotherapy. *Med Oncol* 2014;31:163.

134. Bonastre J, Marguet S, Lueza B, et al. Cost effectiveness of molecular profiling for adjuvant decision making in patients with node-negative breast cancer. *J Clin Oncol* 2014;32:3513-9.

135. Cusumano PG, Generali D, Ciruelos E, et al. European inter-institutional impact study of MammaPrint. *Breast* 2014;23:423-8.

136. Drukker CA, Elias SG, Nijenhuis MV, et al. Gene expression profiling to predict the risk of locoregional recurrence in breast cancer: a pooled analysis. *Breast Cancer Res Treat* 2014;148:599-613.

137. Exner R, Bago-Horvath Z, Bartsch R, et al. The multigene signature MammaPrint impacts on multidisciplinary team decisions in ER+. *Br J Cancer* 2014;111:837-42.

138. Segui MA, Crespo C, Cortes J, et al. Genomic profile of breast cancer: cost-effectiveness analysis from the Spanish National Healthcare System perspective. *Expert Rev Pharmacoecon Outcomes Res* 2014;14:889-99.

139. Yin ZQ, Liu JJ, Xu YC, et al. A 41-gene signature derived from breast cancer stem cells as a predictor of survival. *J Exp Clin Cancer Res* 2014;33:49.

140. Zanotti L, Bottini A, Rossi C, et al. Diagnostic tests based on gene expression profile in breast cancer: from background to clinical use. *Tumour Biol* 2014;35:8461-70.

141. Delahaye L, Wehkamp D, Floore AN, et al. Performance characteristics of the MammaPrint® breast cancer diagnostic gene signature. *Personalized Medicine* 2013;10:

142. Kuijer A, Straver M, den Dekker B, et al. Impact of 70-Gene Signature Use on Adjuvant Chemotherapy Decisions in Patients With Estrogen Receptor-Positive Early Breast Cancer: Results of a Prospective Cohort Study. *J Clin Oncol* 2017;35:2814-9.

143. Esserman LJ, Yau C, Thompson CK, et al. Use of Molecular Tools to Identify Patients With Indolent Breast Cancers With Ultralow Risk Over 2 Decades. *JAMA Oncol* 2017;3:1503-10.

144. Straver ME, Glas AM, Hannemann J, et al. The 70-gene signature as a response predictor for neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res Treat* 2010;119:551-8.

145. Knauer M, Mook S, Rutgers EJT, et al. The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer. *Breast Cancer Res Treat* 2010;120:655-61.

146. Gluck S, de Snoo F, Peeters J, et al. Molecular subtyping of early-stage breast cancer identifies a group of patients who do not benefit from neoadjuvant chemotherapy. *Breast Cancer Res Treat* 2013;139:759-67.

147. Groenendijk FH, Zwart W, Floore A, et al. Estrogen receptor splice variants as a potential source of false-positive estrogen receptor status in breast cancer diagnostics. *Breast Cancer Res Treat* 2013;140:475-84.

148. Krijgsman O, Roepman P, Zwart W, et al. A diagnostic gene profile for molecular subtyping of breast cancer associated with treatment response. *Breast Cancer Res Treat* 2012;133:37-47.

149. Whitworth P, Stork-Sloots L, de Snoo FA, et al. Chemosensitivity predicted by BluePrint 80-gene functional subtype and MammaPrint in the Prospective Neoadjuvant Breast Registry Symphony Trial (NBRST). *Ann Surg Oncol* 2014;21:3261-7.

9 APPENDIX 1

9.1 History/background

Table A 1. Descriptive summary of the scoping search results[1]

Authors (year)	Sponsor/country	Search period and outcome measures	Evaluated tests	Conclusions
Marchionni (2008) [6]	AHRQ/USA	1990-2006 Clinical effectiveness: analytical and clinical validity	Oncotype DX®, MammaPrint®, and H/I.	The body of evidence shows that the tests offer clinically relevant, improved risk stratification over standard predictors. Oncotype DX® has the strongest evidence, closely followed by MammaPrint® and H/I (which is still maturing).
Smartt (2009) [7]	HSAC/NZ	2007-2009 Effectiveness	Oncotype DX®, MammaPrint® HOXB13:IL17BR	This update did not identify any studies providing direct, high-quality evidence that the investigated gene expression profiling tests lead to improvement in outcomes or are able to predict response to chemotherapy in any sub-set of women diagnosed with breast cancer.
OHTA (2010) [81]	OHTA/CA	2006-2010 Laboratory performance, prognostic and predictive value and cost-effectiveness	Oncotype DX®	There is currently insufficient evidence investigating how Oncotype DX® would impact clinician/patient decision-making in a setting generalizable to Ontario.
Ward (2011) [8]	NICE/UK	2009-2011 (for Oncotype DX®) and 2002-2011 (for MammaPrint®): Clinical effectiveness (analytic and clinical validity, clinical utility) and cost-effectiveness for nine tests	BluePrint®, Breast Cancer Index (BCI), IHC4, MammaPrint®, Mammostrat, NPI plus (NPI+), Oncotype DX®, PAM50 and Randox Breast Cancer Array.	The clinical evidence base for Oncotype DX® is considered to be the most robust. For MammaPrint® and Mammostrat there were significant gaps in the available evidence. Evidence for the remaining five tests (PAM50, NPI+, BCI, BluePrint® and Randox®) was limited.
San Miguel (2015) [5]	KCE/Belgium	2011-2014 Clinical validity and cost-	Oncotype DX®, PAM50, MapQuant	The evidence for Oncotype DX® is more robust than the evidence for other tests. Important evidence gaps are still present. It is not yet

		effectiveness	DX®, H/I (replaced by BCI), EndoPredict®, Mammaprint®, BluePrint®, Randox Breast Cancer Array, Mammostrat®, NPI+, IHC4, uPA/PAI-1	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®.
Marrone (2015) [10]	Academic/USA	2007-2013 Clinical utility	Oncotype DX® and MammaPrint®,	Indirect evidence showed Oncotype DX® was able to predict treatment effects of adjuvant chemotherapy, whereas no evidence of predictive value was found for MammaPrint®. Both tests influenced a change in treatment recommendations in 21 to 74% of participants.
IQWiG (2016) [11]	IQWiG/Germany	?-2016 Clinical validity and clinical utility	Oncotype DX®, PAM50 risk-of-recurrence score/Prosigna® Breast Cancer Prognostic Gene Signature Assay, Breast Cancer Index, IHC4	There is currently no hint of a benefit or harm of a biomarker-based strategy to support the decision for or against adjuvant chemotherapy.

[1] Clinical trial databases were searched to identify studies evaluating the clinical effectiveness of GES tests: MammaPrint®, Oncotype DX®, PAM50, MapQuant DX, H/I*, EndoPredict®, BluePrint®, Randox Breast Cancer Array, Prosigna®, and the Breast Cancer Index.

9.2 Documentation of the search strategies

Search terms were:
Medline (PubMed): "Gene Expression Profiling"[Mesh] AND (mammaprint[tiab] OR 70 gene[tiab] OR MINDACT[tiab]) AND "Breast Neoplasms"[Mesh]
Limit: from June 2014, English, Dutch
22 references

Not yet indexed publications
((mammaprint[tiab] OR 70 gene[tiab] OR MINDACT[tiab]) AND breast*[tiab]) NOT Medline[sb]
Limit: from June 2014, English, Dutch
25 references

Embase
mammaprint:ab,ti OR '70 gene':ab,ti AND 'breast tumor'/exp AND [embase]/lim NOT [medline]/lim
Limit: from 2015, English, Dutch
6 references (after removing duplicates)

9.3 Description of the evidence used

Table A 2. Baseline Characteristics, According to Risk Group.[a]

Characteristics		**Low clinical risk**		**High clinical risk**		**All Patients (n=6693)**
		Low genomic risk (n=2745)	**High genomic risk (n=592)**	**Low genomic risk (n=1550)**	**High genomic risk (n=1806)**	
		Number (%)				
Age, yr						
	<35	24 (0.9)	13 (2.2)	20 (1.3)	65 (3.6)	122 (1.8)
	35 to <50	774 (28.2)	165 (27.9)	514 (33.2)	651 (36.0)	2104 (31.4)
	50 to 70	1928 (70.2)	403 (68.1)	1000 (64.5)	1080 (59.8)	4411 (65.9)
	>70	19 (0.7)	11 (1.9)	16 (1.0)	10 (0.6)	56 (0.8)
Tumour size, cm[b]						
	<1	655 (23.9)	198 (33.4)	38 (2.5)	29 (1.6)	920 (13.7)
	1 to 2	1968 (71.7)	383 (64.7)	610 (39.4)	914 (50.6)	3875 (57.9)
	>2 to 5	122 (4.4)	11 (1.9)	843 (54.4)	843 (46.7)	1819 (27.2)
	>5	0	0	58 (3.7)	20 (1.1)	78 (1.2)
Tumour grade[c]						
	1	1242 (45.2)	92 (15.5)	98 (6.3)	15 (0.8)	1447 (21.6)
	2	1457 (53.1)	414 (69.9)	995 (64.2)	421 (23.3)	3287 (49.1)
	3	36 (1.3)	83 (14.0)	443 (28.6)	1365 (75.6)	1927 (28.8)
	Missing data	10 (0.4)	3 (0.5)	14 (0.9)	5 (0.3)	32 (0.5)
Lymph node status[d]						

