THE CLINICAL UTILITY OF MOLECULAR DIAGNOSTICS IN BREAST CANCER MANAGEMENT

SOFIE BERGHUIS

THE CLINICAL UTILITY OF MOLECULAR DIAGNOSTICS IN BREAST CANCER MANAGEMENT

USING BIOMARKERS TO OPTIMIZE PERSONALIZED MEDICINE

ANNE MARGREET BERGHUIS

THE CLINICAL UTILITY OF MOLECULAR DIAGNOSTICS IN BREAST CANCER MANAGEMENT

USING BIOMARKERS TO OPTIMIZE PERSONALIZED MEDICINE

DISSERTATION

To obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. T.T.M. Palstra,
on account of the doctorate board,
to be publicly defended
on Friday the 10th of January 2020 at 12.45 hours

By

Anne Margreet Berghuis

Born on the 4th of February 1992
in Deventer, the Netherlands

THIS DISSERTATION HAS BEEN APPROVED BY:

SUPERVISORS

Prof. dr. Maarten J. IJzerman

Dr. Hendrik Koffijberg

PARANYMPHS

Annelies Klos

Deon Lingervelder

The clinical utility of molecular diagnostics in breast cancer management. Using biomarkers to optimize personalized medicine.

This thesis is part of the Health Science Series, HSS 20-30, department Health Technology and Services Research, University of Twente, Enschede, The Netherlands. ISSN: 1878-4968.

FINANCIAL SUPPORT FOR PRINTING OF THIS THESIS WAS KINDLY PROVIDED BY:

UNIVERSITY OF TWENTE.

Cover Design: Adam de Beer

Lay-out: Kevin van Dijk

Printed by: Ipskamp printing, Enschede

ISBN: 978-90-365-4906-6

DOI: 10.3990/1.9789036549066

© Copyright 2019: Sofie Berghuis, Enschede, the Netherlands. No parts of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author.

GRADUATION COMMITTEE

CHAIRMAIN

Prof. Dr. Th. A.J. Toonen	University of Twente
---------------------------	----------------------

SUPERVISORS

Prof. dr. M.J. IJzerman	University of Twente
Dr. H. Koffijberg	University of Twente

COMMITTEE MEMBERS

Prof. Dr. H. Neubauer	Heinrich-Heine University Düsseldorf
Prof. Dr. E.M.D. Schuurin	University Medical Center Groningen
Prof. Dr. G.H. de Bock	University Medical Center Groningen
Dr. V.M.H. Coupé	VU University Medical Center
Prof. Dr. J. Prakash	University of Twente
Prof. Dr. S. Siesling	University of Twente
Dr. C.G.M. Groothuis-Oudshoorn*	University of Twente

* referee

CONTENTS

Chapter 1 - General Introduction	12
Chapter 2 - Evidence on the cost of breast cancer drugs is required for rational decision making	24
Chapter 3 - Real world data on discordance between estrogen, progesterone and HER2 receptor expression on diagnostic tumor biopsy versus tumor resection material	38
Chapter 4 - Treatment choices for neo- and adjuvant systemic therapy following repeated assessment of estrogen, progesterone and HER2 in tumor biopsy and resection in invasive breast cancer patients	58
Chapter 5 - Detecting blood based biomarkers in metastatic breast cancer: A systematic review of their current status and clinical utility	80
Chapter 6 - Health Economic Impact of Liquid Biopsies in Cancer Management	98
Chapter 7 - Rapid assessment of the health economic impact of using CTCs to guide systemic therapy in metastatic breast cancer - an online tool to inform the design of prospective biomarker studies	116
Chapter 8 - Health economic impact of using CTCs to guide systemic therapy in metastatic ER+ HER2- breast cancer patients – Results from a randomized controlled multicenter trial	134
Chapter 9 - General Discussion	152
Summary	168
Samenvatting	174
Acknowledgements	180
Curriculum Vitae	190
List of publications	194

CHAPTER I

GENERAL INTRODUCTION

BREAST CANCER PROGNOSIS AND SURVIVAL

Cancer is a disease which is characterized by normal cells changing into abnormal cells which subsequently causes abnormal cell growth. Several factors may cause healthy cells to evolve into tumor cells, which are generally referred to as the hallmarks of cancer. These hallmarks are biological characteristics that enable cells to be able to keep growing by overcoming cell death, sustaining proliferative signaling or evading growth suppressors, inducing angiogenesis and form new blood vessels around the tumor, activating the invasion of these blood vessels and potentially form a metastasis, reprogramming the energy metabolism and evading the destruction of these cells by the immune system. When cells successfully acquire (or have) one of these biological characteristics, they simultaneously increase their potential to grow into tumor cells [1].

Approximately 38% of the whole population will develop cancer at some point in their lives. The most common form of cancer among women is breast cancer, accounting for 30% of all female cancers [2]. Currently, already 1 in 6.6 women develop cancer and the incidence of breast cancer is still rising [3]. Although the mortality rates of breast cancer have significantly decreased over the last decades, due to better diagnostics and improved and more personalized treatment options, breast cancer still is the second cause of cancer death in women, with a standardized mortality rate of 12.9 (per 100,000) [2-6]. Nonetheless, for most subgroups of breast cancer patients, the average 5-year survival is relatively high, especially in Western countries, with 5-year survival rates of over 85% [5]. For metastasized breast cancer, the survival is substantially worse than for most other subgroups of breast cancer, with 5-year survival rates of even below 25% [7].

Currently, breast cancer is detected either by self-examination and clinical examination or by discovering suspicious tissue on the mammography taken in the population screening. Following detection, additional diagnostic imaging such as mammography, echography, MRI or combinations of those will be performed to confirm the suspicion of cancer. Based on the results of these tests, obtaining an MRI guided biopsy is indicated to confirm the diagnosis of breast cancer based on the histology and its grade. Subsequently, the expression of the estrogen receptor (ER), progesterone receptor (PR), Human Epidermal Growth Factor Receptor 2 (HER2) and Ki67 is determined [8].

Following, further staging of the tumor is performed, in which the tumor characteristics, tumor size and the exact location of the lesion are described. In case

of suspicion for metastases, additional tests are performed to determine whether any lymph nodes are involved or whether there are distant metastases present [9]. To improve on the survival rates and to potentially prevent that metastatic breast cancer develops, it is important to detect cancer as early as possible.

COST OF CANCER TREATMENT

While significant improvements have been made in the delivery of care for breast cancer patients, the main challenge in healthcare and specifically in oncology nowadays, is the persistent increase in healthcare costs. The share of the total health care expenditures that can be allocated to the cost of cancer treatment is rising and is currently already accounting for almost 6% of the total health care expenditures in the Netherlands [10]. The largest share of the cost of cancer treatment can be allocated to the costs of the treatment of breast cancer patients, accounting for a total of 13% of all cancer related health care costs [11]. For most cancers, the major component of these costs is hospitalization, or inpatient care, that accounts for up to 30% of the total hospital expenditures [12]. However, for breast cancer, it is the vast majority of the costs (46%) that can be allocated to drug expenditures [11]. In addition, a growing component of the total costs of cancer treatment in general can be allocated to drug expenditures, ranging from 3 to 61% of the cancer health expenditures in different European countries [10]. Especially the molecularly targeted treatments, such as Trastuzumab and Pertuzumab, are rather expensive [13].

In addition, more and more costly new cancer treatments keep being developed of which the costs may cause inequity in the access to the particular cancer treatments between countries [14]. Therefore, it is important to further improve on selection mechanisms or diagnostics to identify patients who are likely to benefit (or not) from those treatments. Better targeted and more personalized medicine may reduce over- and under treatment, thereby reducing costs and improving on future cancer care affordability and potentially survival as well.

PERSONALIZED TREATMENT BASED ON BIOMARKER TARGETS

Whether breast cancer patients are eligible for surgery, eventually in combination with radiation, neo- or adjuvant systemic therapy, is dependent on the exact tumor stage, other patient- and tumor characteristics and potential contra indications [9]. ER, PR and HER2 are biological cell characteristics, or so called biomarkers, measured and used to define whether it is potentially valuable to prescribe hormone or targeted therapy and to establish a prognosis [15]. Currently, the expression of ER, PR and

HER2 is typically determined by immunohistochemistry (IHC) or in situ hybridization (ISH) using small samples of tumor specimen, or tissue biopsies, derived from the primary tumor [16].

Although other biomarkers are available, these three biomarkers still are the most important in characterizing the tumor as the optimal type of systemic treatment prescribed, subsequent treatment responses and survival are substantially influenced by the expression of these biomarkers [17]. Patients expressing ER and PR may be eligible for hormone therapy whereas patients expressing HER2 may be eligible for receiving targeted anti-HER2 agents [9]. For patients presenting without any receptor expression, the triple negative breast cancer patients, prognosis and survival is worse than for patients that do express one of these receptors [18].

In the past decades, new treatments that substantially improve the clinical outcome for breast cancer patients became available on the market. Significant survival gain has been realized using targeted treatments such as Trastuzumab and Pertuzumab tailoring the HER2 receptor and Bevacizumab tailoring the Vascular Endothelial Growth Factor Receptor (VEGFR) [19-21]. Most of these new targeted therapies are tailoring the HER2 receptor [22]. Even though significant progress has been made regarding treatments available for different subtypes of breast cancer, there still is a lack of diagnostics that are sensitive enough to detect the presence of potential (micro-) metastases.

Furthermore, supposing that the metastases can be detected, the biological characteristics of the metastasis are not necessarily similar to those of the primary tumor. This biological heterogeneity makes it even more difficult to select the optimal (combination of) treatments. In addition, even though initially the metastatic tumor characteristics are not different from the primary tumor, cancer cells have shown to have the potential to change over time in order to survive [1]. Despite the prescription of targeted treatment tailoring a specific receptor that expressed positive, the tumor cell characteristics may change in order to escape the toxic tailored drugs administered.

To be able to detect the tumor characteristics correctly and subsequently initially prescribe the most appropriate treatment, sensitive diagnostics that capture the tumor characteristics that cover both, the primary tumor and potential (micro-) metastases, should be used. The value of such diagnostics further increases if these can also be used as part of monitoring strategies. In this context, frequently repeated measurements of the treatment response and potential changes in the biological

characteristics of the tumor should be performed. However, taking repeated tumor needle biopsies to test the characteristics of the primary lesion or metastasis is a very high burden for the patients, if it is even possible to do so for the particular location of the (metastatic) lesion. Therefore, to enable diagnostics to additionally be used as monitoring strategies, less invasive and more sensitive diagnostic tests should be developed.

LIQUID BIOPSIES

In the last decades, several initiatives have aimed at developing better diagnostic and monitoring imaging tools [23]. Before cancer metastasizes, cells detach from the primary tumor and enter the blood stream [24]. The presence of such cells, or parts of these cells, in the blood stream may already provide insight in the potential of the tumor to metastasize. In addition, the characterization of these tumor particles may enable earlier and better targeted treatment prescriptions which may subsequently eliminate those tumor cells that are potentially most dangerous.

An emerging and innovative field which aims to develop more sensitive and less invasive diagnostics is focusing on liquid biopsies (LBs). LBs allow for earlier detection of cancer and the detection of potential tumor heterogeneity, by sampling biomarker characteristics from particles that have detached from the primary tumor. Such particles can either be whole circulating cells, cell particles, circulating tumor DNA (ctDNA) or RNA molecules derived from body fluids (mainly blood) [25].

CHALLENGES OF USING NEW BIOMARKERS TO IMPROVE PATIENT CARE

Several studies have evaluated the predictive and prognostic value of multiple different types of LBs [26, 27]. The most extensively investigated LBs so far are circulating tumor cells (CTCS) and ctDNA. Both have demonstrated to be valuable companion diagnostics at initial diagnosis and they have both shown their suitability as monitoring diagnostics, as these were able to detect drug resistance mutations [27, 28]. For CTCs it has been demonstrated that their presence is associated with a significantly lower progression free survival (PFS) and overall survival (OS) [29]. Increased levels of mutations within ctDNA reflect progressive disease in metastatic breast cancer, but can also identify early disease relapse in early breast cancer [30, 31].

Such innovative techniques can, however, only be successfully implemented in clinical practice when the technologies are analytically validated and have demonstrated their value in clinical trials [26, 32]. The gold standard to evaluate and validate the effectiveness and clinical utility of innovations are randomized clinical trials (RCTs)

[33]. However, biomarkers in general and LBs are frequently measured as a number or percentage of cells that express a particular biomarker. Based on expression thresholds, such biomarkers are then categorized as positive or negative and subsequently used to guide treatment decisions. However, the impact of the selected value of these thresholds is difficult to evaluate in clinical trials, as evaluating multiple thresholds directly requires including a multiple of the otherwise included number of patients, thereby substantially increasing trial costs.

Before new technologies or biomarkers will be implemented in routine clinical practice, validation and reimbursement of these technologies is necessary, which is quite often an even more complex process for diagnostics, as they may influence particular pharmaceutical prescriptions. Therefore, it is important that their clinical impact and health economic value is known. Even though RCTs are currently most frequently used to evaluate the effectiveness, prognostic value or any other performance measure for new technologies, alternative methods may be valuable to evaluate the validity and health economic potential of new biomarkers. Furthermore, observational studies that closely reflect (the variation in) clinical practice may even be more suitable than RCTs to determine the external validity of diagnostic or monitoring technology [33]. Therefore, it is valuable to investigate and apply novel methods or tools that enable to evaluate the clinical impact and health economic value of molecular diagnostics in an alternative way.

OUTLINE THESIS

This thesis presents a collection of papers regarding the clinical and health economic value of molecular diagnostics in breast cancer patients. Several issues regarding the extent to which these diagnostics have been investigated, their potential impact on clinical practice, and the difficulties in addressing their cost-effectiveness or health economic potential, have been determined.

In the first part of this thesis, several components of usual breast cancer care were evaluated. To provide a rationale for better use of biomarkers, we addressed which drugs are currently approved in the Netherlands and at which costs an average dose of these drugs may be prescribed. **Chapter 2** shows the results of the review in which prescription information and the costs of all breast cancer drugs in the Netherlands were identified [34].

Especially the ER and HER2 status are crucial in guiding systemic breast cancer treatment [35]. Following, it was determined whether traditionally used molecular

diagnostics have the potential to capture tumor heterogeneity. Potential differences in the status of the ER, PR and HER2 were determined between tumor needle biopsy and tumor resection material in **Chapter 3**. Discordance in either one of these receptors may cause deviations from the expected type of treatment prescribed. **Chapter 4** identified whether treatment patterns associated with particular receptor expressions on biopsy and resection material are present.

In the second part of this thesis, we have evaluated the health economic potential of LBs. All new and emerging developments regarding LBs were identified for patients with metastatic breast cancer. **Chapter 5** presents a systematic review on the use of LBs in metastatic breast cancer and addressed which type of LBs are currently most frequently investigated and in which research phase [36]. Following, **Chapter 6** presents an expert review in which the health economic potential and remaining challenges of using LBs in cancer management is further described [37].

However, evidence becomes frequently available as a multiple of studies towards CTCs are currently still ongoing. Updating the evidence of early health economic models or exploring different scenarios or CTC thresholds for treatment decisions other than investigated in those studies may be time consuming or difficult. **Chapter 7** presents an online tool that may be used to explore and update this evidence to rapidly get new insights in the health economic potential of CTCs.

Finally, **Chapter 8** presents the cost-effectiveness analysis of the STIC METABREAST trial, which is a large multicenter RCT that has investigated the value of using CTCs to guide the decision to use either chemo- or hormone therapy in metastatic breast cancer patients.

REFERENCES

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-74.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(1):7-30.
3. van der Waal D, Verbeek AL, den Heeten GJ, Ripping TM, Tjan-Heijnen VC, Broeders MJ. Breast cancer diagnosis and death in the Netherlands: a changing burden. *European journal of public health*. 2015;25(2):320-4.
4. McDonald ES, Clark AS, Tchou J, Zhang P, Freedman GM. Clinical Diagnosis and Management of Breast Cancer. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*. 2016;57 Suppl 1:9s-16s.
5. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet (London, England)*. 2018;391(10125):1023-75.
6. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359-86.
7. Altekruse SF KC, Krapcho M, Neyman N, Aminou R, Waldron W, Ruhl J, Howlader N, Tatalovich Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Cronin K, Chen HS, Feuer EJ, Stinchcomb DG, Edwards BK. SEER Cancer Statistics Review, 1975-2007, National Cancer Institute.
8. Senkus E, Kyriakides S, Ohno S, PenaultLlorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2015;26 Suppl 5:v8-30.
9. NABON. Breast Cancer Guideline. 2012.
10. Jonsson B, Hofmarcher T, Lindgren P, Wilking N. The cost and burden of cancer in the European Union 1995-2014. *Eur J Cancer*. 2016;66:162-70.
11. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *The Lancet Oncology*. 2013;14(12):1165-74.
12. van Harten WH, Wind A, de Paoli P, Saghatchian M, Oberst S. Actual costs of cancer drugs in 15 European countries. *The Lancet Oncology*. 2016;17(1):18-20.
13. Sullivan R, Peppercorn J, Sikora K, Zalcborg J, Meropol NJ, Amir E, et al. Delivering affordable cancer care in high-income countries. *The Lancet Oncology*. 2011;12(10):933-80.
14. Cherny N, Sullivan R, Torode J, Saar M, Eniu A. ESMO European Consortium Study on the availability, out-of-pocket costs and accessibility of antineoplastic medicines in Europe. *Ann Oncol*. 2016;27(8):1423-43.
15. Nalejska E, Maczynska E, Lewandowska MA. Prognostic and predictive biomarkers: tools in personalized oncology. *Molecular diagnosis & therapy*. 2014;18(3):273-84.
16. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, 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(8):1533-46.