Characteristics		Low clinical risk		High clinical risk		All Patients (n=6693)
		Low genomic risk (n=2745)	High genomic risk (n=592)	Low genomic risk (n=1550)	High genomic risk (n=1806)	
	Negative	2570 (93.6)	577 (97.5)	812 (52.4)	1329 (73.6)	5288 (79.0)
	Positive					
	1 node	131 (4.8)	10 (1.7)	505 (32.6)	296 (16.4)	942 (14.1)
	2 nodes	26 (0.9)	3 (0.5)	157 (10.1)	114 (6.3)	300 (4.5)
	3 nodes	18 (0.7)	2 (0.3)	69 (4.5)	65 (3.6)	154 (2.3)
	≥4 nodes	0	0	6 (0.4)	2 (0.1)	8 (0.1)
Hormone receptor status[e]						
	ER-positive, PR positive, or both	2741 (99.9)	535 (90.4)	1520 (98.1)	1118 (61.9)	5914 (88.4)
	ER-negative and PR-negative	4 (0.1)	57 (9.6)	29 (1.9)	688 (38.1)	778 (11.6)
HER2 status[f]						
	Negative	2641 (96.2)	518 (87.5)	1423 (91.8)	1461 (80.9)	6043 (90.3)
	Positive	97 (3.5)	73 (12.3)	124 (8.0)	344 (19.0)	638 (9.5)
	Missing data	7 (0.3)	1 (0.2)	3 (0.2)	1 (0.1)	12 (0.2)
Clinicopathological subtype[g]						
	Luminal HER2-negative: ER-positive, PR-positive, or both	2638 (96.1)	467 (78.9)	1402 (90.5)	895 (49.6)	5402 (80.7)
	Luminal HER2-positive: ER-positive, PR-positive, or both	96 (3.5)	68 (11.5)	115 (7.4)	222 (12.3)	501 (7.5)
	Non-luminal HER2-positive: ER-negative, PR negative	1 (<0.1)	5 (0.8)	9 (0.6)	122 (6.8)	137 (2.0)
	Triple negative: ER-negative, PR-negative, HER2-negative	3 (0.1)	51 (8.6)	20 (1.3)	566 (31.3)	640 (9.6)
	Missing data	7 (0.3)	1 (0.2)	4 (0.3)	1 (0.1)	13 (0.2)
WHO performance status[h]						
	0	2644 (96.3)	565 (95.4)	1491 (96.2)	1734 (96.0)	6434 (96.1)
	1	101 (3.7)	27 (4.6)	58 (3.7)	71 (3.9)	257 (3.8)
	2	0	0	1 (0.1)	1 (0.1)	2 (<0.1)

Abbreviations: ER: oestrogen receptor; ; PR: progesteron receptor HER2: human epidermal growth factor 2 receptor;

[a] Data was missing for one patient at high clinical and low genomic risk with respect to tumour size, lymph node status, and hormone receptor status. Percentages may not total 100% because of rounding. ER denotes oestrogen receptor, HER2 human epidermal growth factor receptor, and PR progesterone receptor.

[b] A majority of patients at high clinical and low genomic risk (54%) had tumours measuring 2 to 5 cm in diameter. Most of the patients at low clinical and genomic risk (96%) and at low clinical and high genomic risk (98%) had tumours measuring 2 cm or less, as did 52% of the patients at high clinical and genomic risk.

[c] More than three quarters (76%) of patients at high clinical and genomic risk had grade 3 tumours. Most patients at low clinical and genomic risk, low clinical and high genomic risk, and high clinical and low

genomic risk had grade 1 or 2 tumours (98%, 85%, and 71%, respectively).
[d] The presence of negative lymph nodes was substantially more frequent among patients at low clinical and genomic risk (94%) and low clinical and high genomic risk (97%) than among patients at high clinical and low genomic risk (52%) and high clinical and genomic risk (74%).
[e] Almost all tumours were positive for hormone receptors except among patients at high clinical and genomic risk, in whom 38% of tumours were hormone receptor-negative. Hormone receptor positivity was defined as the presence of at least 1% immunoreactive cells, an Allred score of greater than 2 (on a scale from 0 to 8, with higher scores indicating a greater number of receptors), or a level of cytosolic protein of at least 10 fmol per milligram.
[f] HER2 positivity was reported in 4% of patients at low clinical and genomic risk, 12% of those at low clinical and high genomic risk, 8% of those at high clinical and low genomic risk, and 19% of those at high clinical and genomic risk.
[g] Specifically, among patients at high clinical and low genomic risk, 48% had node-positive disease, 58% of tumours measured 2 cm or more, and 90% had the luminal HER2-negative subtype.
[h] The World Health Organization performance scores range from 0 to 5, with 0 denoting perfect health and 5 death.

9.3.1 Guidelines for diagnosis and management

Table A 3. Overview of 'surrogate' intrinsic breast cancer subtypes [45]

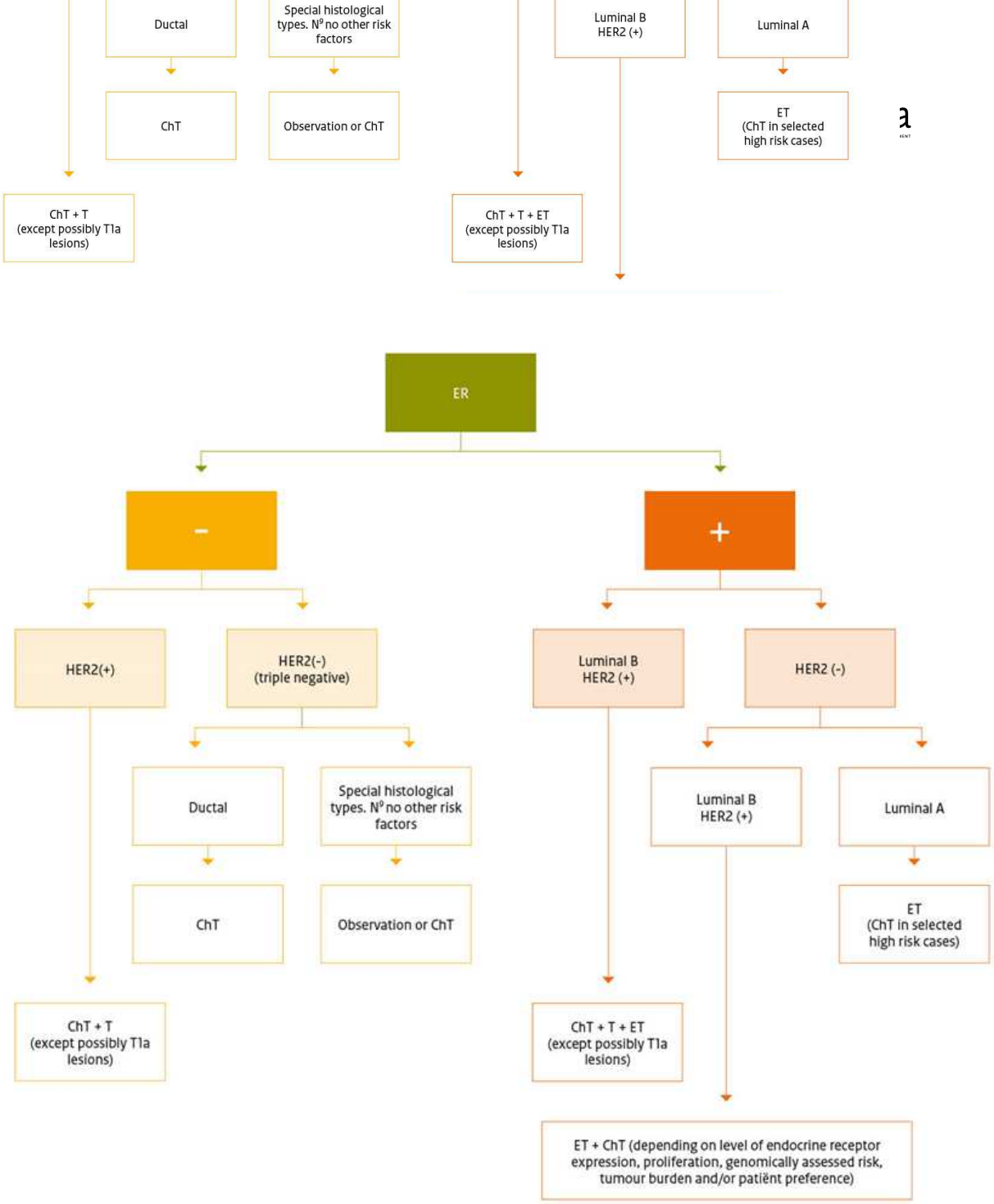
Intrinsic subtype	Prognosis
Luminal A (50-60%) 'Luminal A-like' ER-positive, HER2-negative, Ki67 low*, PR high** Low-risk molecular signature (if available)	Good Recurrence common in bone
Luminal B (15-20%) 'Luminal B-like (HER2-negative)' ER-positive, HER2-negative, and either Ki67 high or PR low High-risk molecular signature (if available)	Poor Increased relapse rate in the first five years after diagnosis [82]
'Luminal B-like (HER2-positive)' ER-positive, HER2-positive, any Ki67, any PR	
HER2 overexpression (15-20%) 'HER2-positive (non-luminal)' HER2-positive, ER and PR absent	Poor
'Basal-like' (8-37%)*** 'Triple-negative (ductal)' ER and PR absent HER2-negative	Poor Metastasis to brain and lung

Abbreviations: ER: oestrogen receptor; PR: progesteron receptor HER2: human epidermal growth factor 2 receptor;

* Ki67 scores should be interpreted in the light of local laboratory values: as an example, if a laboratory has a median Ki67 score in receptor-positive disease of 20%, values of 30% or above could be considered clearly high; those of 10% or less clearly low.

** Suggested cut-off value is 20%; quality assurance programmes are essential for laboratories reporting these results.

*** There is ~80% overlap between 'triple-negative' and intrinsic 'basal-like' subtype, but 'triple-negative' also includes some special histological types such as (typical) medullary and adenoid cystic carcinoma with low risks of distant recurrence [45].



Abbreviations: ER: oestrogen receptor; HER2: human epidermal growth factor 2 receptor; ChT: chemotherapy; ET: endocrine therapy; T: trastuzumab.

Table A 4. Overview of guidelines

Name of society/organisation issuing guidance	Date of issue	Country	Summary of recommendation
Clinical practice guidelines published after the MINDACT publication			
St Gallen[45]	2017	International	In ER-positive, HER2-negative, lymph node-negative breast cancer, MammaPrint® is endorsed as a prognostic marker for adjuvant endocrine therapy in node-negative breast cancers. It is also recommended for guiding the decision on adjuvant chemotherapy in node-negative tumours, identifying cases at low risk, with an excellent prognosis that would not warrant chemotherapy. In ER-positive, HER2-negative, lymph node-positive breast cancer, gene expression signatures were not uniformly endorsed for making treatment decisions regarding adjuvant chemotherapy in node-positive cases. Only MammaPrint® and Oncotype DX® were recommended. Patients with low-risk tumour scores and a limited degree of nodal involvement appear to have a good prognosis with or without chemotherapy.
American Joint Committee on Cancer (AJCC)[40]	2017	USA/ International	the AJCC Task Panel recognizes that MammaPrint® has Level 1 evidence, based on MINDACT, for determining clinical prognosis. MammaPrint® is currently the only breast cancer risk of recurrence test that is recognized by ASCO and AJCC for use in clinically high risk patients. Additionally, they do not endorse the use of any specific multigene genomic panel. These statements clarify that the use of MammaPrint® is compliant with Commission on Cancer – National Accreditation Program for Breast Centers (CoC-NAPBC) accreditation standards.
American Society of Clinical Oncology (ASCO) [41]	2017	USA	If a patient has ER/PR–positive, HER2-negative, node-negative breast cancer, the MammaPrint® (Agendia, Irvine, CA) assay may be used in those with high clinical risk per MINDACT categorisation to inform decisions on withholding adjuvant systemic chemotherapy due to its ability to identify a good-prognosis population with potentially limited chemotherapy benefit. If a patient has ER/PR–positive, HER2-negative, node-negative breast cancer, the MammaPrint® assay **should not** be used in those with low clinical risk per MINDACT categorisation to inform decisions on withholding adjuvant systemic chemotherapy, because women in the low clinical risk

			category had excellent outcomes and did not appear to benefit from chemotherapy even with a genomic high-risk cancer. If a patient has ER/PR–positive, HER2-negative, node positive breast cancer, the MammaPrint® assay **may** be used in patients with one to three positive nodes and at high clinical risk per MINDACT categorisation to inform decisions on withholding adjuvant systemic chemotherapy due to its ability to identify a good prognosis population with potentially limited chemotherapy benefit. However, such patients should be informed that a benefit of chemotherapy cannot be excluded, particularly in patients with greater than one involved lymph node. If a patient has ER/PR–positive, HER2-negative, node-positive breast cancer, the MammaPrint® assay should not be used in patients with one to three positive nodes and at low clinical risk per MINDACT categorisation to inform decisions on withholding adjuvant systemic chemotherapy. There are insufficient data on the clinical utility of MammaPrint® in this specific patient population. If a patient has HER2-positive breast cancer, the clinician should not use the MammaPrint® assay to guide decisions on adjuvant systemic therapy. Additional studies are required to address the role of MammaPrint® in patients with this tumour subtype who are also receiving HER2-targeted therapy. If a patient has ER/PR negative and HER2-negative (triple negative) breast cancer, the clinician should not use the MammaPrint® assay to guide decisions on adjuvant systemic chemotherapy.
National Comprehensive Cancer Network[83]	2017	USA	The 70-gene signature assay is approved by the FDA to assist in assignment of women with ER-positive or ER-negative breast cancer into a high versus a low risk of recurrence, but not for predicting benefit from adjuvant systemic therapy. The prospective RASTER study reported that breast cancer patients classified by the 70-gene signature as low risk (of whom 85% did not receive adjuvant chemotherapy) had an overall 97% distant recurrence free interval at five years. The NCCN panel members acknowledge that many assays have been clinically validated for prediction of prognosis. The MINDACT trial is a phase III trial comparing the 70-gene signature with the commonly used clinicopathological criteria in selecting patients for adjuvant chemotherapy in breast cancer with 0 to 3 nodes. The early results from the MINDACT trial suggest that the 70-gene signature can help avoid

			chemotherapy in certain patients regardless of larger tumour size and nodal status, without compromising the outcome.
AETSA[84]	2017	Spain	Mammaprint® to assess necessity of adjuvant chemotherapy in females or males with recently diagnosed breast tumors, where all of the following criteria are met: A. Breast cancer is nonmetastatic (node negative[1]) or with 1-3 involved ipsilateral axillary lymph nodes; *and* B. Breast tumor is estrogen receptor positive or progesterone receptor positive; *and* C. Breast tumor is HER2 receptor negative (Rationale: adjuvant chemotherapy with trastuzumab (Herceptin) is considered to be medically necessary regardless of Mammaprint® score for HER2 receptor positive lesions); *and* D. Member is determined to be at "high clinical risk" of recurrence using Adjuvant! Online ((see page 20 of MINDACT study supplement for definitions of high clinical risk; *and* E. Adjuvant chemotherapy is not precluded due to any other factor (e.g., advanced age and/or significant co-morbidities); *and* F. Member and physician (prior to testing) have discussed the potential results of the test and agree to use the results to guide therapy.
AGO (German Gynecological Oncology Group)	2017	Germany	2017 Guidelines of the AGO Breast Committee acknowledges Agendia's MammaPrint® test with highest medical evidence level 1A for identification of patient subgroups who can potentially forgo chemotherapy for breast cancer
Guidelines before the MINDACT publication.			
Cancer Care Ontario[85]	2016	Canada	Recommendations •Clinicians may offer multigene profile assay testing to potential chemotherapy candidates with invasive breast carcinoma that is ER-positive/HER2-negative. •In patients with node-negative ER-positive/HER2-negative disease, clinicians may use a low-risk result from Oncotype DX®, Prosigna, or EndoPredict/EPclin assays to support a decision to withhold chemotherapy. •In patients with node-negative ER-positive/HER2-negative disease, clinicians may use a high-risk result from Oncotype DX® to support a decision to offer chemotherapy. A high-risk Oncotype DX® result in this subpopulation has been associated with both poor prognosis