17. Weiss A, Chavez-MacGregor M, Lichtensztajn DY, Yi M, Tadros A, Hortobagyi GN, et al. Validation Study of the American Joint Committee on Cancer Eighth Edition Prognostic Stage Compared With the Anatomic Stage in Breast Cancer. *JAMA Oncol.* 2018;4(2):203-9.
18. Li X, Yang J, Peng L, Sahin AA, Huo L, Ward KC, et al. Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast Cancer Res Treat.* 2017;161(2):279-87.
19. Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis JP. Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *J Natl Cancer Inst.* 2008;100(24):1780-91.
20. Loibl S, Gianni L. HER2-positive breast cancer. *Lancet (London, England).* 2017;389(10087):2415-29.
21. Cardoso F, Costa A, Senkus E, Aapro M, Andre F, Barrios CH, et al. 3rd ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 3). *Ann Oncol.* 2017;28(1):16-33.
22. Cardoso F, Spence D, Mertz S, Corneliussen-James D, Sabelko K, Gralow J, et al. Global analysis of advanced/metastatic breast cancer: Decade report (2005-2015). *Breast.* 2018;39:131-8.
23. Baltzer PAT, Kapetas P, Marino MA, Clauser P. New diagnostic tools for breast cancer. *Memo.* 2017;10(3):175-80.
24. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70.
25. Bardelli A, Pantel K. Liquid Biopsies, What We Do Not Know (Yet). *Cancer Cell.* 2017;31(2):172-9.
26. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. *Nat Rev Clin Oncol.* 2017;14(9):531-48.
27. Ravelli A, Reuben JM, Lanza F, Anfossi S, Cappelletti MR, Zanotti L, et al. Breast cancer circulating biomarkers: advantages, drawbacks, and new insights. *Tumour Biol.* 2015;36(9):6653-65.
28. Alix-Panabieres C, Pantel K. Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. *Cancer discovery.* 2016;6(5):479-91.
29. Bidard FC, Peeters DJ, Fehm T, Nole F, Gisbert-Criado R, Mavroudis D, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *The Lancet Oncology.* 2014;15(4):406-14.
30. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med.* 2013;368(13):1199-209.
31. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Science translational medicine.* 2015;7(302):302ra133.
32. Ignatiadis M, Piccart M. Liquid biopsy to test new treatment strategies in breast cancer: are we there yet? *Ann Oncol.* 2012;23(7):1653-5.
33. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of clinical nursing.* 2003;12(1):77-84.
34. Berghuis AMS, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ. Evidence on the cost of breast cancer drugs is required for rational decision making. *Ecancermedicalscience.* 2018;12:825.

35. Berghuis AMS, van Deurzen CHM, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ. Real-world data on discordance between estrogen, progesterone, and HER2 receptor expression on diagnostic tumor biopsy versus tumor resection material. *Breast Cancer Res Treat.* 2019.

36. Berghuis AMS, Koffijberg H, Prakash J, Terstappen LWMM, MJ IJzerman. Detecting Blood-Based Biomarkers in Metastatic Breast Cancer: A Systematic Review of Their Current Status and Clinical Utility. *Int J Mol Sci.* 2017;18(2).

37. IJzerman MJ, Berghuis AMS, de Bono JS, Terstappen LWMM. Health economic impact of liquid biopsies in cancer management. *Expert Rev Pharmacoecon Outcomes Res.* 2018:1-7.

CHAPTER 2

EVIDENCE ON THE COST OF BREAST CANCER DRUGS IS REQUIRED FOR RATIONAL DECISION MAKING

A.M.S. BERGHUIS, H. KOFFIJBERG, L.W.M.M. TERSTAPPEN,
S. SLEIJFER, M.J. IJZERMAN

THIS CHAPTER WAS PUBLISHED AS:
A.M.S. BERGHUIS, H. KOFFIJBERG, L.W.M.M. TERSTAPPEN, S. SLEIJFER, M.J. IJZERMAN. EVIDENCE ON
THE COST OF BREAST CANCER DRUGS IS REQUIRED FOR RATIONAL DECISION MAKING. E-CANCER MEDICAL
SCIENCE, 16-04-2018. DOI: [10.3332/ECANCER.2018.825](https://doi.org/10.3332/ECANCER.2018.825)

ABSTRACT

BACKGROUND

For rational decision making, assessing the cost-effectiveness and budget impact of new drugs and comparing the costs of drugs already on the market is required. In addition to value frameworks, such as the American Society of Clinical Oncology Value Framework and the European Society of Medical Oncology–Magnitude of Clinical benefit Scale, this also requires a transparent overview of actual drug prices. While list prices are available, evidence on treatment cost is not. This paper aims to synthesise evidence on the reimbursement and costs of high-cost breast cancer drugs in The Netherlands (NL).

METHODS

A literature review was performed to identify currently reimbursed breast cancer drugs in the NL. Treatment costs were determined by multiplying list prices with the average length of treatment and dosing schedule.

RESULTS

Comparing list prices to the estimated treatment cost resulted in substantial differences in the ranking of costliness of the drugs. The average mean treatment length was unknown for 11/31 breast cancer drugs (26.2%). The differences in the 15 highest-cost drugs were largest for Bevacizumab, Lapatinib and everolimus, with list prices of € 541, € 158, € 1,168 and estimated treatment cost of € 174,400, € 18,682 and € 31,207, respectively. The lowest-cost (patented) targeted drug is € 1,818 more expensive than the highest-cost (off-patent) generic drug according to the estimated drug treatment cost.

CONCLUSIONS

A lack of evidence on the reimbursement and cost of high-cost breast cancer drugs complicates rapid and transparent evidence synthesis, necessary to focus strategies aiming to limit the increasing healthcare costs. Interestingly, the findings show that off-patent generics (such as paclitaxel or doxorubicin), although substantially cheaper than patented drugs, are still relatively costly. Extending standardisation and increasing European and national regulations on presenting information on costs per cancer drug is highly recommended.

BACKGROUND

Globally, breast cancer is the most common form of cancer among women. In the past decades, several new treatments that improve clinical outcome became available on the market [1]. Systemic treatment with e.g. Trastuzumab, for which specific biomarkers are employed, is one of these survival improving new treatments [2]. When trials have shown the effectiveness or clinical benefit of a specific treatment, the treatment needs to be approved before it can be implemented and used in clinical practice. Several international regulatory agencies are responsible for these drug approval processes, such as the European Medicines Agency (EMA) and the Food and Drug Administration [3]. In addition, national bodies typically apply country specific regulations concerning drug approval.

A major challenge in oncology today is the swift increase in healthcare costs [4]. Currently, the costs of cancer treatment can even account for up to 30% of the total hospital expenditures [5]. As more and more new cancer treatments are being developed, it is important to anticipate on the increasing cancer costs in hospitals already during early development stages. To be able to compare future treatment cost of upcoming new drugs to the prices that are currently paid, and to put the costs of new drugs into perspective, an overview of the current drug treatment cost should be available. However, such an overview of the actual cost of breast cancer treatments in European countries is lacking [5].

List prices are available and of interest but are not truly reflecting the actual drug treatment cost, as these prices only show the maximum purchase costs for a particular drug [5]. Furthermore, these list prices are not suitable for the comparison of treatment cost, as drugs may be prescribed in different dosages, have different administration forms and can have different toxicities, thereby impacting the total drug treatment cost for each drug. In addition, these costs should have been determined for currently reimbursed drugs but are even more important for new drugs that are currently investigated in trials. As there is currently no overview of the actual cost of currently used breast cancer drugs, there is no reference towards what we are currently able to afford cancer drugs [6]. To anticipate on future treatment costs and to strive towards maintaining good affordability of cancer care, it is important that such a framework will be established. A preliminary step would be to gain insight in the actual treatment costs for high-cost breast cancer drugs by using European and national assessment reports.

In addition, there is increasing attention being given to the lack of transparency in drug pricing. Both the American Society of Clinical Oncology (ASCO) and the European Society for Medical Oncology (ESMO) have already developed value frameworks, which aim to guide decisions by comparing costs and net health benefits. The recent update of the ASCO Value Framework (ASCO-VF) shows that costs included in this framework are based on monthly drug acquisition cost and patient co-payments [7]. Comparisons between the ASCO-VF and the ESMO Magnitude of Clinical benefit Scale (ESMO-MCBS) frameworks have shown that there are substantial differences in estimated costs depending on whether drugs meet the criteria from the ESMO-MCBS or the ASCO-VF [8]. Therefore, this study aims to estimate the total drug treatment costs based on list prices and by using guidance from official approval and reimbursement documents. As this analysis is preliminary, no breast cancer subtypes (based on stage, receptor, and/or metastatic status) were distinguished.

METHODS

The EMA is the first agency to give market approval for newly developed drugs. Individual countries can decide themselves whether or not these drugs will be approved and/or reimbursed based on the evaluation of the cost-effectiveness of the new drug. Cost-effectiveness thresholds and approval and/or reimbursement mechanisms differ between countries [5, 6]. This study focusses on The Netherlands (NL), as this is one of the countries which has an open-access, reliable database containing list prices for all drugs, hosted by the Dutch Administrative Health Authority (ZINL). To extract all information that is necessary to calculate the average drug treatment cost for all approved breast cancer drugs in the NL, information from multiple databases was aggregated. Information was extracted using a stepwise approach of four steps.

First, an overview of all currently approved breast cancer drugs or therapies in the NL was created. Documents of ZINL were investigated for all drugs or therapies that are approved for use in any part of the whole care pathway in breast cancer treatment.

Second, data on the approval and costs of all drugs was retrieved via ZINL. As ZINL frequently updates their cost data, the cost data used were last extracted on 1 June 2017 [9]. For some drugs, different forms and strengths of the active drug substance exist. For each of these forms of the particular drug, the list price of the high-volume, high-dose package was used in the analysis as it was assumed that total cost for a particular quantity of a drug would then be less costly compared to using low-volume, low-dose packages. High-volume, high-dose packages were chosen because

producing, packaging and shipping larger (aggregated) products is generally cheaper than doing the same for multiple smaller products when the overall quantity remains fixed. For each of the forms of these drugs (e.g. 100 mL infusion fluid of 2 mg/mL), the lowest and highest list prices are presented by ZINL [9]. It was assumed that hospitals buying these drugs would pick the lowest-price option when multiple drug manufacturers produce similar drugs (same dose, same form, other manufacturer and/or brand name). Therefore, use of the currently available lower list price was deemed the most realistic approach. Furthermore, due to collaborations between hospitals, potential purchase benefits may arise resulting in lower purchase prices paid for particular drugs [9].

Third, further evidence on the effective dosage and average (if available) or median (if an average was lacking) duration of treatment was used in the calculations. This evidence was identified by using the European Public Assessment Reports (EPARs) when these were available for drugs that are approved in the NL. EPARs are the full scientific assessment reports that are published for every medicine for which authorisation was applied at the EMA. When no EPAR was available concerning the particular drug, the official reports of the national health regulating agency ZINL were used. When both, EPARs and ZINL reports, did not present any information on the average or median duration of treatment with that particular drug, corresponding trials were searched for estimates of progression-free survival (PFS) or disease-free survival (DFS). When these were available, the PFS or DFS were used in the further estimation of treatment costs, as for most, drugs treatment continues until progression is reported. When EPARs, ZINL assessment reports or the corresponding trials did not provide any of this information, the treatment duration for that particular drug was set equal to the average treatment length of all other drugs for which this information was available.

Fourth, all information was accumulated in the estimation of the drug treatment costs for all of the breast cancer drugs that are approved in the NL. The price per milligram of the specific form of the drug was calculated from the list prices per package of the drug. Following, dosing information was used to calculate the total number of milligrams of the drug necessary per treatment cycle. Dosing information in the EPARs and ZINL documents is presented in milligram per square meter or per kilogram of the patient. The average Dutch female has been used as a reference here, with a body surface of 1.6 square meters and a weight of 71.3 kilograms. This dosage per treatment cycle, corresponding list prices and the average length of treatment were used to determine the total estimated drug treatment cost.

RESULTS

According to ZINL, 31 drugs are approved for breast cancer treatment in the NL. In Table 1, the list prices for a single package of the specific drug packages are presented for the 15 highest-cost approved breast cancer drugs in the NL. Several drugs are available in multiple forms and concentrations of that particular drug. In Table 1, the lowest list prices are presented for one package of the highest volume and highest dose substance available for each specific form of every drug. Besides, the reimbursement status of that particular form and strength of the drug was reported. A further extension of Table 1 presenting all approved breast cancer drugs in all available forms and strengths of the drug substance and their reimbursement status is presented in Appendix 1. Appendix 1 shows that only one of all approved breast cancer drugs is not reimbursed in any form or strength of the drug substance, which is pamidronate disodium. For doxorubicin and Bevacizumab, only one of the forms or strengths of the drug substance is not reimbursed. Pertuzumab, Trastuzumab Emtansine and Trastuzumab are the three breast cancer drugs with the highest list prices.

Table 1. Average list prices and reimbursement for the 15 highest-cost breast cancer drugs.

List price rank	Active Substance	Form and strength of drug substance	Reim-bursed	List Price
1	Pertuzumab	Concentrate for solution infusion 30mg/ml 14 ml	Yes	€ 3,030.12
2	Trastuzumab emtansine	Powder for solution infusion 160 mg	Yes	€ 2,984.96
3	Trastuzumab	Injection fluid 120mg/ml 5 ml	Yes	€ 1,751.19
4	Paclitaxel	Concentrate for solution infusion 6 mg/ml 100 ml	Yes	€ 1,584.70
5	Bevacizumab	Concentrate for solution infusion 25mg/ml 16 ml	Yes	€ 1,310.65
6	Doxorubicin	Concentrate for solution infusion 2mg/ml 25 ml	Yes	€ 1,079.75
7	Thiotepa	Powder for solution infusion 100 mg	Yes	€ 1,028.20
8	Docetaxel	Concentrate for solution infusion 10mg/ml 16 ml	Yes	€ 889.68
9	Docetaxel	Concentrate for solution infusion 20mg/ml 8 ml	Yes	€ 877.47
10	Doxorubicin	Powder for solution infusion 50 mg	No	€ 689.49
11	Eribulin	Injection fluid 0,44mg/ml 3 ml	Yes	€ 614.32
12	Trastuzumab	Powder for solution infusion 150 mg	Yes	€ 589.99
13	Epirubicin	Injection fluid 2mg/ml 100 ml	Yes	€ 362.45
14	Paclitaxel	Powder for solution infusion 100 mg	Yes	€ 344.44
15	Goserelin Acetate	Implantation stick 10.8 mg	Yes	€ 341.56

If for a particular form of a drug, multiple list price were presented by ZINL (n = 9), the highest list prices were on average 86.3 % higher than the lowest prices. However, this high percentage was caused by two outliers (ibandronic acid—difference of 717.2 % and capecitabin—difference of 47.0%).

Table 2: Total estimated drug treatment cost and rankings

List price rank	Estimated treatment cost rank	Active Substance	Form and strength of drug substance	Number of cycles (3 weeks)	Treatment duration	Estimated treatment cost
31	1	Bevacizumab	Injection fluid 25mg/ml 0.15 ml	15.1	Mean TL	€ 174,399.68
5	2	Bevacizumab	Concentrate for solution infusion 25mg/ml 16 ml	15.1	Mean TL	€ 51,836.10
2	3	Trastuzumab emtansine	Powder for solution infusion 160 mg	9.5	Median TL	€ 44,505.75
28	4	Everolimus	Dispersible tablet 5 mg	8.0	Mean TL	€ 31,207.25
12	5	Trastuzumab	Powder for solution infusion 150 mg	13.0	Mean TL	€ 21,947.45
25	6	Everolimus	Tablet 10 mg	8.0	Mean TL	€ 21,254.31
36	7	Lapatinib	Tablet 250 mg	15.0	Median TL	€ 18,682.30
1	8	Pertuzumab	Concentrate for solution infusion 30mg/ml 14 ml	4.5	Mean TL	€ 16,665.64
3	9	Trastuzumab	Injection fluid 120mg/ml 5 ml	13.0	Mean TL	€ 16,286.10
7	10	Thiotepa*	Powder for solution infusion 100 mg	7.6	Assumption	€ 14,467.30
32	11	Vinblastin*	Injection fluid 1mg/ml 10 ml	7.6	Assumption	€ 10,175.00
6	12	Doxorubicin*	Concentrate for solution infusion 2mg/ml 25 ml	4.3	Median PFS	€ 9,950.96
11	13	Eribulin*	Injection fluid 0,44mg/ml 3 ml	4.8	Mean TL	€ 8,551.01
14	14	Paclitaxel*	Powder for solution infusion 100 mg	5.6	Mean TL	€ 8,024.00
10	15	Doxorubicin*	Powder for solution infusion 50 mg	4.3	Median PFS	€ 6,354.32

* Generic drugs

If an assumption was used to estimate the average treatment length, this assumption was based on the average treatment length of all other drugs for which a treatment length was available in the EPARs or ZINL assessment reports. Mean TL = Mean Treatment Length. Median PFS = Median Progression Free Survival.

The median of the difference between the highest cost and the lowest cost was only 0.5%. To convert these list prices to the estimated average treatment cost, evidence on the effective dosage and length of treatment was used.

In Table 2, the estimated average treatment costs are presented and ranked for the 15 highest-cost drugs. In the column for treatment duration is shown which type of treatment length (mean/median actual treatment length or mean/median PFS) for each drug has been identified from the EPAR or ZINL assessment reports. For thiotepa and vinblastin, an assumption on treatment length was made as there was no information on treatment length available within these reports for these drugs. These assumptions were based on the average length of treatment for all other drugs for which the treatment lengths were available in EPARs or ZINL assessment reports ($n = 31$, 73.8%), and was calculated to be 7.6 cycles or 22.8 weeks.

Furthermore, Table 2 shows that the highest factorial differences between list prices and estimated treatment cost are found for Bevacizumab (list price: €541.08—treatment cost: €174,399.68—factorial difference: 322.3 times more costly), Lapatinib (list price €158.37—treatment cost €18,682.30—factorial difference: 118.0 times more costly) and everolimus (list price €1,168.65—treatment cost €31,207.25—factorial difference: 26.7 times more costly). The full overview of the ranking and prices of all drugs is presented in Appendix 1. Table A1 in Appendix 1 shows that similar differences can be found for all approved breast cancer drugs according to their estimated treatment cost.

Absolute differences between low-cost targeted drugs (which are all drugs in rank 1–9) and the high-cost generic drugs (which are all drugs in rank 10–15) are relatively small. Of these generic drugs, doxorubicin, paclitaxel, vinblastine and thiotepa are already off-patent. The difference in estimated drug treatment cost between the relatively low-cost targeted drug Trastuzumab (rank 9) and the relatively high-cost generic drug thiotepa (rank 10) is €1,818.81.

In the estimation of the estimated drug treatment cost, list prices of high-volume, high-dose packages were used as these were assumed to be less costly than low-volume, low-dose packages. A comparison of the estimated drug treatment cost using high-volume, high-dose packages with the estimated drug treatment cost using low-dose, low-volume packages is presented in Appendix 2.

DISCUSSION

This study aimed to give an estimation of the total drug treatment cost in breast cancer based on list prices and official approval and reimbursement documents. However, there are still several difficulties encountered in the process of estimating drug treatment costs. The first difficulty concerns the identification of the average length of treatment from EPARs or ZINL assessment reports. For some drugs, the average or median length of treatment is not reported standardised in these reports. By extending the search to the clinical trials reports on which the EPAR and ZINL assessment reports were based, for most drugs, the duration of treatment could be defined based on the PFS or DFS.