			without chemotherapy, and a prediction of benefit from giving chemotherapy. •In some patients with ER-positive/HER2-negative tumours and 1-3 nodes involved (N1a disease), clinicians may withhold chemotherapy based on a low-risk Oncotype DX® or Prosigna score if the decision is supported by other clinical, pathological, or patient-related factors. •In patients with ER-positive disease, there is insufficient evidence to recommend the use of multigene profiling assays to inform clinical decision making for late risk of recurrence. A high-risk score using Prosigna or EndoPredict prognosticates for late recurrence; however, evidence is lacking that these tests predict for benefit of extended adjuvant endocrine treatment beyond five years. Mammaprint® Three identified studies assessed the prognostic ability of MammaPrint®. Based on the tumour marker utility grading system, all three studies were assessed as category C studies. The three studies reported consistent findings; however, since the studies were assessed as category C studies, the overall evidence supporting the prognostic ability of MammaPrint® is considered to be level II. The original prospective observational study by van de Vijver et al. classified patients into poor or good prognosis based on MammaPrint® findings and found that scores were prognostic of early distant recurrence. The other two prospective observational studies analysed RNA extracted from the tumour samples of patients enrolled in the microarRAy prognoSTics in breast cancer (RASTER) study. The first publication compared the risk of five-year distant-recurrence free interval predicted by MammaPrint® with the predicted risk determined by Adjuvant! Online and found that MammaPrint® added prognostic value above the clinicopathological risk estimate of Adjuvant! Online. The second publication found that MammaPrint® score was able to prognosticate early distant recurrence independent of a clinical risk estimate. MammaPrint® has Level II evidence supporting its use as a prognostic tool for distant recurrence due to the category C studies assessing this clinical utility.
European Society for Medical Oncology ESMO[42]	2015	Europe	Gene expression profiles, such as MammaPrint® (Agendia, Amsterdam, the Netherlands), Oncotype DX® Recurrence Score (Genomic Health, Redwood City, CA), Prosigna (Nanostring technologies, Seattle, WA) and Endopredict (Myriad Genetics), may be used to gain additional prognostic and/or predictive information to complement pathology assessment and to predict the benefit of adjuvant chemotherapy. The three latter tests are designed for patients with ER-positive early breast cancer only.

			The clinical utility of MammaPrint® and Oncotype DX® is still being prospectively evaluated in large randomised clinical trials such as MINDACT for MammaPrint®, WSG PLAN B trial, TAILORx and RxPONDER for Oncotype DX®. A IB evidence level has been achieved from retrospective analyses of data from prospective trials regarding the prognostic value of MammaPrint®, Oncotype DX®, Prosigna, and Endopredict in ER-positive breast cancers. In addition, the prognostic value of MammaPrint® has been validated in the Raster trial, a prospective but nonrandomised, clinical trial.
KCE[5]	2015	Belgium	Revision of this guideline is currently in progress. The current guideline recommends (adapted from Ward et al. [9]) 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® (microarray 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, 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.
AHRQ[86]	2014	Maryland, USA	**Conclusions** We found no evidence to determine whether using the tests to estimate prognosis leads to improved outcomes for patients. **MammaPrint®** Our meta-analysis suggests that patients classified as having a poor prognosis by MammaPrint® consistently do worse than those with a good prognosis signature with respect to distant metastatic-free survival and CSS. Studies included in the search varied somewhat in the subpopulations they tested; the fact that the signature added prognostic value across these various subpopulations suggests that it has broad applicability. Because no studies looked at the value of MammaPrint® with respect to loco-regional recurrence (LRR) and only one considered overall survival, there is scant or

			no evidence regarding the clinical validity of the MammaPrint® signature in terms of these outcomes.
Institut National du Cancer (INCa)[87]	2013	France	In 2013, The French National institute of Cancer (InCa) concluded there was not enough evidence of added value with MP compared to usual clinico-pathological prognostic factors. Without this prove of added value, clinical utility of a novel marker as MP couldn't be established at time of the French report. The National institute of Cancer concludes there is not enough evidence for recommendations.
NICE[88]	2013	UK	Revision of this guideline is currently in progress. The current guideline recommends: MammaPrint®, IHC4 and Mammostrat are only recommended for use in research in people with ER+, LN- and HER2- early breast cancer to collect evidence about potentially important clinical outcomes and to determine the ability of the tests to predict the benefit of chemotherapy. The tests are not recommended for general use in these people because of uncertainty about their overall clinical benefit.
DKG[89]	2012	Germany	Revision of this guideline is currently in progress Recommendations according to the guideline from 2012: The EGAPP (Evaluation of Genomic Applications in Practice and Prevention) Working Group comes to the conclusion that because of inadequate evidence, no recommendation regarding the use of gene expression profiles can be given (see http://www.cdc.gov/genomics/gtesting/EGAPP/recommend/GEP_provider.htm). Since the studies published in the meantime do not abolish these general deficits, it is currently not possible to make a recommendation for everyday practice therapy decisions regarding gene expression-analysis.
Nationale BorstKanker Overleg Nederland (NABON)[90]	2012	Netherlands	Revision of this guideline is currently in progress The current NABON guidelines recommend that MammaPrint® may be used in individual cases with a hormone sensitive invasive ductal carcinoma if there is doubt about the indication for adjuvant chemotherapy on the basis of traditional prognostic factors •In 2016, NABON issued an update letter in response to the impact of the publication of the MINDACT trial. It provisionally concluded that MammaPrint® may have added value in patients with pT1-2N0 and pT1-N1, ER-positive and HER2-negative invasive ductal breast cancer who are considered for chemotherapy

Unidad de Evaluacion de Tecnologias Sanitarias (UETS)[91]	2012	Spain	The document contains an economic evaluation of Oncotype DX®. This model does not provide information on the comparison with other tests such as MammaPrint®. There are more studies needed in this context between different commercialised tests, with more patients included, since the obtained results are not conclusive.

Abbreviations: ER: oestrogen receptor; PR: progesteron receptor; HER2: human epidermal growth factor 2 receptor; GRADE: Grading of Recommendations, Assessment, Development and Evaluation; RCT: Randomised Controlled Trial.

9.3.1.1 Evidence tables of individual studies included for clinical effectiveness and safety

Table A 5. Characteristics of randomised controlled study

Author and year or study name	Study type	Number of patients	Intervention(s)	Main endpoints	Conclusion Authors	Included in clinical effectiveness and/or safety domain
Cardoso et al. 2016 MINDACT EORTC[4]	RCT, open-label, randomisations centrally, PP and ITT analysis reported, multicentre (112) in 9 European countries. Follow-up: primary endpoint five-year DMFS. In addition ten-years data will be collected. 10 years and in case of endocrine therapy 15 years. Sponsoring see Table 3.1	6693 19-70 years Female Histologically confirmed primary breast cancer (stage T1, T2, or operable T3) lymph node-negative or up to three positive axillary nodes. In the high clinical risk group and the low genomic risk group (PPS analysis), 503 patients received chemotherapy based on clinical risk and 542 patients did not receive chemotherapy based on genomic risk.	MammaPrint® added to Adjuvant! Online for making decision chemotherapy yes or no compared with decision based only on Adjuvant! Online. Chemotherapy randomised: anthracycline or docetaxel-plus-capecitabine regimen. Hormone+ randomized tamoxifen-letrozole or letrozole-only regimen.	Primary: five-year distant metastasis-free survival[$] Secondary: Proportion of patients that received chemotherapy according to the clinical risk as compared with the genomic risk Overall survival* Disease-free survival[#]	Among women with early-stage breast cancer who were at high clinical risk and low genomic risk for recurrence, the receipt of no chemotherapy on the basis of the 70-gene signature led to a five-year rate of survival without distant metastasis that was 1.5 percentage points lower than the rate with chemotherapy. Given these findings, approximately 46% of women with breast cancer who are at high clinical risk might not require chemotherapy.	Eff /Saf

Abbreviations:

[$]: Survival without distant metastasis was defined as the time until the first distant metastatic recurrence or death from any cause.

[#]: Disease-free survival was defined as the time until first disease progression (locoregional, distant relapse, ipsilateral or contralateral invasive breast cancer, ductal carcinoma *in situ* or an invasive second primary cancer or death from any cause.

* Overall survival was defined as the time until death from any cause.