However, there are still some drugs for which this evidence is lacking in these reports or the trials on which these reports are based. The second difficulty concerned the lack of evidence stating which particular form or strength of the drugs is most likely to be used in clinical practice. For multiple drugs, there are several forms and strengths available. In this study, high-volume, high-dose packages were used to calculate the drug treatment cost. However, there are several options in combining different package sizes to different doses. This prohibits accurate cost estimations as well as estimations on the amount of waste arising when packages are not fully used

As an increasing number of cancer drugs keep being developed and the cost of cancer care continues to increase, it is important to take the estimated total treatment costs of new drugs into account from early development stages onwards. One of the most important means to manage cancer drug prices and to further ensure the affordability of cancer care, is the targeting of those patients who are most likely to respond to particular treatments. One of the ways to identify those patients is the use of biomarkers, which are biological characteristics that can be identified from tissue or body fluids. Biomarkers enable the identification of patients who are highly likely or unlikely to benefit from a particular treatment or can be used as response markers to establish treatment effectiveness very rapidly after treatment initiation. Several studies have been focusing on the use of liquid biopsies, which are tests for biomarkers traceable in body fluids such as blood and urine, to be able to further target and personalise treatment [10]. One of the advantages of using these liquid biopsies is that tests can be cheaper, less invasive, and might even be more sensitive in targeting the treatment population than expensive imaging diagnostics [11–14].

However, as new drug development and cost-effectiveness studies (focusing on the personalised prescription of high-cost drugs based on liquid biopsies) should

refer to the prices that we should be able to pay currently, it is important that we are able to create an overview of the actual treatment prices that are currently being reimbursed. As it is hard to find evidence on especially the average length of treatment, it would be valuable to further extend standardisation of reporting this evidence in national or international drug assessment reports. An extended standardisation in which all treatment measures (such as the treatment duration, dosage and forms and strengths of the active substance on which the results are based) are presented in the same format in the same sections of these reports will improve the availability and transparency of such evidence, and thereby also the accuracy of cost estimations. Even though these standardised measures are currently unavailable for such cost estimations, the results of this review have already shown that there are substantial differences in the ranking of costliness of drugs according to their list prices or their estimated treatment costs. However, even though drugs are costly and have substantial influence on the budget impact, when developing new drugs and referring to such frameworks, it is important to not solely focus on potential budget impact constraints. Cost-effectiveness, that is, the balance between the additional costs and the associated improvement in health outcomes should also be taken into account.

Furthermore, in such cost-effectiveness analyses, the (higher) full patient treatment costs should also be taken into account, as the drug treatment costs identified in this study do not include further patient treatment costs (e.g. the drug treatment costs do not include costs for differences in the administration such as hospitalisation or home medication of different treatments). Although only one drug was identified which was approved but not reimbursed in the NL, other countries have other reimbursement regulations and might, therefore, have different drugs (not) reimbursed [6, 15]. The cost calculations that were performed in this study provide a stepwise approach to establishing a reference framework of drug treatment costs (necessary for rational decision making) when there is limited open-access to cost information.

Even though list prices are not the prices that are actually paid, these are currently the only drug prices which are openly available. Furthermore, though these list prices exist for all drugs in most countries, in some countries these list prices are only openly available in a standardised way for generic drugs and not for the typically more costly targeted drugs. When more countries aim for a transparent costing system, this would enable faster, easier and a more realistic collection of drug treatment costs.

CONCLUSIONS

This study presented an initial estimation of drug treatment cost based on list prices. An important finding was that off-patent generics, such as paclitaxel and doxorubicin, are still relatively costly. Although most generics are substantially cheaper than (patented) targeted drugs, the estimated drug treatment costs for the highest-cost generics were still higher than expected for off-patent drugs. Furthermore, it was shown that there are still several barriers which prohibit accurate estimations of drug treatment cost. List prices are available for all breast cancer drugs in the Netherlands. However, in other European countries this availability is less standardized. In addition, substantial differences can exist between the list prices of drugs and the associated average actual treatment costs which further complicates accurate estimations of drug treatment cost.

To improve on cost estimations and increase the relevance of such estimates for research on the development of new drugs, it is important to further extend standardization and increase European and national regulations on the presentation of treatment or drug costs, dosage and further treatment specifications. Furthermore, such calculations give average estimates of probable drug treatment costs and would improve if real world data can be collected.

DISCLOSURES

Abbreviations

ASCO	American Society of Clinical Oncology
ASCO-VF	ASCO Value Framework
DFS	Disease Free Survival
EPAR	European Public Assessment Report
EMA	European Medicines Authority
ESMO	European Society of Medical Oncology
ESMO-MCBS	ESMO Magnitude of Clinical Benefit Scale
FDA	Food and Drug Administration
IHC	Immunohistochemistry
ISH	In Situ Hybridization
MBC	Metastatic Breast Cancer
NL	the Netherlands
PFS	Progression Free Survival
TL	Treatment Length
TTP	Time to progression
ZINL	Dutch Administrative Healthcare Authority (Zorginstituut Nederland)

Keywords: breast cancer; drug; treatment; cost; medication; health economics

Conflicts of interests: The authors declare no conflict of interest

Financial conflicts of interest: This study was not funded

Disclosure of results at meeting: Preliminary results of this study were presented as a poster during ESMO 2017, Madrid.

Institutional review: All authors have read and approved the final manuscript

Acknowledgements: Not applicable

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016;66(1):7-30.
2. Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis JP. Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *J Natl Cancer Inst.* 2008;100(24):1780-91.
3. Samuel N, Verma S. Cross-comparison of cancer drug approvals at three international regulatory agencies. *Curr Oncol.* 2016;23(5):e454-e60.
4. Jakka S, Rossbach M. An economic perspective on personalized medicine.
5. van Harten WH, Wind A, de Paoli P, Saghatchian M, Oberst S. Actual costs of cancer drugs in 15 European countries. *The Lancet Oncology.* 2016;17(1):18-20.
6. van Harten W, MJ IJzerman. Responsible pricing in value-based assessment of cancer drugs: real-world data are an inevitable addition to select meaningful new cancer treatments. *Ecancermedalscience.* 2017;11:ed71
7. Schnipper LE, Davidson NE, Wollins DS, Blayney DW, Dicker AP, Ganz PA, et al. Updating the American Society of Clinical Oncology Value Framework: Revisions and Reflections in Response to Comments Received. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2016;34(24):2925-34.
8. Del Paggio JC, Sullivan R, Schrag D, Hopman WM, Azariah B, Pramesh CS, et al. Delivery of meaningful cancer care: a retrospective cohort study assessing cost and benefit with the ASCO and ESMO frameworks. *The Lancet Oncology.* 2017;18(7):887-94.
9. Zorginstituut Nederland. www.zorginstituutnederland.nl, 2017.
10. Matthew EM, Zhou L, Yang Z, Dicker DT, Holder SL, Lim B, et al. A multiplexed marker-based algorithm for diagnosis of carcinoma of unknown primary using circulating tumor cells. *Oncotarget.* 2016;7(4):3662-76.
11. Berghuis AM, Koffijberg H, Prakash J, Terstappen LW, MJ U. Detecting Blood-Based Biomarkers in Metastatic Breast Cancer: A Systematic Review of Their Current Status and Clinical Utility. *International journal of molecular sciences.* 2017;18(2).
12. Giuliano M, Giordano A, Jackson S, Hess KR, De Giorgi U, Mego M, et al. Circulating tumor cells as prognostic and predictive markers in metastatic breast cancer patients receiving first-line systemic treatment. *Breast cancer research : BCR.* 2011;13(3):R67.
13. Cristofanilli M, Broglio KR, Guarneri V, Jackson S, Fritsche HA, Islam R, et al. Circulating tumor cells in metastatic breast cancer: biologic staging beyond tumor burden. *Clinical breast cancer.* 2007;7(6):471-9.
14. Ravelli A, Reuben JM, Lanza F, Anfossi S, Cappelletti MR, Zanotti L, et al. Breast cancer circulating biomarkers: advantages, drawbacks, and new insights. *Tumour Biol.* 2015;36(9):6653-65.
15. Cheema PK, Gavura S, Migus M, Godman B, Yeung L, Trudeau ME. International variability in the reimbursement of cancer drugs by publically funded drug programs. *Curr Oncol.* 2012;19(3):e165-

CHAPTER 3

REAL WORLD DATA ON DISCORDANCE BETWEEN ESTROGEN, PROGESTERONE AND HER2 RECEPTOR EXPRESSION ON DIAGNOSTIC TUMOR BIOPSY VERSUS TUMOR RESECTION MATERIAL

A.M.S. BERGHUIS, C.H.M. VAN DEURZEN, H. KOFFIJBERG,
L.W.M.M. TERSTAPPEN, S. SLEIJFER, M.J. IJZERMAN

THIS CHAPTER WAS PUBLISHED AS:

A.M.S. BERGHUIS, C.H.M. VAN DEURZEN, H. KOFFIJBERG, L.W.M.M. TERSTAPPEN, S. SLEIJFER, M.J. IJZERMAN. REAL WORLD DATA ON DISCORDANCE BETWEEN ESTROGEN, PROGESTERONE AND HER2 RECEPTOR EXPRESSION ON DIAGNOSTIC TUMOR BIOPSY VERSUS TUMOR RESECTION MATERIAL. BREAST CANCER RESEARCH AND TREATMENT, 13-02-2019. DOI: 10.1007/S10549-019-05141-Y

ABSTRACT

PURPOSE

The estrogen (ER), progesterone (PR) and HER2 status are essential in the guiding treatment decisions in breast cancer patients. In daily life, the ER/PR/HER2 status is expected to be commonly tested twice, i.e. at diagnosis using material from tumor needle biopsies, and after tumor resection using full tumor tissue material. This study explored the discordance of ER/PR/HER2 between tumor needle biopsies and full tumor resection material using real world patient-level data from Dutch breast cancer patients.

METHODS

Pathology reports of 11,054 breast cancer patients were derived from PALGA (Dutch Pathology Registry). Discordance was calculated for multiple combinations of the ER/PR/HER2 receptor status. The influence of patient- and tumor characteristics on the probability of having discordant test results was analyzed using multiple logistic regression models (separately for ER, PR and HER2).

RESULTS

For 1,279 patients (14.4%), at least one of the receptors (ER/PR/HER2) was determined on both, biopsy and tumor tissue material. The majority had concordant test results for ER (n=916;94.8%), PR (n=1170;86.7%) and HER2 (n=881;98.1%). Patients having an ER and HER2 positive but PR negative biopsy classification, BR grade III and <10% tumor tissue remaining after neo adjuvant therapy (NAT), have the highest probability of ER discordant test results (OR=4.991;p=83.31%). The probability of discordance in PR is based on different sets of patient- and tumor characteristics. Potential cost savings from omitting multiple tests if concordance can be perfectly predicted can be up to €196,000 yearly.

CONCLUSIONS

Double testing of ER/PR/HER2 is less common than expected. Discordance in ER/PR/HER2 test results between tumor needle biopsy taken at the time of diagnosis and tumor resection material is very low, especially in patients not receiving any form of neo adjuvant therapy. Following, these results imply that a substantial number of tests can potentially be omitted in specific subgroups of breast cancer patients.

BACKGROUND

Breast cancer is the most frequently diagnosed form of cancer among women with yearly incidence rates of almost 15,000 women in the Netherlands [38]. Of all women diagnosed with breast cancer, 90% presents with primary breast cancer without distant metastases [39]. In patients with primary breast cancer, the risk of relapse is determined based on factors such as tumor size, tumor grade, and lymph node involvement, and if considered high, patients are candidates for peri-operative systemic treatment aiming to reduce the relapse risk. The type of systemic treatment prescribed, is currently mainly dependent on the determination of the ER/PR and HER2 status of the tumor [40]. The ER/PR/HER2 status can be determined on tumor needle biopsy material or on tumor resection specimen using immunohistochemistry (IHC). For the HER2 status, additional in-situ hybridization (ISH) test is recommended to confirm the HER2 status in case IHC results are equivocal [9, 16].

In daily life, the ER/PR/HER2 status is determined multiple times in many breast cancer patients; first on specimen derived from a needle biopsy taken for initial diagnosis, and second on the whole tumor specimen obtained by tumor resection [9]. Usually, the receptor status results of the tumor needle biopsy and the tumor resection are expected to be concordant. In patients not receiving any form of neo-adjuvant therapy (NAT) between taking the needle biopsy and resection, small series indeed strongly suggest that the concordance between both measurements (biopsy and resection) is relatively high (range 90.8%-97.5%) [41, 42]. However, discordance in test results might arise since tumor characteristics can change over time, in particular in patients treated with NAT, or because of sampling or analytical errors [43-45].

Given the importance of assessing the ER/PR/HER2 status for treatment decision making, it is essential to get more insight into the factors that may underlie discordant test results. Apart from consequences for treatment decision making, this is also important to improve the cost-effectiveness of the diagnostic pathway. For those patients in whom it is highly unlikely that the second test would yield a discordant test result or a change in clinical management, one of the ER/PR/HER2 determinations can potentially be omitted. Although the costs of the ER/PR/HER2 tests are relatively low, around €100 for an IHC and between €300 - 400 for a ISH [46, 47], the cumulative costs of the use of such tests can still be high given the large number of patients yearly diagnosed with breast cancer.

Several studies have reported on the discordance in ER/PR/HER2 between tumor needle biopsies and tumor resection. However, most (recent) studies reported on relatively small sample sizes in total, small subgroups of breast cancer patients, or included only patients diagnosed at a single hospital [42, 48-50]. In the study presented here, we have evaluated the discordance in ER/PR/HER2 between tumor needle biopsy and tumor resection material in the majority of invasive breast cancer patients diagnosed in 2016 and 2017 in the Netherlands. In addition, the influence of several tumor- and patient characteristics on the probability of discordance in either of the receptors were addressed. Furthermore, potential cost savings due to eliminating over testing in patients with concordant test results, was estimated.

METHODS

DATA SOURCES AND DESCRIPTION

Data on the ER/PR/HER2 status of invasive breast cancer patients was requested from the Dutch Pathology Registry (PALGA), which archives all pathology reports [51]. Since 2009, reporting modules are available for creating those pathological reports. Within these reports, information on patient-, tumor- and test characteristics is captured in numerous variables instead of in free text fields, which improves the possibility of analyzing high numbers of reports simultaneously.

Pathology reports of invasive breast cancer patients diagnosed between January 12, 2016 and January 1, 2018 were extracted from PALGA, as on January 12, 2016 a new synoptic reporting module for breast cancer biopsies became available that enabled saving more data in a standardized way.

Performing more than two tests, thereby creating additional excerpt records, can have multiple underlying reasons which are not always well documented within the pathology reports. Therefore, patients having more than two excerpt records were removed from further analysis. Discordance, by definition, can only exist between multiple measurements. Consequently, patients for which only one excerpt record was available were also removed from further analysis.

DISCORDANCE IN TEST RESULTS

For each excerpt for which the ER and PR status were tested, the percentage of tumor cells that stained positive for ER and PR and the final classification of the excerpt (positive or negative) were registered. According to the Dutch breast cancer guideline [9], excerpts with 10% or more cells stained positive for ER or PR, are classified as

being positive. ER or PR classifications that deviate from this protocol, i.e. excerpts that had less than 10% of cells staining positive for ER or PR that were classified as being positive, were reclassified according to this 10% threshold.

Tumor needle biopsies will be referred to as biopsies in this study, whereas material derived from surgically removed tumor tissue will be referred to as the tumor resections. Test results between the biopsy and tumor resection were considered discordant, when the final corrected classification per marker, i.e. positive or negative, was different between both excerpts. Discordance was calculated for ER, PR, and HER2 separately, and for ER and PR combined.

LOGISTIC REGRESSION ANALYSES

To estimate whether discordance was more likely to occur in particular subgroups of patients, three independent logistic regression analyses were performed for ER, PR and HER2. These analyses were performed using the glm function from the stats package (version 3.5.1) in R (version 3.5.1). Discordance in either of the receptors was categorized as a binomial variable and was assigned to be the dependent variable, whereas the type of NAT (i.e. hormonal- chemotherapy, or no neo adjuvant therapy), response to therapy, tumor subtype, Bloom Richardson (BR) grade, TNM stage, ER biopsy classification, percentage of cells positive for ER on biopsy material, PR biopsy classification, percentage of cells positive for PR on biopsy material, HER2 biopsy classification, HER2 IHC result on biopsy material and HER2 ISH result on biopsy material were assigned as independent variables in the initial model.

In each of the three analyses, a subset of the data was created which contained records of those patients for whom the particular receptor under investigation was tested twice. The logistic regression models were evaluated using the Akaike Information Criterion (AIC) in a stepwise algorithm in which both, backward and forward selection of variables, was combined. The stepwise approach was performed using the step function from the stats package (version 3.5.1) in R (version 3.5.1) and was started with a null model, in which no variables were manually added to the initial model. The approach was continued with additional steps in which variables were either added forwards or eliminated backwards in each step. The model that was derived when either eliminating or adding a particular variable to the model did not result in a lower AIC of the model, was considered to be the model with the best performance.

Before running the logistic regression analyses and to overcome implications of too small patient subgroups, several variables were grouped to maintain a sufficient number of patients in each subgroup. Tumor subtypes were merged into three main categories: ductal carcinoma, lobular carcinoma, and other tumor subtypes. For this reason, also the type of neo-adjuvant therapy was reclassified into 4 categories: chemotherapy, hormonal therapy, other or combined therapy and no therapy. The primary tumor (T), regional lymph nodes (N) and distant metastases (M) stage (TNM) was defined according to the eight edition of the TNM staging system [52]. In this staging system, a N1mi classification can only exist in combination with T1. The N-stage of patients for whom a higher T-stage was recorded in combination with this N1mi, was assumed to be a N1-N-stage.

POTENTIAL COST SAVINGS OF OMITTING MULTIPLE TESTS

Potential cost savings from avoiding unnecessary tests (i.e. with outcomes similar to the initial test) were calculated. Based on the total number of concordant test results in either ER/PR/HER2, or combinations of those, the number of tests that potentially could have been omitted was calculated. Costs for each test, or combinations of tests, were derived from the Dutch Healthcare Authority (NZa). The ER and PR status are generally tested by using an IHC. In the calculation of the potential cost savings, it was assumed that the costs for testing both, ER and PR on the same excerpt, were equal to those of a single test for either ER or PR. If HER2 was additionally tested, additional costs of the HER2 antibody were added to the costs of performing an IHC, whereas IHC in combination with either ER, PR or both was assumed to be equally expensive to performing an IHC for HER2 singularly.