Table A 6: Reason for exclusion studies

	Citation	Reason for exclusion
1	Duffy et al, 2017 [92]	Guideline- publication/consensus meeting
2	Makama et al, 2017[93]	Association study
3	Whitworth et al, 2017[94]	neo-adjuvant setting
4	Xu et al. 2017[95]	Concordance study with other prognostic tests (Miscellaneous)
5	Poh, 2016[96]	Comment/letter
6	Araki et al. 2016[97]	Other language
7	Baron et al. 2016[98]	neo-adjuvant setting
8	Bartlett et al. 2016[99]	Concordance study with other prognostic tests
9	Beitsch et al. 2016[100]	neo-adjuvant setting
10	Beumer et al. 2016[101]	Technical and/or clinical validation
11	Beumer et al. 2016[102]	Technical and/or clinical validation in invasive lobular carcinoma (duplicate with 55)
12	Blok et al. 2016[103]	Comment/letter
13	Falato et al. 2016[104]	Clinical validation
14	Gregoire et al. 2016[105]	Pre-clinical study
15	Gyanchandani et al. 2016[106]	Pre-clinical study
16	Kuijer et al. 2016 [107]	Decision impact study no outcomes on clinical utility
17	Kuijer et al. 2016 [108]	Decision impact study no outcomes on clinical utility
18	Li et al. 2016[109]	Miscellaneous
19	Lopez et al. 2016[110]	Miscellaneous
20	Ma et al. 2016[111]	Review
21	Markopoulos et al. 2016[112]	Review
22	Myers et al. 2016[113]	Review
23	Nagarajan et al. 2016[114]	Miscellaneous
24	Nunes et al. 2016[115]	Technical and/or clinical validation
25	Patil et al. 2016[116]	Miscellaneous
26	Pohl et al. 2016 [117]	Decision impact study no outcomes on clinical utility
27	Ribnikar et al. 2016[118]	Review
28	Schilsky et al. 2016[119]	Miscellaneous
29	Schmidt C et al. 2016[120]	Comment/letter
30	Schmidt M et al. 2016[121]	Review
31	Stein et al. 2016[122]	Technical and/or clinical validation
32	Thewes et al. 2016[71]	Comment/letter
33	Van 't Veer et al. 2016[123]	Comment/letter
34	Viale et al. 2016 [124]	Technical and/or clinical validation
35	Yerlikaya et al. 2016[125]	Concordance study with other prognostic tests
36	Gyorffy et al. 2015[126]	Review
37	Hadi et al. 2015[127]	Review
38	Marrone et al. 2015[10]	Review
39	Rahilly-Tierney et al. 2015[128]	Related to costeffectiveness
40	Segui et al. 2015[129]	Comment/letter/review
41	Shimizu et al. 2015[130]	Technical and/or clinical validation
42	Suo et al. 2015[131]	concordance study
43	Zemmour et al. 2015[132]	Technical and/or clinical validation
44	Bayraktar et al. 2014[133]	neo-adjuvant setting

45	Bonastre et al. 2014[134]	Related to cost-effectiveness
46	Cusumano et al. 2014[135]	Decision impact study no outcomes on clinical utility
47	Drukker et al. 2014 [136]	Technical and/or clinical validation
48	Drukker et al. 2013 [12]	No supportive evidence on clinical utility
49	Exner et al. 2014[137]	Decision impact study no outcomes on clinical utility
50	Segui et al. 2014[138]	Related to cost effectiveness
51	Yin et al. 2014[139]	concordance study
52	Zanotti et al. 2014[140]	Review
53	Krop et al. 2017[41]	Guideline publication/consensus meeting
54	Kuijer et al. 2016[107]	Decision impact study no outcomes on clinical utility
55	Beumer et al. 2016[102]	Technical and/or clinical validation (duplicate with 11)
56	Delahaye et al. 2017[141]	Technical and/or clinical validation
57	Kuijer et al. 2017[142]	Decision impact study no outcomes on clinical utility
58	Esserman et al. 2017[143]	Technical and/or clinical validation
59	Curigliano et al. 2017[39]	Guideline- publication/consensus meeting
60	Straver et al. 2010[144]	neo-adjuvant setting
61	Knauer et al. 2010[145]	Decision impact study direct evidence on clinical utility
62	Glück et al. 2013[146]	neo-adjuvant setting
63	Groenendijk et al. 2013[147]	concordance study
64	Krijgsman et al. 2012[148]	concordance study in neo-adjuvant setting
65	Whitworth et al. 2014[149]	neo-adjuvant setting
66	Retel et al. 2013[13]	No supportive evidence on clinical utility

Table A 7: Reason enrolment was not successful (screening failure)

Reason for screening failure	**All screening failures (n=4595) n(%)**
MammaPrint® not feasible (mostly <50%/<30% tumour cells)	1182 (26%)
Patient/investigator decision	899 (20%)
Ineligible: LN status	772 (17%)
Inadequate/absent sample	768 (17%)
Ineligible: other	447 (10%)
Unknown or other	527 (11%)

9.3.2 List of relevant on-going and planned studies

Table A 8. List of relevant on-going studies with gene-expression signature

Study identifier	Estimated completion date	Study type	Number of patients	Intervention	Comparator	Patient population	Endpoints
NCT00310180 (TAILORx)	31 December 2017	RCT	11248	No chemotherapy based on Oncotype DX®	Chemotherapy	Breast cancer patients with node-negative, estrogen-receptor positive breast cancer by using a special test (Oncotype DX®), and whether hormone therapy alone or hormone therapy together with combination chemotherapy is better for women who have an Oncotype DX® recurrence score of 11-25.	To determine whether adjuvant hormonal therapy is not inferior to adjuvant chemohormonal in women whose tumours meet established clinical guidelines for adjuvant chemotherapy and fall in the "primary study group" category (Oncotype DX® Recurrence Score 11-25). •Distant recurrence-free interval •Overall survival •Recurrence-free interval •Perceived cognitive function evaluated using the Functional Assessment of Cancer Therapy - Cognitive Function •Quality of life measured using FACT-COG, fatigue (FACT-Fatigue and Patient Reported Outcomes Measurement Information System Fatigue SF), fear of recurrence (Assessment of Survivor Concerns), endocrine symptoms (FACT-ES) and HRQL (FACT-General) [time frame: up to 36 months]
NCT01272037 RxPONDER	1 February 2022	RCT	10000	No chemotherapy based on Oncotype DX®.	Chemotherapy	A Phase III, randomized clinical trial of standard adjuvant endocrine therapy +/- chemotherapy in patients With 1-3 positive	To determine the effect of chemotherapy in patients with node-positive breast cancer who do not have high recurrence scores (RS) by Oncotype DX®. •Overall survival

						nodes, hormone receptor-positive and HER2-negative breast cancer with Recurrence Score (RS) of 25 or Less. RxPONDER: A Clinical Trial Rx for Positive Node, Endocrine Responsive Breast Cancer	•Distant disease-frees survival (DDFS) •Local disease-free interval (LDFI) •Toxicity •Patient-reported anxiety •Initial management cost of node-positive, hormone receptor-positive, HER2-negative breast cancer. •Patient-reported utilities (e.g., quality of life) for those randomised to chemotherapy versus no chemotherapy.

Source: clinicaltrials.gov

9.3.3 Risk of bias tables

Table A 9. Risk of bias – study level (RCTs)

Trial	Adequate generation of randomisation sequence	Adequate allocation concealment	Blinding		Selective outcome reporting unlikely	No other aspects which increase the risk of bias	Risk of bias – study level
			Patient	Medicinal personnel and other staff			
MINDACT CH/GL	Yes	Yes	No[1]	No[1]	Yes	No[2]	High[3]
MINDACT CL/GH	Yes	Yes	No[1]	No[1]	Yes	Yes	Low[4]

Abbreviations: CH/GL: clinical high and genomic low, CL/GH = clinical low and genomic high; PP: per protocol; OS overall survival; DMFS: distant metastasis-free survival; DFS: disease-free survival.
Source: MINDACT

[1] Open-label study

[2] Of the patients randomised to chemotherapy (n=749) 174 (23%) patients were not included in the PP analysis (of whom 128 did not receive chemotherapy and 26 had a change of risk). Of those assigned to no chemotherapy (n=748), 119 (16%) were not included in the PP analysis (of whom 85 received chemotherapy and 21 had a change of risk). No lost to follow up data is mentioned. Because a considerable number of patients could not be analysed in the PP analysis, it is not clear if the baseline characteristics were still comparable or there was selective dropout. The results do not show if both groups are comparable.

[3] Because of both 1 and 2 it is concluded that risk of bias is high.

[4] We have rated the risk of bias as low because in the case of OS, DFS, and DMFS, we find downgrading only for the fact that study was open-label too strict.