RESULTS

Patient level data derived from PALGA included 11,054 patients with invasive breast cancer whom were diagnosed (i.e. had a biopsy) after January 12, 2016. Patients whom had either one excerpt record (n=92; 0.66%) or more than two excerpt records (n=2,081; 15.00%) were excluded from further analysis. The final dataset used for further analyses consisted of 8,881 unique patients, whom all had one biopsy and one tumor resection record. An overview of all patient characteristics is provided in table 1.

Table 1: Summarized overview of clinically relevant patient characteristics

Category	Sub-categories	Total number of patients	Therapy			
			No NAT	Chemotherapy	Hormonal therapy	Other therapy
Bloom Richardson	Unknown	618 (6.96 %)	112 (18.12 %)	437 (70.71 %)	47 (7.61 %)	22 (3.56 %)
Grade	Grade I	2464 (27.74 %)	2324 (94.32 %)	105 (4.26 %)	34 (1.38 %)	1 (0.04 %)
	Grade II	4086 (46.01 %)	3827 (93.66 %)	205 (5.02 %)	49 (1.2 %)	5 (0.12 %)
	Grade III	1713 (19.29 %)	1595 (93.11 %)	105 (6.13 %)	11 (0.64 %)	2 (0.12 %)
Response to therapy	No response	7984 (89.9 %)	7858 (98.42 %)	86 (1.08 %)	35 (0.44 %)	5 (0.06 %)
	< 10 % tumor remaining	232 (2.61 %)	0 (0.00 %)	216 (93.10 %)	6 (2.59 %)	10 (4.31 %)
	10-50 % tumor remaining	325 (3.66 %)	0 (0.00 %)	281 (86.46 %)	36 (11.08 %)	8 (2.46 %)
	> 50 % tumor remaining	340 (3.83 %)	0 (0.00 %)	269 (79.12 %)	64 (18.82 %)	7 (2.06 %)
TNM Stage	0	506 (5.70 %)	476 (94.07 %)	25 (4.94 %)	4 (0.79 %)	1 (0.20 %)
	IA	688 (7.75 %)	673 (97.82 %)	13 (1.89 %)	2 (0.29 %)	0 (0.00 %)
	IB	4226 (47.58 %)	3862 (91.39 %)	317 (7.50 %)	37 (0.88 %)	10 (0.24 %)
	IIA	330 (3.72 %)	299 (90.61 %)	26 (7.88 %)	3 (0.91 %)	2 (0.61 %)
	IIB	1967 (22.15 %)	1680 (85.41 %)	229 (11.64 %)	50 (2.54 %)	8 (0.41 %)
	IIIA	739 (8.32 %)	589 (79.70 %)	119 (16.10 %)	27 (3.65 %)	4 (0.54 %)
	IIIB	264 (2.97 %)	158 (59.85 %)	89 (33.71 %)	15 (5.68 %)	2 (0.76 %)
	IIIC	63 (0.71 %)	43 (68.25 %)	15 (23.81 %)	2 (3.17 %)	3 (4.76 %)
Tumor type	Ductal	98 (1.10 %)	78 (79.59 %)	19 (19.39 %)	1 (1.02 %)	0 (0.00 %)
	Lobular	6999 (78.81 %)	6166 (88.10 %)	715 (10.22 %)	91 (1.30 %)	27 (0.39 %)
	Other	1212 (13.65 %)	1067 (88.04 %)	101 (8.33 %)	41 (3.38 %)	3 (0.25 %)
	Total	8881 (100 %)	7858 (88.48 %)	852 (9.59 %)	141 (1.59 %)	30 (0.34 %)

DISCORDANCE

To calculate discordance, only those patients were selected for whom the ER, PR or HER2 was determined on both, biopsy and tumor resection material. For 1,279 patients, either ER, PR or HER2 was determined on both excerpts (14.40%). For those patients who only had 1 ER, PR or HER2 determination, the majority of patients had their status determined on resection material (ER n=4,343; 57.13%, PR n=4,363; 57.39%, HER2 n=4,772; 62.78%). Discordance in each of the receptors for patients with ER, PR and HER2 determined on both excerpts, is shown in table 2.

Table 2: Discordance in test results between biopsy and tumor resection material.

Receptor	Population size*	Discordant	Biopsy positive, resection negative	Biopsy negative, resection positive
ER	684 (7.7 %)	31 (4.53 %)	20 (64.52 %)	11 (35.48 %)
PR	890 (10.02 %)	135 (15.17 %)	67 (49.63 %)	68 (50.37 %)
HER2	707 (7.96 %)	6 (0.85 %)	0 (0.00 %)	6 (100 %)

* The population size described shows the number of patients for whom each of the specific receptors was determined on both, biopsy and tumor resection material.

In a substantial number of patients (n=590; 47.20%), the ER and PR receptor were both tested on biopsy and tumor resection material. Discordance in both receptors was found in a small number of patients (n=7; 1.19%).

A substantial percentage of all patients for whom either the ER, PR or HER2 status was determined on both, biopsy and tumor resection material, received no NAT (n=1,070; 85.60%), which is slightly lower than the relative number of patients not receiving any NAT in the full dataset (n=7,858; 88.48%). The percentage of patients having discordance in ER or PR test results, is higher in patients receiving either neo adjuvant chemotherapy or hormonal therapy, as shown in table 3.

Table 3: Discordance in ER, PR or HER2 receptor status and neo adjuvant therapy

	Chemotherapy		Hormonal therapy		No NAT	
	Population	Discordant	Population	Discordant	Population	Discordant
ER	96 (11.27%)	9 (9.38%)	17 (12.06%)	1 (5.88%)	568 (7.23%)	21 (3.7%)
PR	110 (12.91%)	21 (19.09%)	30 (21.28%)	12 (40%)	745 (9.48%)	99 (13.29%)
HER2	51 (5.99%)	3 (5.88%)	9 (6.38%)	0 (0%)	645 (8.21%)	3 (0.47%)

* Row sums of the subpopulations that were tested multiple times for each receptor do not add up to population size as presented in column 2, table 2. Some patients (range 2-5 for the different receptors) received another therapy than chemotherapy or hormonal therapy. None of these patients had discordant test results for any of the receptors.

Of those patients with discordance in both, the ER and the PR receptor (n=7; 1.19%), 1 patient received chemotherapy, whereas all others received no neo adjuvant therapy.

INFLUENCE ON DISCORDANCE OF TUMOR- AND PATIENT CHARACTERISTICS

Logistic regression models for ER, PR and HER2 discordance separately were established using the stepwise algorithm. For HER2 discordance, however, the model could not be established properly, as only a few patients had discordance in this receptor ($n=6$), and was therefore not further evaluated. In the remaining regression analyses (ER and PR), the TNM stages were merged into stage 0-1, stage 2 and stage 3. Patients with different types of NAT, i.e. other than chemotherapy, hormonal therapy or no NAT, were excluded from this analysis ($n=30$; 7.67%).

3

In the regression analysis modelling ER discordance, the stepwise algorithm indicated that including ER and PR biopsy classification, the BR grade, the response to therapy (if any), and the HER2 IHC result of the biopsy results in a model with the lowest AIC. For PR discordance, the best performing model contains the type of neo adjuvant therapy (if any), the ER and PR biopsy classification, the TNM stage and the tumor type.

The derived odds ratios from both models, 95% confidence intervals (CIs), standard errors, and the corresponding p-values are shown in table 4.

The outcomes of the logistic regression analyses, as shown in table 4, show that the probability of discordance in the ER receptor is higher for patients with an ER positive biopsy classification than for patients with an ER negative biopsy classification ($OR=3.629$), whereas the probability of discordance in this ER receptor is lower for patients with a PR positive biopsy classification compared to patients with a PR negative biopsy classification ($OR=0.132$). The probability of discordance in the PR receptor, is higher for patients with an ER positive biopsy classification compared to patients with an ER negative biopsy classification ($OR=6.031$), whereas the probability of discordance in the PR receptor is lower for patients with a positive biopsy classification for HER2 derived by IHC compared to patients for whom no IHC was performed ($OR=0.356$)

Table 4: Results of the logistic regression analyses per receptor

	ER Discordance	Odds ratio	Std. Error	P-value	CI 2.5 %	CI 97.5 %
	Intercept*	0.050	2.382	0.001	0.008	0.254
ER biopsy classification (ref. biopsy negative, n=463)	Positive (n=463)	3.629	1.654	0.010	1.378	10.073
PR biopsy classification (ref. biopsy negative, n=267)	Positive (n=267)	0.132	1.709	0.000	0.043	0.361
BR Grade (ref. unknown BR grade, n=39)	Grade 1 (n=121)	0.990	2.262	0.990	0.199	5.202
	Grade 2 (n=265)	0.191	2.412	0.060	0.031	1.074
	Grade 3 (n=198)	1.482	2.100	0.596	0.367	7.006
ISH biopsy result (ref. no ISH performed, n=420)	Amplified (n=31)	1.399	2.013	0.631	0.293	4.939
	Not Amplified (n=172)	0.191	2.228	0.039	0.028	0.736
NAT (ref. no NAT, n=518)	Chemotherapy (n=90)	3.355	1.695	0.022	1.141	9.245
	Hormonal therapy (n=15)	3.230	3.171	0.310	0.157	22.473
	PR Discordance	Odds ratio	Std. Error	P-value	2.5 %	97.5 %
	Intercept	0.036	1.874	0.000	0.009	0.113
ER biopsy classification (ref. biopsy negative, n=142)	Positive (n=673)	6.031	1.589	0.000	2.609	16.505
TNM Stadium (ref. stadium 0 -1, n=458)	TNM stadium 2 (n=312)	1.910	1.250	0.004	1.236	2.965
	TNM stadium 3 (n=45)	0.973	1.622	0.955	0.353	2.396
NAT (ref. no NAT, n=683)	Chemotherapy (n=104)	2.188	1.361	0.011	1.177	3.959
	Hormonal therapy (n=28)	4.088	1.531	0.001	1.741	9.381
PR biopsy classification (ref. biopsy negative, n=340)	Positive (n=475)	0.500	1.261	0.003	0.317	0.790
IHC biopsy result (ref. no IHC performed, n=131)	IHC 2+ (n=101)	0.927	1.430	0.833	0.457	1.865
	IHC 0 (n=212)	0.479	1.390	0.025	0.250	0.913
	IHC 1+ (n=323)	0.560	1.349	0.053	0.313	1.015
	IHC 3+ (n=48)	0.356	1.833	0.089	0.094	1.065
Tumor type ** (ref. different type, n=73)	Ductal (n=618)	1.237	1.585	0.644	0.537	3.370
	Lobular (n=124)	2.146	1.645	0.125	0.855	6.192

* The intercept reflects the probability of discordance for the total reference category of patients. For the prediction of ER discordance, these patients have an ER negative but PR positive biopsy classification, an unknown BR grade, had no ISH performed on the biopsy material and did not receive any form of NAT.

** Both tumor types appear to have a non-significant influence compared to a different type of tumor. This is due to recoding this variable. Ductal and lobular are significantly different from each other.

When combining the different patient and tumor characteristics specific subgroups of patients can be determined with the highest and lowest probabilities of deriving discordant test results in either of the receptors. Patients with a BR Grade 3, a positive classification of ER on biopsy, a positive classification of HER2 on biopsy derived by an amplified ISH who did receive neo adjuvant chemotherapy, have the highest probability of an ER discordant test result (OR=1.274; p=56.01%), whereas patients with a BR Grade 2, a positive classification of PR on biopsy, a negative classification of HER2 on biopsy derived by a not amplified ISH who did not receive any NAT at all have the lowest probability of an ER discordant test result (OR= <0.005; p=0.02%). Patients with a lobular tumor type staged as TNM stage of 2A or 2B, a positive classification of ER, a negative classification of PR and no HER2 IHC performed on biopsy and who did receive neo adjuvant hormonal therapy, have the highest probability of a PR discordant test result (OR=3.627; p=78.39%), whereas patients with another tumor type than lobular or ductal, a TNM stage 3A, 3B or 3C, a positive classification of PR on biopsy, a positive classification of HER2 derived by a 3+ IHC result and who did not receive any form of NAT, had the lowest probability of a PR discordant test result (OR=0.006; p=0.62%).

POTENTIAL COST SAVINGS

For the majority of patients in whom multiple tests were performed, the test results were concordant, ranging from 84.83% in PR status to a concordance of 99.15% in the HER2 status. The cost of determining the ER and PR status together is €151 whereas additionally determining the HER2 on an IHC is €100 more expensive. The cost of an ISH are approximately €408. Determining the ER, PR and complete HER2 status (IHC and ISH) results in a cost of approximately €659.

Assuming (hypothetically) that test results can be perfectly predicted for patients with concordant test results, either one of the tests performed can potentially be omitted. In table 2, the total number of patients having discordant test results was presented. However, as tests can be performed simultaneously, it was additionally calculated how many patients had concordant test results in all receptor statuses (ER, PR and HER2). A total of 325 full ER/PR/HER2 tests can potentially be omitted, as all receptors were found to be concordant (approximately €214,000). In addition, also ER and PR are often tested simultaneously. A total of 179 tests can additionally be potentially omitted, as 179 patients had a combined ER/PR test and concordant test results for both receptors whilst not having had two HER2 determinations (approximately €27,000). Following, the number of patients with double tests for

ER, PR or HER2 singularly (double tests in only 1 receptor) were calculated, resulting in additional cost savings of approximately €169,000. In total, potential cost savings can sum up to approximately €410,000 when concordance in test results can be perfectly predicted, and second tests can therefore safely be omitted. Determining the ER/PR/HER2 status on both, biopsy and tumor resection material is currently already selectively performed (n=1279; 14.40%). However, cost savings for this particular patient group (from 2016 till 2018) may still result in yearly cost savings of approximately €205,000. Combined, these results indicate that an average cost saving of up to € 320 per patient can potentially be reached.

DISCUSSION

The majority of patients had two excerpt records in their pathology reports. These results are in accordance with what was expected for most breast cancer patients and in line with the Dutch breast cancer guideline, as usually a biopsy is done to confirm breast cancer after which the tumor is resected in invasive breast cancer. In general, for most patients (n=7,631; 85.9%) the ER, PR, and HER2 status were, however, only tested once on either biopsy or tumor resection material. Discordance in the ER/PR/HER2 receptor was shown to be limited, with percentages of patients with discordant test results in either one of the receptors ranging from 0.9-15.2%. When testing for both, the ER and PR receptor, discordance in test results was also found to be very low (1.19%). The percentages of patients with discordant test results were substantially higher in patients who received any form of NAT (i.e. chemotherapy or hormonal therapy) compared to those patients who did not receive any form of NAT.

Two separate multivariable stepwise logistic regression analyses were performed to calculate the probability of deriving discordant test results for the ER and PR receptor as a function of patient and tumor characteristics that are available directly after the biopsy. Even though the numbers of patients having discordant test results in specific subgroups are relatively low, the data suggest that several of these variables significantly influence the probability of deriving a discordant test result. However, before predicting discordance using such models, these should be validated, which is often complex or even impossible in datasets in which a minority of observations is used to predict the outcome. In this study, small percentages of patients had a discordant test result, especially for ER and even more for HER2, whereas this percentage was slightly higher for PR. Consequently, validating the model for PR discordance may be easier, given the larger number of expected discordant test results. Although potentially feasible and valid in predicting discordance, singly

testing the PR receptor is clinically less relevant than single tests for the ER and PR receptors. Therefore, such a model would not serve its initial clinical aim, which is of course, predicting discordance in clinically meaningful receptors. However, the regression analyses do have provided valuable insight into the combination of patient- and tumor characteristics which appear to be influencing the probability of discordance. For example, the probability of discordance can even be less than 1% in specific subgroups of patients.

Based on the results of the logistic regression model and the low percentages of patients with discordant test results that were found in this study, performing multiple tests is probably unnecessary for the majority of patients. However, a number of significant risk factors for discordance were identified in this study, suggesting that testing twice is likely beneficial in very specific subgroups of patients. As a next step, discordance risks should be validated in future studies, focusing particularly on patients with high predicted discordance risks, to further tailor multiple ER/PR/HER2 testing.

Potential cost savings of omitting tests can potentially sum up to approximately €205,000 yearly. However, prices used in this calculation are list prices, which are not always reflecting the actual money paid for each test. In addition, potential cost savings were based on the assumption that discordance can be perfectly predicted, whereas it is hard to make perfect predictions in practice. Furthermore, to be able to estimate full potential cost impact, treatment decisions based on discordance in these test results and guideline adherence regarding the use of these tests should also be taken into account.

No adjuvant treatment decisions based on the discordance in this study were taken into account. Future studies should focus on the impact of discordant results on future treatment decisions. In addition, no survival results could yet be derived for this cohort of patients as these patients were diagnosed after January 12, 2016. Consequently, the follow-up period of this patient cohort is (yet) too short to draw firm conclusions regarding survival. Future studies should incorporate the impact on treatment decisions and associated survival outcomes in further predicting the probability of potential discordance. In addition, further treatment decisions related to discordance should be taken into account. Especially when treatment decisions will not differ, even though test results are discordant, additional testing may be omitted.

CONCLUSIONS

Testing for ER/PR/HER2 on both, tumor biopsy and tumor resection material, is not performed in the majority of patients. The discordance in test results of the ER/PR/HER2 status in patients with invasive breast cancer is limited. Of all three receptors, remarkably the PR receptor status is most frequently measured on both, biopsy and tumor resection material and was subsequently found to be discordant most frequently. In addition, the probability of a discordant test result is generally low, but can however, be significantly influenced by several patient- and tumor characteristics.

Future research investigating either discordance, their corresponding predictive models, or its potential consequences, should also take further treatment decisions and treatment outcomes into account for each of the receptors.

DECLARATIONS

Table of abbreviations

ASCO	American Society of Clinical Oncology
CAP	College of American Pathologists
ER	Estrogen
FISH	Fluorescent In Situ Hybridization
HER2	Human Epidermal Growth Factor 2
IHC	Immunohistochemistry
ISH	In Situ Hybridization
PALGA	Dutch Pathology Registry
PR	Progesterone

Acknowledgements: The authors like to thank PALGA for their commitment towards improving the analyses and providing valuable suggestions for the final manuscript.

Funding: This study was not funded

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable

Data availability: The datasets used and analyzed in this study are not publicly available but can be requested at PALGA.