Table A 10. Risk of bias – outcome level (RCTs)

Outcome Trial	Blinding – outcome assessors	ITT principle adequately realized	Selective outcome reporting unlikely	No other aspects according to risk of bias	Risk of bias – outcome level
Overall survival 5y					
MINDACT (CH/GL)	Low	High (PP)[#]	Low	Low	High
MINDACT (CL/GH)	Low	Low	Low	Low	Low
DMFS 5y					

Outcome Trial	**Blinding – outcome assessors**	**ITT principle adequately realized**	**Selective outcome reporting unlikely**	**No other aspects according to risk of bias**	**Risk of bias – outcome level**
MINDACT (CH/GL)	High*	High (PP)[#]	Low	Low	High
MINDACT (CL/GH)	High*	Low	Low	Low	Low
DFS 5y					
MINDACT (CH/GL)	High*	High (PP)#	Low	Low	High
MINDACT (CL/GH)	High*	Low	Low	Low	Low
QoL					
MINDACT (CH/GL)	Not performed[¥]				
MINDACT (CL/GH)	Not performed[¥]				
Side effects of chemotherapy					
MINDACT (CH/GL)	Not reported				
MINDACT (CL/GH)	Not reported				

Abbreviations: CH/GL: clinical high and genomic low, CL/GH = clinical low and genomic high; ITT: intention to treat; OS overall survival; DMFS: distant metastasis-free survival; DFS: disease-free survival; QoL: quality of life.
Source: MINDACT

* Because outcome assessor was not blinded there is a risk of bias because there is some degree of subjectivity in assessing DMFS and DFS

[#] PP analysis is used. Because a considerable amount of patients couldn't be analysed in the PP population it is not sure if the baseline characteristics were still comparable or selective drop out is the case.

¥ QoL was measured by Retel et al. [13] in 347 (566 enrolled) patients of the MINDACT. The primary aims of the study were to evaluate the association between breast cancer patients' well-being and the results of a gene expression profile on to compare different recurrence risk groups, according to their genomic and standard clinical risk assessment. Different questionnaires were taken to assess the QoL. The QoL was assessed 6-8 weeks after surgery. This study does not compare QoL in the long term between CH/GL patients receiving treatment based on the MammaPrint® result and receiving treatment based on the AO! [13].

Table A 11. GRADE assessment; population with a CH/GL risk (PPS)

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Adjuvant! Online and MammaPrint®**	**Only Adjuvant! Online**	**Relative (95% CI)**	**Absolute (95% CI)**		
Overall survival after 10 years (follow up: median 5 years; assessed with: surrogate OS 5 year)												
1	Randomised trials	Not serious [a]	Not serious [b]	Serious [c]	Very serious [d]	None	525/542 (96.9%) *	495/503 (98.4%)*	**HR 0.54** (0.23 to 1.26)	**91 fewer per 1.000** (from 10 more to 370 fewer)	⊕◯◯◯ VERY LOW	CRUCIAL
Overall survival after 10 years (follow up: median 5 years; assessed with: surrogate DMFS 5y)												
1	Randomised trials	Not serious [e]	Not serious [b]	Serious [e]	Serious [e]	None	509/542 (93.9%)*	485/503 (96.4%)*	**HR 0.60** (0.34 to 1. 06)	**100 fewer per 1.000** (from 6 more to 287 fewer)	⊕⊕◯◯ LOW	CRUCIAL
Overall survival after 10 years (follow up: median 5 years; assessed with: DFS 5y)												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Adjuvant! Online and MammaPrint®**	**Only Adjuvant! Online**	**Relative (95% CI)**	**Absolute (95% CI)**		
1	Randomised trials	Serious [f]	Not serious [b]	Serious [f]	Serious [f]	None	481/542 (88.7%)*	471/503 (93.6%)*	**HR 0.57** (0.37 to 0.87)	**144 fewer per 1.000** (from 27 fewer to 297 fewer)	⊕⊕◯◯ LOW	CRUCIAL
Quality of life - not measured[g]												
-	-	-	-	-	-	-					-	CRUCIAL
Short-term and long-term side effects from chemo - not reported												
-	-	-	-	-	-	-					-	CRUCIAL

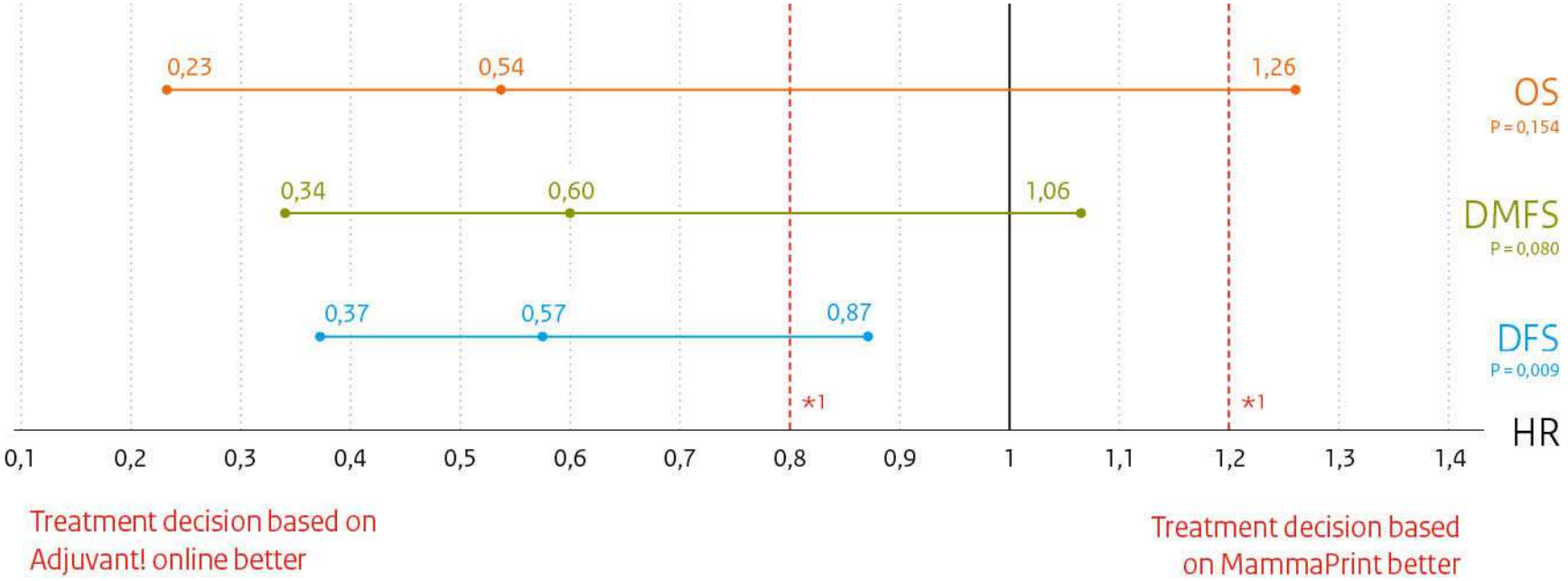
Abbreviations: CI: Confidence interval; HR: Hazard Ratio

[a] We did not downgrade separately for risk of bias, although there is a risk of attrition bias. Of the patients randomised to chemotherapy (n=749) ,174 (23%) patients were not included in the PP analysis (of whom 128 did not receive chemotherapy and 26 had a change of risk). Of those assigned to no chemotherapy (n=748), 119 (16%) were not included in the PP analysis (of whom 85 received chemotherapy and 21 had a change of risk). No lost to follow up data is mentioned. Because a considerable number of patients could not be analysed in the PP analysis, it is not clear if the baseline characteristics were still comparable or there was selective dropout. The results do not show if both groups are comparable. Furthermore, patients, caregivers, and assessors were not blinded, but we did not deem it necessary to downgrade because OS is an objective outcome measure. Overall, we find it too strict to downgrade again.
[b] There is only one publication so inconsistency is not possible.

[c] Because ten-year follow-up is essential, we have downgraded because only five-year data is available.
[d] Using the thresholds of clinical relevance of the ESMO (ESMO-MCBS) (HR<0.65 or <0.65 HR <0.80), the 95% CI crosses the thresholds at both sides (independent if you make use of the threshold of 0.65 or 0.80). That is the reason for downgrading twice.
[e] In total, we have downgraded twice because of: risk of bias (attrition bias and detection bias), indirectness (the relation between 5-year DMFS and 10-year OS is not validated. If the standard care AO! is not used in your country, you should consider downgrading an extra time for indirectness), and imprecision (using the ESMO-MCBS criteria, the lower boundary of 95% CI crosses line of clinical relevance (see **Figure A2**).
[f] In total we have downgraded twice because of: risk of bias (attrition bias and detection bias), indirectness the relation between 5-year DMFS and 10-year OS is not validated. If the standard care AO! is not used in your country, you should consider downgrading an extra time for indirectness), and imprecision (using the ESMO-MCBS criteria, the lower boundary of 95% CI crosses line of clinical relevance (see **Figure A2**)).
[g] QoL was measured by Retel et al. [13] in 347 (566 enrolled) patients of the MINDACT. The primary aims of the study were to evaluate the association between breast cancer patients' well-being and the results of a gene expression profile on to compare different recurrence risk groups, according to their genomic and standard clinical risk assessment. Different questionnaires were taken to assess the QoL. The QoL was assessed 6-8 weeks after surgery. This study does not compare QoL in the long term between CH/GL patients receiving treatment based on the MammaPrint® result and receiving treatment based on the AO! [13].