Conflicts of interest: The authors declare that they have no conflict of interest.

Authors' contributions: The initial data transformations and analyses were performed by AMSB. Reclassification and clinically meaningful use of variables was discussed with CHMD, SS and LWMMT. Model algorithms and stepwise approaches were discussed with HK and MIJ. All authors were involved in writing and revising the manuscript.

Keywords: Discordance, breast cancer, estrogen, progesterone, human epidermal growth factor receptor 2, diagnostics

REFERENCES

1. Netherlands Cancer Registry (NCR) 2018. Nationwide breast cancer incidence.
2. Dawood S, Broglio K, Gonzalez-Angulo AM, Buzdar AU, Hortobagyi GN, Giordano SH (2008) Trends in survival over the past two decades among white and black patients with newly diagnosed stage IV breast cancer. *J Clin Oncol* 26 (30):4891-4898. doi:10.1200/JCO.2007.14.1168
3. Mueller C, Haymond A, Davis JB, Williams A, Espina V (2018) Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev Proteomics* 15 (2):131-152. doi:10.1080/14789450.2018.1421071
4. NABON (2012) Breast Cancer Guideline.
5. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel M (2015) 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* 26 (8):1533-1546. doi:10.1093/annonc/mdv221
6. Singh K, Wang Y, Marketkar S, Kalife ET, Steinhoff MM (2017) Comparison of estrogen receptor, progesterone receptor and HER2 results in concurrent ipsilateral samples with invasive breast carcinoma: a retrospective study of 246 biopsies from 119 patients. *Hum Pathol* 65:123-132. doi:10.1016/j.humpath.2017.04.016
7. Hadad SM, Jordan LB, Roy PG, Purdie CA, Iwamoto T, Pusztai L, Moulder-Thompson SL, Thompson AM (2016) A prospective comparison of ER, PR, Ki67 and gene expression in paired sequential core biopsies of primary, untreated breast cancer. *BMC Cancer* 16 (1):745. doi:10.1186/s12885-016-2788-x
8. Gahlaut R, Bennett A, Fatayer H, Dall BJ, Sharma N, Velikova G, Perren T, Dodwell D, Lansdown M, Shaaban AM (2016) Effect of neoadjuvant chemotherapy on breast cancer phenotype, ER/PR and HER2 expression - Implications for the practising oncologist. *Eur J Cancer* 60:40-48. doi:10.1016/j.ejca.2016.03.006
9. Jin X, Jiang Y-Z, Yu K-D, Shao Z-M, Di G-H (2015) Prognostic value of receptor conversion after neoadjuvant chemotherapy in breast cancer patients: a prospective observational study *Oncotarget* 6 (11):9600-9611
10. Yang L, Zhong X, Pu T, Qiu Y, Ye F, Bu H (2018) Clinical significance and prognostic value of receptor conversion in hormone receptor positive breast cancers after neoadjuvant chemotherapy. *World J Surg Oncol* 16 (1):51. doi:10.1186/s12957-018-1332-7
11. Carlson B (2008) HER2 tests: how do we choose?
12. Garrison LP, Jr., Lalla D, Brammer M, Babigumira JB, Wang B, Perez EA (2013) Assessing the potential cost-effectiveness of retesting IHC0, IHC1+, or FISH-negative early stage breast cancer patients for HER2 status. *Cancer* 119 (17):3113-3122. doi:10.1002/cncr.28196
13. Asogan AB, Hong GS, Arni Prabhakaran SK (2017) Concordance between core needle biopsy and surgical specimen for oestrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 status in breast cancer. *Singapore Med J* 58 (3):145-149. doi:10.11622/smedj.2016062
14. Ensani F, Omranipour R, Jahanzad I, Jafari A, Nafarzadeh S, Aminishakib P (2017) The Core Needle and Surgical Biopsy Concordance to Detect Estrogen, Progesterone, and Her-2 Receptors in Breast Cancer:

A Comparative Study. Iranian journal of pathology 12 (3):202-208

15. You K, Park S, Ryu JM, Kim I, Lee SK, Yu J, Kim SW, Nam SJ, Lee JE (2017) Comparison of Core Needle Biopsy and Surgical Specimens in Determining Intrinsic Biological Subtypes of Breast Cancer with Immunohistochemistry. Journal of breast cancer 20 (3):297-303. doi:10.4048/jbc.2017.20.3.297

16. Casparie M, Tiebosch ATMG, Burger G, Blauwgeers H, van de Pol A, van Krieken JHJM, Meijer GA (2007) Pathology databanking and biobanking in the Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cellular Oncology 29:19-24

17. Giuliano AE, Edge SB, Hortobagyi GN (2018) Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. Annals of surgical oncology 25 (7):1783-1785. doi:10.1245/s10434-018-6486-6

CHAPTER 4

TREATMENT DECISIONS OF ADJUVANT SYSTEMIC THERAPY FOLLOWING REPEATED ASSESSMENT OF ESTROGEN AND THE HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR 2 ON TUMOR BIOPSY AND RESECTION MATERIAL IN INVASIVE PRIMARY BREAST CANCER PATIENTS

A.M.S. BERGHUIS, C.H.M. VAN DEURZEN, H. KOFFIJBERG,
L.W.M.M. TERSTAPPEN, S. SLEIJFER, M.J. IJZERMAN

THIS CHAPTER WAS SUBMITTED AS:
A.M.S. BERGHUIS, C.H.M. VAN DEURZEN, H. KOFFIJBERG, L.W.M.M. TERSTAPPEN, S. SLEIJFER,
M.J. IJZERMAN. TREATMENT DECISIONS OF ADJUVANT SYSTEMIC THERAPY FOLLOWING REPEATED
ASSESSMENT OF ESTROGEN AND THE HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR 2 ON
TUMOR BIOPSY AND RESECTION MATERIAL IN INVASIVE PRIMARY BREAST CANCER PATIENTS.

ABSTRACT

PURPOSE

Presence of estrogen (ER), progesterone (PR) and HER2 on tumor tissue are used to guide treatment decisions in breast cancer patients. Discordance between these receptor statuses may arise between the tumor needle biopsy and the tumor resection. No previous studies evaluated whether treatment decisions are similar in patients with concordant or discordant test results. This study evaluates the adjuvant treatment consequences in primary breast cancer patients with discordant and patients with concordant test results, based on real world data.

METHODS

Pathology reports derived from the Dutch Pathology Registry (PALGA) and consecutive treatment information derived from the Netherlands Cancer Registry (NCR) were linked. Treatment patterns for different combinations of receptor statuses (ER and HER2) on biopsy and resection material were identified. To evaluate whether discordance and other clinically relevant variables influence the treatment prescribed, logistic multivariable regression models were developed.

RESULTS

ER was tested twice in 658 patients, HER2 in 681 patients. For patients with concordant test results, the majority received treatment that matched the treatment expectations (ER: $n=472$; 73.5%, HER2: $n=670$; 98.0%). Approximately half of the patients with discordant test results received treatment that matched the expectations (ER: $n=15$; 51.7%, HER2: $n=3$; 50.0%). Discordance in ER and HER2 test results significantly decreases the probability that patients receive treatment that is not matching the expectations, assuming that other patient and tumor characteristics remain the same (ER concordant OR=0.087, HER2 concordant OR=0.026).

CONCLUSIONS

This study demonstrated that primary breast cancer patients with concordant test results for ER and HER2 usually received the expected matching adjuvant treatment, while treatments are more diversified in patients with discordant test results. To identify potential underlying reasons causing this diversification and to evaluate their consequences, future research should include larger numbers of patients and take survival, quality of life and the costs of these treatments into account.

BACKGROUND

The most commonly diagnosed type of cancer in women is breast cancer, with an incidence of nearly 15,000 women yearly in the Netherlands [1]. The majority of breast cancers nowadays are diagnosed at an early stage, when no distant metastases are present (yet) [2]. The risk of recurrence is determined by characteristics such as tumor size, lymph node involvement and the tumor grade. Together with the expression of estrogen (ER), progesterone (PR) and human epidermal growth factor receptor 2 (HER2) in tumor tissue, the risk of recurrence indicates whether patients may be eligible for adjuvant or neo-adjuvant systemic therapy (NAT) [3].

For patients with an indication for adjuvant or neo-adjuvant treatment, the particular type of treatment is based on the outcomes of ER, PR, and HER2 assessments [4]. In the majority (85.6%) of patients, the expression of ER, PR and HER2 is currently tested once, either on samples derived from tumor needle biopsies that were derived to establish the diagnosis, or on full tumor resection specimen derived after surgery [5]. In the remaining group of patients, receptor expression is determined multiple times and consequently, discordance in test results may be found [5-7].

If discordance in test results is found between assessments on the tumor needle biopsy and the tumor resection material, this may typically be expected to be found for patients receiving NAT, as the NAT may cause the biological characteristics of the tumor to change over time [8]. Several studies strongly suggest that, especially for those patients not receiving any form of NAT, discordance in test results is limited (range 2.5% - 9.2%) [9,10]. A more recent study including both, patients receiving NAT and patients not receiving NAT, confirmed that clinically meaningful discordance between repeated tests in general is limited. It was demonstrated that discordance in ER and HER2 is typically higher in patients that received chemotherapy (9.4% for ER, and 5.9% for HER2) or hormonal therapy (5.9% for ER and 0% for HER2) compared to patients who received no NAT at all (3.7% for ER and 0.5% for HER2) [5].

However, even though subgroups of patients having discordance in ER, PR and HER2 may be small, the expression of these receptors is of great importance to inform treatment decisions in both the neo-adjuvant and the adjuvant setting. When ER expresses positively on a patient's tumor, it is expected that patients will be sensitive to hormone therapy, whereas patients are eligible for anti-HER2 targeted therapy when HER2 is overexpressed [11,12]. For example, if ER expressed positively on tumor resection material, it is recommended that some type of adjuvant hormone therapy will be prescribed singly or in combination with chemotherapy in the adjuvant setting,

whereas it is recommended to prescribe at least one anti-HER2 targeted drug when HER2 overexpresses [13]. Nonetheless, when ER or HER2 expresses positively on tumor needle biopsy material but negative on full tumor resection material, treatment recommendations are less straightforward. However, no studies have yet evaluated whether discrepant findings in the receptor expressions indeed may lead to changes in clinical management. Therefore, this study evaluates whether discordance in ER and HER2 between the tumor needle biopsy and the tumor resection is accompanied by the expected treatment decisions concerning hormone- and HER2-targeting therapy.

METHODS

DATA SOURCES AND DESCRIPTION

Pathology reports of invasive primary breast cancer patients diagnosed after January 12, 2016 were derived from the Dutch Pathology Registry (PALGA). Discordance in tests results for the included patients was calculated in a previous study and defined as having a difference in results, i.e. positive or negative, between the tests performed on tumor biopsy and tumor resection material [5]. Data on the diagnostic tests performed and treatments prescribed were derived from the Netherlands Cancer Registry (NCR). Only patients with multiple tests, i.e. tests for either ER, PR or HER2 on both, tumor needle biopsy and tumor resection material, were included in the analysis. According to the Dutch breast cancer guideline, estrogen (ER), progesterone (PR) and Human Epidermal Growth Factor Receptor 2 (HER2) expression is determined using immunohistochemistry (IHC) for all three receptors and may be followed by an additional in-situ hybridization (ISH) to confirm the HER2 expression [3,13].

Data derived from the NCR contained detailed treatment information for both NAT and AT, including the specific drug, or type of treatment, that was prescribed. Both datasets were linked based on externally encrypted unique patient identifiers.

EXPECTED TREATMENT CONSEQUENCES

Drug treatment was classified into three groups, i.e. chemotherapy, hormone- or targeted therapy (therapy that is targeting specific cell characteristics/expressions, i.e. Trastuzumab which is targeting cells that express HER2) and used to evaluate whether discordance in ER or HER2 is followed by expected treatment decisions. The recommended treatment decisions for multiple types of discordance in each of these receptors are presented in table 1 [3]. As PR on its own is not singularly guiding treatment, this receptor is less relevant in clinical treatment decisions and is therefore not evaluated in this analysis.

Table 1: Discordance in receptor status and expected adjuvant treatment types

Receptor	Biopsy	Resection	Expected adjuvant treatment
ER	Negative	Positive	Hormone therapy
ER	Positive	Negative	Multiple therapy options*
HER2	Negative	Positive	Targeted therapy (Trastuzumab)
HER2	Positive	Negative	Multiple therapy options*

* Patients with a receptor status changing from positive to negative may be eligible for multiple types of therapy. For these patients, the different types of treatment will be evaluated during the analyses.

EXPLORATION OF ADJUVANT TREATMENTS

The general types of treatment (i.e. hormone-, targeted-, or chemotherapy) and the specific drugs prescribed (i.e. particular drug names, such as Docetaxel, Tamoxifen, Trastuzumab), were identified using the stats package in R (version 3.5.1). For patients with a biopsy status changing from negative to positive, it is expected that hormone therapy will be prescribed if ER changes to positive, and targeted therapy if HER2 changes to positive, as shown in table 1. Patients for whom the receptor status changed from positive to negative may be eligible for several treatment options. For all patients the different combinations of treatments that have been prescribed were identified. Adjuvant treatment was classified as matching the expected type of treatment based on the receptor expression and the type of adjuvant treatment and was independent of potential NAT prescribed.

LOGISTIC REGRESSION MODEL

Patient- and tumor characteristics other than the ER and HER2 status may explain why a specific treatment prescribed deviates from the expected type of treatment as shown table 1. To evaluate whether discordance in ER and HER2 had a significant influence on the prescription of hormone and targeted therapy, two logistic regression models were established using the glm function from the stats package in R (version 3.5.1).

Whether the treatment prescribed matched the expected type of treatment based on the receptor status, was used as the dependent categorical variable. The specific type of expected matching treatments were described in table 1. Consequently, if ER expressed positive, the adjuvant treatment decision was classified as matching when hormone therapy was prescribed in the first logistic regression model, whereas the prescription of targeted therapy was classified as matching when the tumor material was positive for HER2 in the second logistic regression model. In addition, several clinically relevant independent variables were included in both models, which were:

age, the bloom Richardson grade, sentinel gland involvement, the incidence year and whether a patient received neo-adjuvant chemo- (both models), hormone (ER model), or targeted (HER2 model) therapy. The first logistic regression model used the appropriateness of the treatment prescribed for the ER status. The dependent variable in this model was classified as being appropriate when either ER was positive on tumor resection material and hormone therapy was prescribed or when ER was negative and no hormone therapy was prescribed. The second logistic regression model used the appropriateness of the treatment prescribed for the HER2 status and classified the appropriateness of HER2 based on the prescription of targeted therapy in a similar way as the ER classifications. P-values for the estimated model coefficients were calculated using Chi-squared tests (for categorical variables) or two tailed Z-tests (for continuous variables).

RESULTS

Data on pathological test results was derived from PALGA and linked to the data containing treatment outcomes derived from the NCR. Discordance in test results was calculated for patients for whom the expression of ER and HER2 was measured on both, tumor biopsy and tumor resection specimen in a previous study [5]. For several patients, no information on whether patients received systemic treatment could be extracted from the NCR, so these patients were excluded from further analysis. The total number of patients for whom a record including the treatment information and a record including the information on the discordance in test results was derived, is shown in table 2.

Table 2: Overview of the number of patients included in this study

Receptor	Number of patients tested twice	Concordant test results			Discordant test results		
		Number of patients	Positive concordant	Patients with no treatment record	Number of patients	Biopsy positive, resection negative	Patients with no treatment record
ER	671	642 (95.68%)	463 (72.12%)	11 (1.64%)	29 (4.32%)	19 (65.52%)	2 (0.30%)
HER2	694	688 (99.14%)	33 (4.80%)	13 (1.87%)	6 (0.86%)	0 (0.00%)	0 (0.00%)

THERAPIES PRESCRIBED

As presented in table 3, discordant subgroups of patients in whom the ER and HER2 receptor were tested twice are rather small, ranging from 6 patients with HER2 discordant test results, to 19 patients with initially ER positive discordant test results. A substantial number of patients in did not receive any form of adjuvant therapy, ranging from 16.7% in the HER2 discordant patient group till 50.8% in the ER negative concordant patient group, respectively. The presented treatment combinations reflect the aggregated treatment groups. Specific types of drugs prescribed are presented in appendix 1.

For those patients with an either positive or negative concordant ER status (n=642), the majority received treatment that matched the treatment expectations (n=472; 73.52%). A substantial number of patients with a concordant ER positive receptor status (n=165; 35.6%) did not receive any form of adjuvant hormone therapy (consisting of the patients receiving no therapy, only chemotherapy or chemotherapy and targeted therapy). Most patients were staged in a low class TNM stadium, with 123 patients in stage IA-IB (74.55%), 22 in stage IIA (13.33%), 9 in stage IIB (5.45%) and 5 in stage IIIA-IIIC (3.03%). The vast majority of these patients did not receive any form of neo adjuvant systemic therapy (n=142; 86.06%). A minority of patients with a concordant ER negative receptor status (n=5; 2.8%) received hormone therapy in combination with chemotherapy.

For patients for whom the ER receptor was discordant (n=29), a small majority received the type of therapy that matched the expectations (n=15; 51.72%). A substantial number of patients received no therapy (n=9; 31.03%) or hormone therapy (n=8; 27.59%). Of the remaining patients, the majority received different treatment combinations of hormone therapy with either chemotherapy, targeted therapy or both. In the majority of patients with an ER discordant test result, the ER receptor status changed from positive to negative (n=19; 65.52%). The change in the percentage of cells that expressed positive for ER was substantially diversified over all patients with a discordant test result, ranging from 4% till 100% change in receptor expression. Most patients did not receive any form of neo adjuvant therapy (n=20; 68.97%). The majority of these ER discordant patients had an invasive ductal carcinoma (n=22; 75.86%) and were staged IA (n=18; 62.07%) or IIA (n=9; 31.03%).

Table 3: Adjuvant therapy prescriptions within all combinations of the receptor statuses on biopsy and resection material

	Total number of patients (n)	No therapy	Hormone therapy	Chemo-therapy	Targeted therapy	Chemo- and hormone therapy	Chemo- and targeted therapy	Hormone and targeted therapy	Chemo- hormone and targeted therapy
ER –	10	3	4	1	0	1	0	0	1
to +		(30.00%)	(40.00%)	(10.00%)	(0.00%)	(10.00%)	(0.00%)	(0.00%)	(10.00%)
ER –	179	91	0	64	1	5	18	0	0
to -		(50.84%)	(0.00%)	(35.75%)	(0.56%)	(2.79%)	(10.06%)	(0.00%)	(0.00%)
ER +	19	6	4	2	0	4	1	1	1
to -		(31.58%)	(21.05%)	(10.53%)	(0.00%)	(21.05%)	(5.26%)	(5.26%)	(5.26%)
ER +	463	158	225	6	0	58	1	3	12
to +		(34.13%)	(48.60%)	(1.30%)	(0.00%)	(12.53%)	(0.22%)	(0.65%)	(2.59%)
HER2 – to +	6	1	2	0	1	0	0	0	2
		(16.67%)	(33.33%)	(0.00%)	(16.67%)	(0.00%)	(0.00%)	(0.00%)	(33.33%)
HER2 – to -	655	264	277	47	0	67	0	0	0
		(40.31%)	(42.29%)	(7.18%)	(0.00%)	(10.23%)	(0.00%)	(0.00%)	(0.00%)
HER2 + to +	33	12	6	0	0	0	6	1	8
		(36.36%)	(18.18%)	(0.00%)	(0.00%)	(0.00%)	(18.18%)	(3.03%)	(24.24%)

The vast majority of patients with HER2 concordant test results (n=684) received treatment that was matching the treatment expectations (n=670; 97.95%). Several patients (n=18; 54.5%) did not receive any form of targeted therapy while the HER2 receptor was concordant positive (consisting of those patients receiving no therapy or only hormone therapy). Of these patients, 9 were staged TNM stage IA (50.00%), 4 stage IIA (22.22%), 2 stage IIB (11.11%) and 1 in stage IIIA (5.55%). The majority of these patients did not receive any type of neo adjuvant therapy (n=16; 88.89%) and had an invasive ductal carcinoma (n=16; 88.89%).

No patients with a HER2 negative concordant test result received targeted therapy. For patients with a discordant HER2 test result (HER2 negative biopsy but HER2 positive resection), the treatment prescribed was more diversified than for those patients with a HER2 concordant test result. Half of these patients (n=3; 50.00%) received targeted therapy, either in combination with chemotherapy and hormone therapy or singular. Of those patients that did not received any form of targeted therapy (n=3; 50.00%), the majority (n=2; 33.33) received neo adjuvant chemotherapy (n=2; 33.33%). The majority of patients had a ductal carcinoma (n=5; 83.33%) and a TNM stage IA (n=3; 50.00%) or IIA (n=2; 33.33%).

There were no patients whom had a HER2 positive biopsy but a HER2 negative resection so this patient group was excluded from table 3.

Of those patients with an ER concordant positive test result that did not receive any type of systemic therapy (n=158; 34.13%), the majority of patients also had a negative receptor status for HER2 on either tumor needle biopsy material, tumor resection material, or both (n=149; 94.30%). Regardless of the ER status, several patients have received targeted therapy, either singly or in combination with chemotherapy or hormone therapy. All of these patients had a positive expression of HER2 on either tumor needle biopsy material, tumor resection material, or both. All patients with a concordant HER2 negative test result and who received hormone therapy (n=277; 42.29%) had a positive expression of ER on either tumor needle biopsy material, tumor resection material, or both. An overview of the therapies prescribed in each of these combined subgroups (ER and HER2) is presented in appendix 2.

LOGISTIC REGRESSION MODEL

The results of the multivariable logistic regression model with the appropriateness of ER related therapy as the dependent variable are presented in table 4. Several variables have a significant influence on the probability that a patient will receive a different therapy than the therapy that would match the treatment expectations when only considering ER expression.

Patients with an ER positive classification on tumor resection material have a higher probability of receiving treatment that is not matching the expected type of treatment (OR = 5.27), whereas the probability of receiving treatment that is not matching the expected type of treatment is lower for patients with an ER concordant test result (OR = 0.087). When combining the different patient and tumor characteristics, subgroups of patients can be identified which have the highest and lowest probability of receiving treatment that is not matching the expected type of treatment.

From table 3 it can be identified that 184 (27.42%) patients received treatment that is not matching the expectations when ER was tested twice, whereas 487 (72.58%) patients did receive treatment that did match the expectations. From the results of the logistic regression model, patients with an ER positive test result that did receive neo-adjuvant hormone therapy seem to have a high probability of receiving treatment that did not match the expectations ($OR > 1$), whereas patients with an ER concordant test result that did receive chemotherapy have low probability of receiving treatment that did not match the expectations ($OR < 1$).

Table 4: Raw coefficients derived from the logistic regression model with ER related treatment as the dependent variable

		Odds ratio	2.5% CI	97.5% CI	p-value **
(Intercept)		7.441	1.049	56.713	0.048
Age		0.991	0.971	1.01	0.346
Resection ER positive		5.270	2.484	12.228	< 0.01
(ref = negative)					
ER concordant		0.087	0.028	0.263	< 0.01
(ref = discordant)					
Sentinel Node (ref = negative)	= < 0.2 mm	0.471	0.163	1.218	< 0.01
	Micrometastasis 0.2 - 2 mm	0.066	0.019	0.186	
	Metastasis > 2 mm	0.137	0.057	0.293	
	Sentinel node not found	1.923	0.456	8.483	
Bloom Richardson	Grade 1	1.606	0.423	5.987	< 0.01
	Grade 2	0.295	0.08	1.055	
	Grade 3	0.140	0.037	0.513	
(ref = Unknown)					
Neo-adjuvant chemotherapy		0.220	0.067	0.624	< 0.01
(ref = none)					
Neo-adjuvant hormone therapy		13.397	4.115	52.22	< 0.01
(ref = none)					

* Prob. = Probability (to have received treatment that is not matching the expectations) ** For categorical variables, a Chi-squared test was performed to calculate the p-values. For continuous variables, two-tailed Z-tests were used to calculate the p-values. CI = Confidence Interval

Based on the coefficients of the logistic regression model presented in table 4, the probability of receiving treatment that is not matching the expectations was calculated for all patients in whom the ER receptor was tested twice. Patients with an ER positive tumor resection that did receive neo-adjuvant hormone therapy have a relatively high mean probability of receiving treatment that does not match the expectations (76.19%; n=27). Patients with an ER concordant test result that did receive neo-adjuvant chemotherapy, have a relatively low mean probability of receiving treatment that is not matching the expectations (10.37%; n=81).

A similar logistic regression model was developed for patients for whom the HER2 status was measured on both, tumor needle biopsy and tumor resection material. The HER2 classification on tumor resection material was excluded as one of the dependent variables in the logistic regression model as there were no patients that have received treatment that was not matching the expectations when they had a negative tumor resection status, or in other words, there were no patients with a negative tumor resection status who received adjuvant targeted therapy. The detailed sentinel node status was also excluded from the model, as for several categories there were no patients that received treatment which was not matching the expectations. For a small subgroup of patients (n=18), the Bloom Richardson grade was unknown. As all these patients had appropriate treatment prescriptions, this subgroup was excluded from the model. Neo adjuvant targeted therapy was not included in the model, as there were no patients that received neo adjuvant targeted therapy for whom treatment was prescribed that matched the expected type of treatment.

From table 3 it can be identified that 21 (3.17%) patients received treatment that is not matching the expectations when HER2 was tested twice, whereas 641 (96.83%) patients did receive treatment that did match the expectations. Based on the coefficients of the logistic regression model presented in table 5, the probability of receiving treatment that is not matching the expectations was calculated for all patients in whom the HER2 receptor was tested twice. Patients that received neo-adjuvant chemotherapy have a relatively high mean probability of receiving treatment that does not match the expectations (9.38%; n=32), whereas patients with a concordant test result have a relatively low mean probability of receiving treatment that does not match the expectations (p=2.74%; n=656).

Table 5: Raw coefficients derived from the logistic regression model with HER2 related treatment as the dependent variable

		Odds ratio	2.5% CI	97.5% CI	P-value **
(Intercept)		0.103	0.005	1.919	0.130
Age		1.029	0.991	1.071	0.069
HER2 concordant		0.026	0.004	0.181	< 0.001
(ref = discordant)					
Bloom Richardson Grade (ref = Grade I)	Grade II	1.406	0.446	5.171	0.348
	Grade III	2.560	0.748	9.865	
Neo-adjuvant chemotherapy		2.633	0.383	12.037	0.292
(ref = no chemotherapy)					

* Prob. = Probability (to have received treatment that is not matching the expectations) ** For categorical variables, a Chi-squared test was performed to calculate the p-values. For continuous variables, two-tailed Z-tests were used to calculate the p-values. CI = Confidence Interval

In both models, concordance in receptor expression was a significant predictor of whether treatment will be prescribed that matches the expected type of treatment based on receptor expressions

DISCUSSION

The results of this study imply that for patients with double ER and HER2 measurements for whom the test results were found to be concordant (positive or negative) on both, tumor needle biopsy and tumor resection material, the treatment prescribed is matching the expected type of treatment in the majority of patients. Treatment decisions were however, substantially more diversified for those patients for whom a discordant test result was found for either ER or HER2. These results can partially be explained by the fact that physicians may face substantial uncertainty about the actual receptor expression when its status differs between the measurements on tumor needle biopsy and the tumor resection. In addition, there may be multiple underlying causes of such discordance. Cancer in general has the potential to change receptor expressions over time which subsequently may have caused the discordance. However, also measurement errors in the tests such as sampling errors of the tumor needle biopsies or the sensitivity of the test used to determine the expression may potentially have resulted in false negative or false positive test results.

Though the majority of patients with concordant test results received a treatment that matched the treatment expectations based on the receptor status, there is, however,

a relatively large subgroup of patients with positive concordant test results for HER2 in whom no targeted therapy was prescribed (n=18; 54.54%). The same yields for patients with an ER positive concordant test result, as a substantial subgroup of these patients did not receive adjuvant hormone therapy (n=165; 35.64%). For both concordant positive groups, treatment was not matching the expected type of treatment. Most of these patients were staged in a low TNM stage. These results may imply that the prognosis of the patients potentially was so good that adjuvant systemic therapy was deemed unnecessary.

Both logistic regression models demonstrated that patients with concordance in receptor expression have a very low probability of receiving treatment that is not-matching the treatment expectations when considering the ER and HER2 status. Several clinically relevant variables significantly influence the appropriateness of treatment, especially in the model in which the ER related treatment decisions were modelled. As patient characteristics of those patients for whom the HER2 status was determined on both, tumor needle biopsy and tumor resection material were skewed, several clinically potentially relevant variables could currently not be used to develop the multivariable logistic regression model.

In this study, no information on contra-indications for particular therapies was available in the data used for this analysis. To be able to evaluate whether the most appropriate (i.e. effective) treatments have been prescribed beyond the extent to which these match the receptor expressions, such information should be taken into account. Furthermore, data collection of the NCR is mainly focused on collecting treatment information regarding the first year of treatment. Though prescribing systemic treatment only after the first year is infrequent, this is currently not taken into account in the analysis. In addition, no survival data could be derived (yet), as patients were only recently diagnosed and treated (cohort diagnosed between 2016-2018). Future studies should evaluate whether potential differences in the treatment prescribed do result in different survival outcomes, especially in those patients for whom discordant test results have been found.

CONCLUSIONS

The majority of patients for whom ER and HER2 were determined on both, tumor needle biopsy and tumor resection material, had concordant test results. For those patients, the type of treatment prescribed usually matches the expected type of treatment when only considering the receptor status. However, for a minority of patients with a concordant ER or HER2 positive test result, no hormone or targeted therapy was prescribed in the adjuvant setting, thereby not matching the treatment expectations based on the receptor status.

Several parameters, including discordance in ER and HER2 between tumor needle biopsy and tumor resection material, seem to be significantly influencing whether patients have received the most appropriate treatment. Potential causes that may be underlying the decisions that caused treatment to be not matching to the expected type of treatment based on the receptor status should be identified in future studies. Therefore, future research should focus on including a larger number of patients. In addition, the consequences in terms of survival, quality of life and costs should be evaluated.

DECLARATIONS

Table of abbreviations

ER	Estrogen
HER2	Human Epidermal Growth Factor Receptor 2
IHC	Immunohistochemistry
ISH	In Situ Hybridization
NAT	Neo Adjuvant Therapy
NCR	Netherlands Cancer Registry
PALGA	Dutch Pathology Registry
PR	Progesterone

Acknowledgements: None.

Funding: This study was not funded

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable

Data availability: The datasets used and analyzed in this study are not publicly available but can be requested at PALGA and the NCR.

Conflicts of interest: The authors declare that they have no conflict of interest.

Authors’ contributions: The initial data linking and analysis was performed by AMSB. The identification of clinically meaningful variables and subgroups of patients were discussed with CHMD, SS and LWMMT. The model algorithms and statistical approaches were discussed with and checked by HK and MIJ. All authors were involved in writing and revising the manuscript.

Keywords: Breast cancer; discordance; estrogen; human epidermal growth factor receptor 2; systemic treatment; targeted therapy; hormone therapy; adjuvant therapy

REFERENCES

1. Netherlands Cancer Registry (NKR) 2018. Nationwide breast cancer incidence.
2. eSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A (2017) Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin* 67 (6):439-448. doi:10.3322/caac.21412
3. NABON (2012) Breast Cancer Guideline.
4. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, Zackrisson S, Cardoso F (2015) Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 26 Suppl 5:v8-30. doi:10.1093/annonc/mdv298
5. Berghuis AMS, van Deurzen CHM, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ (2019) Real-world data on discordance between estrogen, progesterone, and HER2 receptor expression on diagnostic tumor biopsy versus tumor resection material. *Breast Cancer Res Treat.* doi:10.1007/s10549-019-05141-y
6. Yang L, Zhong X, Pu T, Qiu Y, Ye F, Bu H (2018) Clinical significance and prognostic value of receptor conversion in hormone receptor positive breast cancers after neoadjuvant chemotherapy. *World J Surg Oncol* 16 (1):51. doi:10.1186/s12957-018-1332-7
7. Asogan AB, Hong GS, Arni Prabhakaran SK (2017) Concordance between core needle biopsy and surgical specimen for oestrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 status in breast cancer. *Singapore Med J* 58 (3):145-149. doi:10.11622/smedj.2016062
8. Gahlaut R, Bennett A, Fatayer H, Dall BJ, Sharma N, Velikova G, Perren T, Dodwell D, Lansdown M, Shaaban AM (2016) Effect of neoadjuvant chemotherapy on breast cancer phenotype, ER/PR and HER2 expression - Implications for the practising oncologist. *Eur J Cancer* 60:40-48. doi:10.1016/j.ejca.2016.03.006
9. Singh K, Wang Y, Marketkar S, Kalife ET, Steinhoff MM (2017) Comparison of estrogen receptor, progesterone receptor and HER2 results in concurrent ipsilateral samples with invasive breast carcinoma: a retrospective study of 246 biopsies from 119 patients. *Hum Pathol* 65:123-132. doi:10.1016/j.humpath.2017.04.016
10. Hadad SM, Jordan LB, Roy PG, Purdie CA, Iwamoto T, Pusztai L, Moulder-Thompson SL, Thompson AM (2016) A prospective comparison of ER, PR, Ki67 and gene expression in paired sequential core biopsies of primary, untreated breast cancer. *BMC Cancer* 16 (1):745. doi:10.1186/s12885-016-2788-x
11. Mueller C, Haymond A, Davis JB, Williams A, Espina V (2018) Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev Proteomics* 15 (2):131-152. doi:10.1080/14789450.2018.1421071
12. Schramm A, De Gregorio N, Widschwendter P, Fink V, Huober J (2015) Targeted Therapies in HER2-Positive Breast Cancer - a Systematic Review. *Breast Care (Basel)* 10 (3):173-178. doi:10.1159/000431029
13. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel M (2015) 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* 26 (8):1533-1546. doi:10.1093/annonc/mdv221

APPENDIX 1 – THERAPIES PRESCRIBED IN EACH RECEPTOR SUBGROUP

Table 1: The specified prescription of all drugs in each receptor status subgroup

	ER- to ER+	ER- to ER-	ER + to ER -	ER + to ER+	HER2 – to HER2 +	HER2 – to HER2 -	HER2+ to HER2+ *
No systemic treatment	3 (30%)	91 (50.8%)	6 (31.6%)	158 (34.1%)	1 (16.7%)	256 (40.3%)	11 (35.5%)
Trastuzumab	0 (0%)	1 (0.6%)	0 (0%)	0 (0%)	1 (16.7%)	0 (0%)	0 (0%)
5-FU, Epirubicine, Cy- clofosfamide (FEC/CEF), Tamoxifen	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)
Anastrozol	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.3%)	0 (0%)
Anthracycline	0 (0%)	2 (1.1%)	0 (0%)	0 (0%)	0 (0%)	2 (0.3%)	0 (0%)
Anthracycline, Hor- monal therapy through drugs	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)
Anthracycline, Tamox- ifen	0 (0%)	0 (0%)	0 (0%)	3 (0.6%)	0 (0%)	2 (0.3%)	0 (0%)
Anthracycline, Taxane	0 (0%)	46 (25.7%)	2 (10.5%)	3 (0.6%)	0 (0%)	35 (5.5%)	0 (0%)
Anthracycline, Taxane, Aromastase inhibitors	0 (0%)	0 (0%)	0 (0%)	11 (2.4%)	0 (0%)	10 (1.6%)	0 (0%)
Anthracycline, Taxane, Aromastase inhibitors, Trastuzumab	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Anthracycline, Taxane, Hormonal therapy through surgery	0 (0%)	2 (1.1%)	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)
Anthracycline, Taxane, Tamoxifen	1 (10%)	3 (1.7%)	3 (15.8%)	40 (8.6%)	0 (0%)	47 (7.4%)	0 (0%)
Anthracycline, Taxane, Tamoxifen, Trastuzum- ab	0 (0%)	0 (0%)	1 (5.3%)	5 (1.1%)	1 (16.7%)	0 (0%)	1 (3.2%)
Anthracycline, Taxane, Trastuzumab	0 (0%)	7 (3.9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (6.5%)
Aromastase inhibitors	0 (0%)	0 (0%)	0 (0%)	50 10.8%)	0 (0%)	50 (7.9%)	0 (0%)