* Our calculations differ a little from the published numbers. The calculation for five-year survival percentage takes into account the time at risk for each individual in the group up to 5 years. So, that includes the time at risk for each patient either with event (event time cut-off), or without event (censoring time cut-off). This is probably the reason why the percentages in the GRADE table differ a little from the percentages in the MINDACT publication.

Figure A 2. PPS-analysis: HR of OS, DMFS and DFS with 95% CI.



Abbreviations: DFS: disease-free survival; DMFS: distant metastasis-free survival; HR: hazard ratio; OS: overall survival.
*1 = thresholds for clinical relevance.

Table A 12. GRADE assessment; population with a CL/GH risk (ITT)

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Adjuvant! Online and MammaPrint®**	**Only Adjuvant! Online**	**Relative (95% CI)**	**Absolute (95% CI)**		
Overall survival after 10 years (follow up: median 5 years; assessed with: OS 5 year)												
1	Randomised trials	Not serious [a]	Not serious [b]	Serious [c,d]	Very serious [e]	None	333/344 (96.8%) *	336/346 (97.1%) *	**HR 1.28[#]** (0.54 to 3.02)	**18 more per 1.000** (from 29 more to 119 fewer)	⊕◯◯◯ VERY LOW	CRUCIAL
Overall survival after 10 years (follow up: median 5 years; assessed with: DMFS 5y)												
1	Randomised trials	Not serious [a]	Not serious [b]	Serious [d,f]	Very serious [e]	None	326/344 (94.8%) *	329/346 (95.1%) *	**HR 1.17[#]** (0.59 to 2.28)	**20 more per 1.000** (from 48 more to 120 fewer)	⊕◯◯◯ VERY LOW	CRUCIAL
Overall survival after 10 years (follow up: median 5 years; assessed with: DFS 5y)												

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Adjuvant! Online and MammaPrint®**	**Only Adjuvant! Online**	**Relative (95% CI)**	**Absolute (95% CI)**		
1	Randomised trials	Not serious [a]	Not serious [b]	Serious [d,g]	Very serious [e]	None	316/344 (91.9%) *	312/346 (90.2%) *	**HR 0.87[#]** (0.53 to 1.45)	**35 fewer per 1.000** (from 64 more to 194 fewer)	⊕◯◯◯ VERY LOW	CRUCIAL
Quality of life - not measured[**]												
-	-	-	-	-	-	-					-	CRUCIAL
Short-term and long-term side effects from chemo - not reported												
-	-	-	-	-	-	-					-	CRUCIAL

Abbreviations: CI: Confidence interval; HR: hazard ratio.
[a] We did not downgrade. There is some risk of bias because the patients, caregivers and assessors were not blinded. For OS, DMFS, and DFS, this will not make a difference if patients and caregivers were blinded or not. However, it may influence the assessor in his opinion if there is a metastasis of other form of progression. We find it too strict to downgrade for risk of bias.
[b] There is only one publication so inconsistency is not possible
[c] Because ten-year follow-up is essential, we have downgraded because only five-year data is available.
[d] We did not downgrade for the standard care (Adjuvant! Online) because in Belgium and the Netherlands this is according to clinical guidelines. If this standard is very different in other countries, you should consider downgrading an extra time for indirectness.
[e] Using the thresholds of clinical relevance of the ESMO (ESMO-MCBS) (HR<0.65 or <0.65 HR <0.80), the 95% CI crosses the thresholds at both sides (independent if you

make use of the threshold of 0.65 or 0.80). That is the reason why it is downgraded twice.
[f] We have downgraded once because the relation between 5-year DMFS and-10 year OS is not validated.
[g] We've downgraded because the relation between 5-year DFS and 10-year OS is not validated.
* Our calculations differ a little from the published numbers. . The calculation for five-year survival percentage takes into account the time at risk for each individual in the group up to 5 years. So, that includes the time at risk for each patient either with event (event time cut-off), or without event (censoring time cut-off). This is probably the reason why the percentages in the GRADE table differ a little from the percentages in the MINDACT publication.
[#] The HRs do not correspond with the survival data (DFS and OS), possibly because the curves are crossing.
** QoL was measured by Retel [13] et al. in 347 (566 enrolled) patients of the MINDACT. The primary aims of the study were to evaluate the association between breast cancer patients' well-being and the results of a gene expression profile on to compare different recurrence risk groups, according to their genomic and standard clinical risk assessment. Different questionnaires were taken to assess the QoL. The QoL was assessed 6-8 weeks after surgery. This study does not compare QoL in the long term between CH/GL patients receiving treatment based on the MammaPrint® result and receiving treatment based on the AO! [13].

Figure A 3. ITT analysis: HR of OS, DMFS and DFS with 95% CI.

Abbreviations: DFS: disease-free survival; DMFS: distant metastasis-free survival; HR: hazard ratio; OS: overall survival.
*1 = thresholds for clinical relevance.

9.3.4 Applicability tables

Table A 13. Summary table characterising the applicability of a body of studies

Domain	Description of applicability of evidence
Population	The majority of enrolled patients are patients with ER-positive and HER2-negative tumours. These patients with ER-positive and HER2-negative tumours are relevant to the target population of early breast cancer patients: in particular in the subset of patients with ER positive and HER2-negative tumours in which there is controversy regarding the benefit of adjuvant chemotherapy, and hence controversy on decisions on adjuvant chemotherapy.
Intervention	MammaPrint® is a genomic expression signature test: see considerations described in the 'Population', 'Comparators' and 'Outcomes' domain.
Comparators	The comparator in the MINDACT study is Adjuvant! Online. Although Adjuvant! Online has been globally investigated in validation studies, some concerns have been raised regarding their applicability in populations other than those used in their validation studies. Currently Adjuvant! Online is being updated. An update potentially can lead to impact on the baseline risks of recurrence and hence this may potentially limit the clinical applicability of the MINDACT results.
Outcomes	Wide confidence intervals were found for all surrogate outcome measures. The confidence intervals show there is a lot of uncertainty and possibly many patients could be harmed in terms of hazards of death.
Setting	See considerations described in the 'Comparators' domain.

Table A 14. Outcome according to discordant risk group CH/GL and treatment strategy

Risk group, Outcome, and treatment strategy	**Chemo-therapy**	**No. of patients**	**No. of events**	**Percentage at 5 years (95% CI)**	**Hazard ratio (95% CI)**	**P value**
High clinical risk and low genomic risk (PPS)						
*Overall survival (5y) (**OS**)*						
• Using clinical risk	Yes	503	8	98.8 (97.1-99.5)	0.57 (0.23 – 1.26)	0.154
• Using genomic risk (add on)	No	542	17	97.0 (94.9-98.2)	1.00	
*Survival without distant metastasis 5y **(DMFS)***						
• Using clinical risk	Yes	503	18	96.5 (94.1-97.9)	0.60 (0.34-1.06)	0.080
• Using genomic risk (add on)	No	542	33	94.0 (91.4-95.8)	1.00	
Disease-free survival 5y **(DFS)**						
• Using clinical risk	Yes	503	32	93.3 (90.3-95.4)	0.57 (0.37-0.87)	**0.009**
• Using genomic risk (add on)	No	542	61	88.8 (85.7-91.3)	1.00	
High clinical risk and low genomic risk (PP)						
*Overall survival (5y) (**OS**)*						
• Using clinical risk	Yes	592	10	98.8 (97.4-99.5)	0.63 (0.29 – 1.37)	0.25
• Using genomic risk (add on)	No	636	18	97.3 (95.6-98.4)	1.00	
*Survival without distant metastasis 5y **(DMFS)***						