Aromastase inhibitors, Trastuzumab	0 (0%)	0 (0%)	0 (0%)	2 (0.4%)	0 (0%)	0 (0%)	0 (0%)
Capecitabin	1 (10%)	5 (2.8%)	0 (0%)	0 (0%)	0 (0%)	2 (0.3%)	0 (0%)
Capecitabin, Tamoxifen	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Docetaxel	0 (0%)	1 (0.6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Doxorubicine, Cyclofos- famide (AC)	0 (0%)	5 (2.8%)	0 (0%)	2 (0.4%)	0 (0%)	4 (0.6%)	0 (0%)
Doxorubicine, Cyclofos- famide (AC), Hormonal therapy through drugs	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Doxorubicine, Cyclo- fosfamide (AC), LHRH Agonist	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Doxorubicine, Cyclofos- famide (AC), Tamoxifen	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)
Hormonal therapy through surgery	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	1 (0.2%)	0 (0%)
LHRH Agonist	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Paclitaxel, Aromastase inhibitors, Trastuzumab	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.2%)
Systemic chemotherapy	0 (0%)	3 (1.7%)	0 (0%)	0 (0%)	0 (0%)	2 (0.3%)	0 (0%)
Systemic chemothera- py, Trastuzumab	0 (0%)	1 (0.6%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Tamoxifen	4 (40%)	0 (0%)	4 (21.1%)	1 (37.4%)	2 (33.3%)	214 (33.7%)	5 (16.1%)
Tamoxifen, Trastuzum- ab	0 (0%)	0 (0%)	1 (5.3%)	1 (0.2%)	0 (0%)	0 (0%)	1 (3.2%)
Taxane	0 (0%)	2 (1.1%)	0 (0%)	1 (0.2%)	0 (0%)	1 (0.2%)	0 (0%)
Taxane, Aromastase inhibitors, Trastuzumab	0 (0%)	0 (0%)	0 (0%)	2 (0.4%)	0 (0%)	0 (0%)	1 (3.2%)
Taxane, Tamoxifen	0 (0%)	0 (0%)	1 (5.3%)	1 (0.2%)	0 (0%)	3 (0.5%)	0 (0%)
Taxane, Tamoxifen, Pertuzumab	0 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	1 (3.2%)
Taxane, Tamoxifen, Trastuzumab	1 (10%)	0 (0%)	0 (0%)	3 (0.6%)	1 (16.7%)	0 (0%)	4 (12.9%)
Taxane, Trastuzumab	0 (0%)	10 (5.6%)	1 (5.3%)	0 (0%)	0 (0%)	0 (0%)	4 (12.9%)

APPENDIX 2 – ALL COMBINATIONS OF RECEPTOR EXPRESSIONS AND TREATMENT

Table 1: The specified prescription of all drugs in each receptor status subgroup

	HER2 biopsy	HER2 resection	Totals	No therapy	Hormone therapy	Chemo-therapy	Targeted therapy	Chemo- and hormone therapy	Chemo- and targeted therapy	Hormone and targeted therapy	Chemo- hormone and targeted therapy
ER - to +	Negative	Negative	6	2	2	1		1			
	Negative	Unknown	1	1							
	Unknown	Negative	1		1						
	Unknown	Unknown	1		1						
	Positive	Unknown	1								1
Totals			10	3	4	1		1			1
ER + to -	Negative	Negative	11	5	2	2		2			
	Negative	Positive	1								1
	Unknown	Negative	3	1				2			
	Positive	Unknown	2		1				1		
	Positive	Positive	2		1					1	
Totals			19	6	4	2		4	1	1	1
ER + to +	Negative	Equivocal	3	2	1						
	Negative	Negative	344	120	174	6		44			
	Negative	Unknown	35	17	15			3			
	Negative	Positive	2		2						
	Unknown	Equivocal	2	1	1						
	Unknown	Negative	42	12	21			9			
	Unknown	Unknown	2		1			1			
	Unknown	Positive	6						1		5
	Positive	Equivocal	2		1					1	
	Positive	Negative	1					1			
	Positive	Unknown	3	1						1	1
	Positive	Positive	21	5	9					1	6
Totals			463	158	225	6		58	1	3	12
ER - to -	Negative	Equivocal	3	3							
	Negative	Negative	86	47		37		2			
	Negative	Unknown	48	30		17		1			
	Negative	Positive	1				1				
	Unknown	Negative	13	2		9		2			
	Unknown	Positive	4	1					3		
	Positive	Unknown	14	5		1			8		
	Positive	Positive	10	3					7		
Totals			179	91		64	1	5	18		

Table 2: Therapies prescribed in all subgroups of HER2 broken down by the ER status

	ER biopsy	ER Resection	Totals	No therapy	Hormone therapy	Chemo-therapy	Targeted therapy	Chemo- and hormone therapy	Chemo- and targeted therapy	Hormone and targeted therapy	Chemo- hormone and targeted therapy
HER2	Negative	Unknown	3	1		2					
- to -	Negative	Negative	80	44		34		2			
	Negative	Positive	2	1	1						
	Positive	Unknown	308	114	150	7		37			
	Positive	Negative	4	3				1			
	Positive	Positive	258	101	126	4		27			
		Totals	655	264	277	47		67			
HER2	Negative	Negative	1				1				
- to +	Positive	Unknown	2	1							1
	Positive	Negative	1								1
	Positive	Positive	2		2						
		Totals	6	1	2		1				2
HER2	Negative	Unknown	3	3							
+ to	Negative	Negative	8	2				6			
+	Positive	Unknown	10	5	1					1	3
	Positive	Negative	1		1						
	Positive	Positive	11	2	4						5
		Totals	33	12	6			6		1	8

CHAPTER 5

DETECTING BLOOD BASED BIOMARKERS IN METASTATIC BREAST CANCER: A SYSTEMATIC REVIEW OF THEIR CURRENT STATUS AND CLINICAL UTILITY

A.M.S. BERGHUIS, H. KOFFIJBERG, J. PRAKASH,
L.W.M.M. TERSTAPPEN, M.J. IJZERMAN

THIS CHAPTER WAS PUBLISHED AS:

A.M.S. BERGHUIS, H. KOFFIJBERG, J. PRAKASH, L.W.M.M. TERSTAPPEN, M.J. IJZERMAN. DETECTING BLOOD BASED BIOMARKERS IN METASTATIC BREAST CANCER: A SYSTEMATIC REVIEW OF THEIR CURRENT STATUS AND CLINICAL UTILITY. INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, 09-02-2017.

DOI: 10.3390/IJMS18020363

ABSTRACT

Reviews on circulating biomarkers in breast cancer usually focus on one single biomarker or a selective group of biomarkers. An overview summarizing the discovery and evaluation of all blood-based biomarkers in metastatic breast cancer is lacking. This systematic review aims to identify the available evidence of known blood-based biomarkers in metastatic breast cancer, regarding their clinical utility and state-of-the-art position in the validation process. The initial search yielded 1078 original studies, of which 420 were assessed for eligibility. A total of 320 studies were included in the final synthesis. A Development, Evaluation and Application Chart (DEAC) of all biomarkers was developed. Most studies focus on identifying new biomarkers and search for relations between these biomarkers and traditional molecular characteristics. Biomarkers are usually investigated in only one study (68.8%). Only 9.8% of all biomarkers was investigated in more than five studies. Circulating tumor cells, gene expression within tumor cells and the concentration of secreted proteins are the most frequently investigated biomarkers in liquid biopsies. However, there is a lack of studies focusing on identifying the clinical utility of these biomarkers, by which the additional value still seems to be limited according to the investigated evidence.

INTRODUCTION

BREAST CANCER SURVIVAL

Globally, breast cancer is the most commonly diagnosed form of cancer among women. Clinical management has improved over the last years, and the development of genetic tests such as Mammaprint and OncoTypeDX have proven to guide treatment in early stage breast cancer. Although the current 5-year survival for primary breast cancer is relatively high (ranging from 80% to 92% in different populations) [1], survival rates decrease to less than 25% when the disease becomes metastatic [1,2]. The most important factor to increase survival for those suffering from metastatic breast cancer, is to prescribe a treatment that has the most likelihood of being effective, guided by the tumor cell characteristics [3,4]. To select the most effective treatment once the metastatic lesions have been detected, it is essential to obtain accurate information on the characteristics of the tumor cells at the time therapy is to be initiated [5].

DETECTION AND TREATMENT OF METASTATIC LESIONS

Technical advances in the molecular characterization of cells has already lead to accurate predictions of survival and treatment efficacy. However, these molecular characterizations require high-quality biopsies, which cannot always be obtained from the primary tumor [6]. Alternatively, taking a biopsy of the metastatic lesion is either difficult or even impossible, for example, due to its location, or the inability to visualize that location with the currently used imaging techniques [6–8]. Furthermore, previous research has shown that molecular aberrations of the primary tumor may differ from that of the metastatic lesion and different metastatic lesions can have different characteristics [9]. Therefore, there remains a need for new tests which are sufficiently sensitive and reflect the composition of the tumor at all sites to guide treatment of metastatic disease.

THE USE OF BLOOD BASED BIOMARKERS

A possible way of enabling better treatment response monitoring or treatment guidance is the use of blood-based biomarkers or liquid biopsies [10]. A large number of single blood-based biomarkers can be distinguished in the blood, of which the most commonly known soluble proteins are Human Epidermal Growth Factor Receptor 2 (HER2), Cancer Antigen 15-3 (CA 15-3), Carcinoembryonic Antigen (CEA) and MUC1 [11]. Furthermore, all kinds of gene expression patterns or mutations can be extracted from circulating mRNA or circulating free DNA [8,12]. However, not only

proteins or gene expression patterns yield prognostic or predictive information, even complete cells found in the blood—such as Circulating Tumor Cells (CTCs) or Cancer Associated Fibroblasts (CAF)—provide this type of information.

Although a range of different biomarkers is known, it is far more difficult to evaluate their usefulness for treatment targeting or prognosis of disease. It therefore is required to develop a classification, both to determine biomarkers with clinical utility and to prioritize future research. For clinical decision making, there are different ways of classifying diagnostic information[10]. Classifications focus, for example, on prognostic or predictive ability, or on a classification according to specific hallmarks of cancer [13].

EVIDENCE ON THE UTILITY OF BIOMARKERS

Up to now, the literature is not clear about the clinical utility of biomarkers in breast cancer. Several systematic reviews on blood-based biomarkers have been published yet [10,13]. However, these studies usually focus on one single biomarker or a selective group of biomarkers. These reviews are helpful to understand specific molecular pathways of oncogenesis, on specific prognostic information and all other outcomes they are related to, or on both.

5

An overview summarizing the discovery and evaluation of blood-based biomarkers for metastatic disease, in terms of their current status and future potential for clinical application, is still lacking. Therefore, this systematic review focusses on identifying known biomarkers, the available evidence regarding their clinical utility and exploring the current state-of-the-art in the validation process of all blood-based biomarkers in metastatic breast cancer. The review aims to identify a set of blood-based biomarkers that may have substantial future potential.

Whereas it is common to focus on outcomes in terms of effectiveness, this review instead focusses on the developmental stage as the primary outcome measure of the included studies. First, all blood-based biomarkers will be identified and classified according to their developmental stage (e.g., from discovery to clinical utility). Second, the set of biomarkers with the highest future potential for clinical application will be identified by the number of studies that have been performed in each of the developmental stages.

CONCLUSIONS

The main aim of research on blood-based biomarkers in metastatic breast cancer is the identification of new biomarkers or relations of these biomarkers with other

original molecular tumor characteristics. Especially gene expression within CTCs is investigated frequently. However, there still is a lack of studies identifying the clinical utility of these biomarkers. Thereby, the additional value for these biomarkers seems to be still limited according to the investigated evidence.

RESULTS

SEARCH RESULTS

The initial search resulted in a total of 1249 studies from all databases searched. After screening all abstracts, 410 studies were further assessed for eligibility. During the assessment for eligibility, 91 studies were excluded. A total of 320 studies were included in this review. The full list of all studies that were included is presented in Appendix C. Most studies were excluded because the biomarkers investigated were not extracted from metastatic breast cancer patients ($n = 22$; 24.4%), because the blood used in the detection of the biomarker was non-human or was injected with a cell line that had just metastatic potential ($n = 19$; 21.1%) or because the study investigated multiple stages of breast cancer but had not reported conclusions for metastatic breast cancer separately ($n = 17$; 18.9%). The flow diagram of the search is presented in Figure 1.

Figure 1: PRISMA Flow Diagram

STUDY CHARACTERISTICS

For each study, the data were extracted and two classifications were made. First, the biomarkers were classified in one of the four general categories. Second, studies were classified in one of the pre-defined developmental stage categories, as defined in Figure 2. To illustrate the classification more clearly, citations of those studies which were classified as being in one of these phases are given in the right column of Figure 2.

5

General Stage	Specified stages	General description of the phase	Example of conclusions in articles classified in this phase
Technical Validation	Basic Research	This phase concerns the evaluation of basic principles of new techniques that are currently developed to detect the biomarkers. In this phase research to improving parts of the whole detection technique is presented.	"With the CellSearch platform, a 10XNA0.45 objective is adequate for this purpose, but we could not image the track-etched filter at this resolution because it had deformed during filtration. The filter material must be sufficiently stiff to prevent deformation during filtration." [16]
	Proof of concept	This phase concerns the proof of concept of detecting the biomarkers. In this phase evidence is given that the biomarker concerned can be detected with this new method and that repeated measurements give the same result.	"This study demonstrates the feasibility of expression profiling of pure CTCs. Using IE/FACS, we isolated small populations of CTCs and amplified their RNA for micro-array analysis." [17]
	Method comparisons	In this phase a new method is compared to the reference standard. It is shown what differences there are in the detection of the biomarker concerned, and what the performance of both methods of detection is.	"In the present study we compared three molecular assays for the detection and molecular characterization of CTC after excluding all errors in the pre-analytic variables, such as sample isolation, sample volume, logistics and storage conditions, as well as important analytic variables such as the CTC isolation methodology." [18]
Identification	Basic predictive research	This phase presents research on thresholds, percentages of patients that are positively diagnosed according to that threshold, concentrations detected and eventually comparisons to other studies reporting these outcome measures.	"In comparing patients with a relatively high number of CTCs to patients with a relatively low number, we found that for all three types of cancer, NK cell cytotoxic activity has diminished significantly." [19]
	Basic prognostic research	This phase presents research that couples all outcome measures recalled in basic predictive research to a patient level outcome measure. Outcome measures included are for example OS, PFS, HR's and other survival measures.	"ROC curve analysis showed that a cutoff value of 1 CTC has the highest sensitivity/specificity to predict OS. Of 194 patients, 115 had ≥ 119 CK19-releasing CTC(59%) per 7.5 mL of blood." [20]
Clinical Validation	Prognostic validation	This phase presents research that compares the result of the new study to already presented final outcomes studies. Research intention is to validate the final outcomes for patients within another, new population.	"These data confirm the independent prognostic effect of CTC count on progression-free survival and overall survival. CTC count also improves the prognostication of metastatic breast cancer when added to full clinicopathological predictive models, whereas serum tumour markers do not." [21]
	Observational Research	This phase concerns observational research in clinical practice. Of all patients involved in these studies the biomarker concentrations are measured (repeatedly) and reported in combination with treatment and a clinical outcome measure.	"Triple-staining experiments revealed that the full-length receptor was expressed in seven out of 14 patients before treatment and in four patients after trastuzumab administration ($p=0,035$)." [22]
Utility	Clinical Trials	This phase concerns clinical trials for the biomarker concerned. These studies report the differences in clinical outcomes based on the new clinical pathway which is led by according to concentrations of the biomarker.	"In this prospective clinical trial, we tested whether changing to an alternative chemotherapeutic regimen might improve outcomes for patients with MBC whose CTCs were not reduced after one cycle of first-line chemotherapy. Unfortunately, this strategy did not improve either OS, or surprisingly, even PFS for such patients." [23]
	Health Economic Evaluations	This phase concerns all health economic evaluations that have been done for these biomarkers. Health outcomes need to be compared with economic outcomes in terms of additional value and eventual cost differences.	No studies performed in this phase

Figure 2: Stages of clinical translation in biomarker discovery

RESULTS ACCORDING TO DEVELOPMENTAL PHASE

Figure 3 presents the DEAC with the distribution of studies over developmental phases. From this figure it is apparent that most studies focused on the identification phase. This means that most studies focus on finding relationships between the concentration of the biomarker, in relation to a new or existing threshold and furthermore, try to evaluate this against an outcome measure in terms of survival (e.g., Overall Survival (OS), Progression Free Survival (PFS) or survival in months). This phase is split up over two sub phases, namely basic predictive and basic prognostic research. For predictive research, only the concentrations in a subgroup of metastatic breast cancer patients were reported. For prognostic research, these concentrations were linked to an outcome measure related to survival (OS or PFS).

Figure 3: Development, Evaluation and Application Chart

RESULTS PER BIOMARKER

The general biomarker category in which most studies were performed on blood-based biomarkers in metastatic breast cancer, concerned whole cells in the blood (n = 181; 56.6%). CTCs made up a large part of this. In 85.1% (n = 154) of all included studies CTC enumeration was performed. In 42.5% of all included studies (n = 136), also genetic profiling for these cells had been done. The markers most frequently investigated are presented in Table 1.

Only those biomarkers for which 5 or more studies have been performed are included in the table. This cut-off had been chosen because these markers represent the most frequently investigated biomarkers. The frequency by which biomarkers are investigated is presented in Table 2, which presents that only 9.8% of all biomarkers is investigated in more than 5 studies. A detailed overview presenting the amount of studies performed for each single biomarker, including an overview of the amount of studies in each developmental stage is presented in Appendix D.

Table 1: Most frequently investigated biomarkers of all studies included

	# articles	% of included studies	End Stage	# Studies at end stage
ALDH1	5	1.6%	Observational	1
CA15-3	51	15.9%	Observational	6
CEA	19	5.9%	Observational	1
CK19	6	1.9%	Observational	1
CTC enumeration	154	48.1%	Clinical trial	29
EGFR	15	4.7%	Observational	6
ER	13	4.1%	Basic prognostic	3
HER2	61	19.1%	Observational	15
PIK3CA	13	4.1%	Observational	1
PR	7	2.2%	Basic prognostic	2
RASSF1A	6	1.9%	Basic predictive	5
THBS-1	9	2.8%	Observational	5
TP53	5	1.6%	Basic prognostic	1
TWIST	7	2.2%	Observational	2
VEGF	22	6.9%	Observational	15
VEGFR	13	4.1%	Observational	12
Vimentin	6	1.9%	Basic prognostic	1

** The abbreviations used are standard abbreviations. Corresponding gene identities encoding for these biomarkers are presented in appendix D.*

Table 2: Number of studies in which single biomarkers are investigated

Number of studies that investigated a specific biomarker	Frequency	% of all included studies
1	190	68.8%
2	38	13.8%
3	12	4.3%
4	9	3.3%
5 -10	19	6.9%
> 10	8	2.9%

The percentages shown in Table 1 present the percentage of total studies that investigated that single marker. The second general biomarker category on which a relatively large amount of studies have been performed ($n = 107$; 33.4%) are proteins. Within this category most research has been focusing on 4 proteins, which are: CA15-3 ($n = 22$; 20.5%), soluble HER2 ($n = 19$; 17.8%), Vascular Endothelial Growth Factor (VEGF) ($n = 18$; 16.8%) and Vascular Endothelial Growth Factor Receptor (VEGFR) ($n = 14$; 13.1%). As discussed before, frequently studies are focusing on investigating multiple biomarkers instead of single biomarkers. As presented in Table 1, a total of 51 studies have investigated CA15-3. This means that also research which mainly focuses on one of the other biomarker categories investigates CA15-3. The same differences in the amount of studies performed were seen for HER2 and VEGF.