• Using clinical risk	Yes	592	22	96.7 (94.7-98.0)	0.65 (0.38-1.10)	0.11
• Using genomic risk (add on)	No	636	37	94.8 (92.6-96.3)	1.00	
Disease-free survival 5y **(DFS)**						
• Using clinical risk	Yes	592	39	93.3 (90.7-95.2)	0.64 (0.43-0.95)	**0.03**
• Using genomic risk (add on)	No	636	66	90.3 (87.6-92.4)	1.00	
High clinical risk and low genomic risk (ITT)						
*Overall survival (5y) (**OS**)*						
• Using clinical risk	Yes	749	14	98.4 (97.0-99.1)	0.69 (0.35-1.35)	0.278
• Using genomic risk (add on)	No	748	22	97.0 (95.4-98.1)	1.00	
*Survival without distant metastasis 5y (**DMFS**)*						
• Using clinical risk	Yes	749	34	95.9 (94.0-97.2)	0.78 (0.50-1.21)	0.267
• Using genomic risk (add on)	No	748	46	94.4 (92.3-95.9)	1.00	
Disease-free survival 5y **(DFS)**						
• Using clinical risk	Yes	749	54	92.9 (90.5-94.7)	0.71 (0.50-1.01)	0.055
• Using genomic risk (add on)	No	748	78	90.1 (87.5-92.1)	1.00	

Table A 15. Outcome according to discordant risk group CL/GH and treatment strategy

Risk group, Outcome, and treatment strategy	**Chemotherapy**	**No. of patients**	**No. of events**	**Percentage with outcome at 5 years (95% CI)**	**Hazard ratio (95% CI)**	**P value**
Low clinical risk and high genomic risk (ITT)						
Survival without distant metastasis 5y ***(DMFS)***						
• Using genomic risk (add on)	Yes	344	18	95.8 (92.9-97.6)	1.17* (0.59-2.28)	0.657
• Using clinical risk	No	346	17	95.0 (91.8-97.0)	1.00	
Disease-free survival 5y **(DFS)**						
• Using genomic risk (add on)	Yes	344	28	92.1 (88.3-94.6)	0.87* (0.53-1.45)	0.603
• Using clinical risk	No	346	34	90.1 (86.1-93.0)	1.00	
*Overall survival (5y) (****OS****)*						
• Using genomic risk (add on)	Yes	344	11	97.1 (94.5-98.5)	1.28* (0.54-3.02)	0.578
• Using clinical risk	No	346	10	97.8 (95.5-99.0)	1.00	

* In the case of DFS and OS, the HR does not correspond with the percentage outcome after five years, a possible explanation being that the curves of the MammaPrint® and AO! groups cross each other. The HRs published in MINDACT are adjusted, but it is not clear if they are adjusted for these crossings.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4304842/

Table A 16. Outcome according for different pre-specified subgroups and treatment strategy

Risk group, Outcome, and treatment strategy	**Chemotherapy**	**No. of patients**	**No. of events**	**Percentage with outcome at 5 year (95% CI)**	**Adjusted Hazard ratio (95% CI)**	**P value**
High genomic risk (clinical high and clinical low)		2398				
*Survival without distant metastasis 5y **(DMFS)** using only genomic risk*				94.7*		
High clinical risk (genomic high and genomic low)		3356				
*Survival without distant metastasis 5y **(DMFS)** using only clinical risk*				95.0*		
Cordant groups						
High clinical risk and high genomic risk	Yes	1806				
OS 5Y			103	94.7 (93.4–95.7)		
DMFS			171	90.6 (89.0-92.0)		
DFS			266	85.3 (83.4-87.0)		
Low clinical risk and low genomic risk	No	2745				
OS 5Y			47	98.4 (97.8–98.9)		
DMFS			77	97.6 (96.9-98.1)		
DFS			211	92.8 (91.7-93.7)		
Subgroup: by nodal status (ITT1 population)						
Clinical high/genomic low LN0						
DMFS						

Using clinical risk	Yes	395		95.7 (93.0-97.4)	0.69 (0.39-1.21)	0.19
Using genomic risk (add on)	No	392		93.2 (90.1-95.4)	1.00	
Clinical high/genomic low N+						
DMFS						
Using clinical risk	Yes	353		96.3 (93.1-98.1)	0.88 (0.42-1.82)	0.72
Using genomic risk (add on)	No	356		95.6 (92.7-97.4)	1.00	
Clinical low/genomic high LN0						
DMFS						
Using clinical risk	No	333		95.1 (91.9-97.1)	1.00	0.82
Using genomic risk (add on)	Yes	333		96.0 (93.1-97.7)	1.09 (0.54-2.19)	
Clinical low/genomic high N+						
DMFS		Group too small to be analysed				
Subgroup: by tumour size (ITT1 population)						
Clinical high/genomic low T2						
DMFS						
Using clinical risk	Yes	402		94.5 (91.4-96.6)	0.90 (0.53-1.54)	0.71
Using genomic risk (add on)	No	406		93.7 (90.6-95.8)	1.00	
Clinical high/genomic low T1						
DMFS						
Using clinical risk	Yes	322		97.6 (95.1-98.9)	0.59 (0.26-1.33)	0.20
Using genomic risk (add on)	No	314		94.8 (91.5-96.8)	1.00	

Clinical low/genomic high T1						
DMFS						
Using clinical risk	No	338		94.9 (91.6-96.9)	1.00	0.85
Using genomic risk (add on)	Yes	333		95.7 (92.6-97.5)	1.07 (0.54-2.12)	
Clinical low/genomic high T2						
DMFS		Group too small to be analysed				
Subgroup: HR-/+/HER2-/+/LN0 (ITT1 population)						
Clinical high/genomic low HR+/HER2-/LN0						
DMFS						
Using clinical risk	Yes	349		95.5 (92.5-97.3)	0.80 (0.44-1.45)	0.46
Using genomic risk (add on)	No	350		93.9 (90.6-96.1)	1.00	
Clinical low/genomic high HR-/HER2+/LN0						
DMFS						
Using clinical risk	No	262		95.5 (91.6-97.6)	1.00	0.33
Using genomic risk (add on)	Yes	272		95.1 (91.5-97.2)	1.45 (0.68-3.08)	

* To have an unbiased estimate, the discordant patients are doubly weighted, because they are underrepresented by a factor of two in the resulting sample. Therefore, comparison by means of classical statistical inference is incorrect and only the estimates of the five-year DMFS are shown.

10 APPENDIX 2: CHECKLIST FOR POTENTIAL ETHICAL, ORGANISATIONAL, PATIENT, SOCIAL AND LEGAL ASPECTS

1. Ethical	
1.1. Does the introduction of the new technology and its potential use/non-use instead of the defined, existing comparator(s) give rise to any new ethical issues?	Yes
Introduction of MammaPrint® at this moment (with only the availability of five-year data with wide CI's and therefore a lot of uncertainty) can lead to some ethical issues both for the patient as well as for the healthcare provider. For the patient, there is the ethical issue of a trade-of between survival benefit versus deterioration in quality of life. For the healthcare provider, the ethical issue is: how to comply with the principle of clinical utility, in this case accepting a potential aggressive treatment without expecting health benefit versus how to comply with the principle to adhere to patient autonomy and shared decision-making.	
1.2. Does comparing the new technology to the defined, existing comparators point to any differences that may be ethically relevant?	Yes
See 1.1. The current data does not point to a statistical difference between chemotherapy following clinical high-risk assignment and no chemotherapy following MammaPrint® low-risk assignment in the primary endpoint of DMFS. However, it cannot be concluded that MammaPrint® is non-inferior because the 95% CI of five-year DMFS and 5-year DFS are crossing the non-inferiority threshold. Only when the entire 95% CI is above the non-inferiority threshold it can be concluded that MammaPrint® is non-inferior compared with AO!. As pointed out in the results and discussion, it cannot be excluded that adjuvant chemotherapy following the clinical risk assignment significantly decreases the hazard of death due to distant metastasis. So, when MammaPrint® is introduced prematurely, omitting chemotherapy following genomic risk assignment may place patients at increased risk of premature death due to metastatic disease which could have been prevented by adjuvant chemotherapy.	
2. Organisational	
2.1. Does the introduction of the new technology and its potential use/non-use instead of the defined, existing comparator(s) require organisational changes?	No
2.2. Does comparing the new technology to the defined, existing comparator(s) point to any differences that may be organisationally relevant?	No
3. **Social**	
3.1. Does the introduction of the new technology and its potential use/non-use instead of the defined, existing comparator(s) give rise to any new social issues?	Yes
Yes, if a patient can safely omit chemotherapy she will probably be able to keep working as normal. On the other side, more deaths can possibly be expected.	

3.2. Does comparing the new technology to the defined, existing comparator(s) point to any differences that may be socially relevant?	Yes
See 3.1	
4. **Legal**	
4.1. Does the introduction of the new technology and its potential use/non-use instead of the defined, existing comparator(s) give rise to any legal issues?	No
4.2. Does comparing the new technology to the defined, existing comparator(s) point to any differences that may be legally relevant?	No

A separate appendix with comments from external experts and the MAH/manufacturer(s), as well as responses from authors, for the purposes of transparency, is published separately.