RESULTS ON THE NUMBER OF STUDIES PERFORMED

Summarized over all general biomarker categories, the total amount of studies included in the results synthesis is 320 as presented in Figure 2. In these studies a total of 275 single biomarkers have been investigated. The average number of studies performed on one single biomarker is 2.6 (range 1–154 studies). In Table 2 results are presented for frequency by which the study investigated a number of biomarkers. Table 2 shows that for 68.8% of all the biomarkers only one study has investigated that particular biomarker. For 13.8% of all biomarkers two studies have investigated that biomarker.

DISCUSSION

In this paper we present a broad overview of research on blood-based biomarkers in metastatic breast cancer, performed since 2006. Of the included studies, most focused on detecting whole cells in the blood, with a focus on the enumeration or genetic characterization of circulating tumor cells. Considering the classification into developmental stages, the identification stage is the stage during which most research

has been performed. Most studies focus on the identification phase, in which they investigate the ability to detect particular biomarkers in the blood and are trying to find connections between these concentrations and potential outcome measures in terms of survival. For proteins CA15-3, soluble HER2, VEGF and VEGFR have been investigated most frequently. However, for CTCs there have been clinical trials, but not for one of these proteins since 2006.

In terms of developmental stages, we expected that the amount of research performed would follow some kind of trend over time. It was expected that per biomarker there would be a substantial amount of studies focusing on the early developmental stages (technical validation), with decreasing numbers of studies the further the research for that particular biomarker proceeded in the developmental process. However, the DEAC shows that this trend does not exist for blood-based biomarkers in metastatic breast cancer. The DEAC shows that the number of studies performed increase until they reach the identification phase, and decrease afterwards. Therefore, it seems that the technical validation and clinical validation phase are currently less performed than research in the identification phase. Another observation from the DEAC is the low amount of research performed in the prognostic validation phase, suggesting this phase is not receiving sufficient attention. However, this may well be due to the fact that the initial search was limited to articles published since 2006, so that a limited amount of studies concerning some of the developmental phases were found. It might have been that specific phases which seemed to have had insufficient attention for several biomarkers were investigated before 2006. In addition, some information might have been missed, as publication bias may have occurred due to excluding non-English studies.

Furthermore, biomarker research may have been performed in a commercial setting or for stakeholders intent to guide internal research and development decisions. As such, selective reporting may occur by which not all findings might have been published. The same holds for studies with negative findings on (some subset) of investigated biomarkers, as it is known that such results are harder to publish than positive findings. As we did not investigate a single outcome measure, no standard methods are available to assess the ensuing risk of bias in our results. Even though the intention of reports might be to inform about recent developments, other stakeholders might use this information differently. Therefore, it seems valuable for future research to be able to have access to all information that was, or can possibly be extracted from the blood samples. Future research should pay attention to selective reporting before publishing, or ensure that samples are publicly available via biobanks.

MATERIALS AND METHODS

This systematic review of blood-based biomarkers in metastatic breast cancer was performed according to the PRISMA guidelines [15]. A review protocol was used and is presented in Appendix A. This review was not registered in the PROSPERO database. All types of studies were included in the initial review, as the aim of this review is to identify the best available evidence exploring the position in the development process of all blood-based biomarkers in metastatic breast cancer. Since all types of primary research studies were included, it was not required that the intervention, control or specific outcome measure was reported in the initial search. Therefore, no specific study characteristics or PICO-statement for inclusion criteria was used. The only restriction applied to the search concerned a time constraint, as studies published since 1 January 2006 were included. Databases that were searched are PubMed, Scopus and OVID. Additionally, articles found by cross-referencing or hand search were included in the initial search. The initial search was performed in June 2016 and updated on 1 December 2016. The detailed search terms applied are presented in Appendix B.

After the initial search and removal of duplicate papers, abstracts were scanned for relevance. Abstracts of articles that either did not present non-primary research data or concerned topics not of interest here (such as, other cancer types, only other stages of breast cancer, and non-blood-based biomarkers—e.g., biomarkers that can be found in other body fluids) were excluded from the full-text review. All abstracts were processed by one reviewer (AMSB) and were discussed with a second reviewer (HK) if necessary.

Full texts of all included papers were assessed for eligibility by one reviewer (AMSB). All studies were then categorized according to the 10 pre-defined developmental stage and per general biomarker type. Four general developmental stages were identified, namely technical validation, identification, clinical validation and clinical utility. A full description of all pre-defined developmental stages is presented in Figure 1. Data was then classified in four general types of biomarkers, namely cells, proteins, circulating DNA and circulating RNA. Final classification of studies was discussed with a second reviewer (HK) if classification in either one of the categories was unclear to the first reviewer (AMSB). For studies on which there was no consensus between these two reviewers, a third reviewer reclassified the study (MIJ).

ARTICLE PROCESSING

Quantitative and qualitative data was manually extracted from the included studies and structured in Excel (version 2013) in pre-defined and labeled columns. The following information was extracted from all the included studies:

- General biomarker classification (classification in one of the four categories: cells, proteins, circulating DNA or circulating RNA)
- Developmental stage (classification according to the stages and general descriptions of these stages shown in Figure 1)
- Specific biomarker name
- Type of test used to quantify or detect biomarker (e.g., ELISA, CellSearch, etc.)
- Whether—and if so, which—survival data was presented (Overall Survival, Progression Free Survival, survival in months)

Given the focus on the translation of biomarkers to clinical practice, the results of all included studies were summarized according to the number of studies performed per developmental stage for each general biomarker category. Results for all single biomarkers were summarized per general biomarker category as studies might investigate more than one biomarker. For all single biomarkers it was determined how many studies investigated that biomarker and in which stage of translation the biomarker was identified. For each general biomarker category it was investigated how many studies presented results on the full range of single biomarkers found.

SYNTHESIS OF RESULTS

Results were presented in a Development, Evaluation and Application Chart (DEAC) that was developed specifically for this review. This figure gives a broad overview of the development of biomarkers in each of the predefined stages of clinical translation. Specifically, the figure shows four bar diagrams above each other, one diagram for each of the general biomarker categories. Each vertical bar, per diagram, represents a developmental stage. The bars are displayed to represent the different stages in the translation, starting with the most basic (developmental) research on the left side and more advanced (evaluation) research (such as clinical trials or health economic evaluations) presented on the right side. The height of the bars reflects the number of included studies. This figure therefore gives an overview of the number of studies published on each of the general biomarker categories according to the developmental stage timeline.

CONCLUSIONS

Since 2006, a substantial amount of research has been done to investigate the potential role of blood-based biomarkers in metastatic breast cancer. There seems to be a focus on research toward the use of CTCs, as most studies investigate these, whether in combination with other markers or as a single marker. The current emphasis of investigating these biomarkers seems to be on developing new techniques or finding new biomarkers that might have predictive or prognostic value, as most studies focus on the identification phase. There is a lack of studies focusing on clinical utility of these biomarkers. This might be because these studies have not yet been performed or suffer from publication bias. However, the lack of studies investigating the utility of blood-based biomarkers causes the additional value in terms of clinical utility, health outcomes or health care efficiency to still be limited according to the investigated evidence.

DECLARATIONS

Abbreviations

CA15-3	Cancer Antigen 15-3
CTC	Circulating Tumor Cell
DEAC	Development, Evaluation and Application Chart
PFS	Progression Free Survival
HER2	Human Epidermal Growth Factor Receptor 2
OS	Overall Survival
VEGF	Vascular Endothelial Growth Factor
VEGFR	Vascular Endothelial Growth Factor Receptor

Author Contributions: The initial search was performed by A. M. Sofie Berghuis. Classification into developmental stages of the articles was discussed with Hendrik Koffijberg and Maarten J. IJzerman. Classification of the general biomarker categories was discussed with Jai Prakash and Leon W. M. M. Terstappen. All authors were involved in writing and revising the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015 2015;65(1):5-29.
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* [Internet]. 2014 Sep 13 [cited 2014 Sep 25];386. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25220842>
3. Ufen M-P, Köhne CH, Wischneswky M, Wolters R, Novopashenny I, Fischer J, et al. Metastatic breast cancer: are we treating the same patients as in the past? *Ann Oncol* [Internet]. 2014 Jan [cited 2015 Feb 11];25(1):95–100. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24276026>
4. Zervoudis S, Iatrakis G, Tomara E, Bothou A, Papadopoulos G, Tsakiris G. Main controversies in breast cancer. *World J Clin Oncol* [Internet]. 2014 Aug 10 [cited 2015 Feb 12];5(3):359–73. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4127607&tool=pmcentrez&rendertype=abstract>
5. Marino N, Woditschka S, Reed LT, Nakayama J, Mayer M, Wetzel M, et al. Breast cancer metastasis: issues for the personalization of its prevention and treatment. *Am J Pathol* [Internet]. American Society for Investigative Pathology; 2013 Oct [cited 2015 Feb 11];183(4):1084–95. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3791679&tool=pmcentrez&rendertype=abstract>
6. Gaeta CM, Sher AC, Kohan A, Rubbert C, Avril N. Recurrent and metastatic breast cancer PET, PET/CT, PET/MRI: FDG and new biomarkers. 2013;(December).
7. NABON. Breast Cancer Guideline. 2012;
8. Eccles S a, Aboagye EO, Ali S, Anderson AS, Armes J, Berditchevski F, et al. Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Res* [Internet]. 2013 Jan [cited 2014 Jul 13];15(5):R92. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3907091&tool=pmcentrez&rendertype=abstract>
9. Ravelli A, Reuben JM, Lanza F, Anfossi S, Cappelletti MR, Zanotti L, et al. Breast cancer circulating biomarkers : advantages , drawbacks , and new insights. 2015;6653–65.
10. Lewandowska MA. Prognostic and Predictive Biomarkers : Tools in Personalized Oncology. 2014;273–84.
11. Banys-paluchowski M, Krawczyk N, Meier-stiegen F. Critical Reviews in Oncology / Hematology Circulating tumor cells in breast cancer — current status and perspectives. Elsevier Ireland Ltd; 2016;97:22–9.
12. Nalejska E, Mączyńska E, Lewandowska MA. Prognostic and predictive biomarkers: tools in personalized oncology. *Mol Diagn Ther* [Internet]. 2014 Jun [cited 2015 Jan 8];18(3):273–84. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4031398&tool=pmcentrez&rendertype=abstract>
13. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell* [Internet]. Elsevier Inc.; 2011;144(5):646–74. Available from: <http://dx.doi.org/10.1016/j.cell.2011.02.013>
14. Janni WJ, Rack B, Terstappen LWMM, Pierga J, Taran F, Fehm T, et al. Pooled Analysis of the Prognostic

Relevance of Circulating Tumor Cells in Primary Breast Cancer. 2016;1–12.

15. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma T. Preferred reporting items for systematic reviews and meta-analyses : The PRISMA statement. 2010;8:336–41.

16. Coumans FAW, van Dalum G, Beck M, Terstappen LWM. Filter Characteristics Influencing Circulating Tumor Cell Enrichment from Whole Blood. PLoS One. 2013;8(4).

17. Lang JE, Scott JH, Wolf DM, Novak P, Punj V, Magbanua MJM, et al. Expression profiling of circulating tumor cells in metastatic breast cancer. Breast Cancer Res Treat. 2015;149(1):121–31.

18. Strati A, Kasimir-Bauer S, Markou A, Parisi C, Lianidou ES. Comparison of three molecular assays for the detection and molecular characterization of circulating tumor cells in breast cancer. Breast Cancer Res [Internet]. BioMed Central Ltd; 2013;15(2):R20. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3672668&tool=pmcentrez&rendertype=abstract>

19. Santos MF, Mannam VKR, Craft BS, Punecky L V., Sheehan NT, Lewis RE, et al. Comparative analysis of innate immune system function in metastatic breast, colorectal, and prostate cancer patients with circulating tumor cells. Exp Mol Pathol [Internet]. Elsevier B.V.; 2014;96(3):367–74. Available from: <http://dx.doi.org/10.1016/j.yexmp.2014.04.001>

20. Ramirez J-M, Fehm T, Orsini M, Cayrefourcq L, Maudelonde T, Pantel K, et al. Prognostic relevance of viable circulating tumor cells detected by EPISPOT in metastatic breast cancer patients. Clin Chem [Internet]. 2014;60(1):214–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24255082>

21. Bidard F, Peeters D, Peeters D, Fehm T, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. Lancet Oncol [Internet]. 2014;15(April):1–9. Available from: [http://dx.doi.org/10.1016/S1470-2045\(14\)70069-5](http://dx.doi.org/10.1016/S1470-2045(14)70069-5)

22. Kallergi G, Agelaki S, Papadaki MA, Nasias D, Matikas A, Mavroudis D, et al. Expression of truncated human epidermal growth factor receptor 2 on circulating tumor cells of breast cancer patients. Breast Cancer Res. Breast Cancer Research; 2015;17:113. Available from: <http://dx.doi.org/10.1186/s13058-015-0624-x>

23. Smerage JB, Barlow WE, Hortobagyi GN, Winer EP, Leyland-Jones B, Srkalovic G, et al. Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. J Clin Oncol [Internet]. 2014 Nov 1 [cited 2015 Mar 19];32(31):3483–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24888818>

CHAPTER 6

HEALTH ECONOMIC IMPACT OF LIQUID BIOPSIES IN CANCER MANAGEMENT

M.J. IJZERMAN, A.M.S. BERGHUIS,
J.S. DE BONO, L.W.M.M. TERSTAPPEN

THIS CHAPTER WAS PUBLISHED AS:
M.J. IJZERMAN, A.M.S. BERGHUIS, J.S. DE BONO, L.W.M.M. TERSTAPPEN. HEALTH ECONOMIC IMPACT OF LIQUID BIOPSIES IN CANCER MANAGEMENT. EXPERT REVIEW OF PHARMACOECONOMICS & OUTCOMES RESEARCH, 07-08-2018. DOI: 10.1080/14737167.2018.1505505

ABSTRACT

INTRODUCTION

Liquid biopsies (LBs) are referred to as the sampling and analysis of non-solid tissue, primarily blood, as a diagnostic and monitoring tool for cancer. Because LBs are largely non-invasive, they are a cheap alternative for serial analysis of tumor progression and heterogeneity to facilitate clinical management. Although a variety of tumor markers are proposed (e.g. free-circulating DNA), the clinical evidence for Circulating Tumor Cells (CTCs) is currently the most developed.

AREAS COVERED

This paper presents a health economic perspective of LBs in cancer management. We first briefly introduce the requirements in biomarker development and validation, illustrated for CTC enumeration. Second, we discuss the state-of-art on the clinical utility of LBs in breast cancer in more detail. We conclude with a general outlook on clinical use and reimbursement of LBs

EXPERT COMMENTARY

A significant increase in clinical research on LBs can be observed and suggest a promising future and a rapid change of cancer management. In addition to studies evaluating clinical utility of LBs, a smooth translation into clinical practice requires systematic assessment of the health economic benefits. This paper argues that (early stage) health economic research is required to facilitate its clinical use and to prioritize further evidence development.

INTRODUCTION

There is a growing consensus that the cost of cancer care is exploding, particularly due to new targeted agents. While the total cost of cancer is between 4-6 % of the healthcare spending, cancer drugs account for approximately one third of the costs and this spending is expected to grow even further (1). Several initiatives aim at controlling drug prices directly, such as the negotiation of drug purchasing rebates based on the results of cost-effectiveness studies and the demand for price transparency (2) (3). Another mechanism to provide affordable cancer care, is to effectively prescribe drugs guided by value frameworks such as the ESMO magnitude of clinical benefit scale (MCBS) (4) and by better targeting of expensive medicines using biomarkers and genomic profiling (5).

Staging and targeting of treatment is currently done by combining information from tissue biopsies and the use of imaging modalities or combinations of those. However, an emerging and promising technology is the use of liquid biopsies (LBs), which allow for the early detection of cancer (6) and/or the analysis of tumor progression and profiling, by sampling and isolating tumor markers from body fluids, mainly blood (7). Several traditional blood tumor markers proteins, such CA15-3, CA19-9 and PSA markers, are used in the management of respectively breast, colorectal and prostate cancer, but with the advances of new detection technologies the molecular composition of the cancer can be studied through circulating tumor DNA (ctDNA), tumor derived extra-cellular vesicles (td-EV), micro-RNA (miRNA) and circulating tumor cells (CTCs). Although substantial effort is invested in the isolation and validation of different tumor markers (ctDNA, td-EV, miRNA) from blood, CTCs are most likely to mirror the tumor and represent a true surrogate for a tissue biopsy. They, however, need to be present, isolated and characterized for this to be materialized. At this time CTC are still the most frequently studied LBs in clinical trials and we will therefore mainly focus on the utility and potential of CTCs to change patient management (8).

The presence of tumor cells in blood was first observed in 1869 (9), a validated assay to reliable measure CTC was introduced in 2004 (10) and results from multicenter prospective studies in metastatic breast, prostate showing their prognostic validity and ability to predict response to therapy were reported from 2004 to 2008 (11-13). The FDA cleared Cellsearch® is based on immunomagnetic enrichment of cells expressing EpCAM and the assignment of CTC based on expression of Cytokeratins and lack of CD45. Since the introduction of CellSearch many studies have confirmed the early findings (14,15) have expanded it to other cancers (16,17), showed promise

in early disease settings (18,19) and moved efforts to the characterization of the CTC by showing its potential for personalized medicine by detection of diagnosing gene aberrations or amplifications (20).

In particular, the non-invasiveness of sampling LBs (7.5 mL blood sample) makes this technology very useful for frequent therapy monitoring and for deciphering tumor heterogeneity and mechanisms of drug resistance. Particularly in tumors where it is hard to take biopsies from the primary tumor, such as lung-cancer, LBs have substantial potential (21).

BIOMARKER DEVELOPMENT AND VALIDATION: THE CLINICAL USE OF CTCs IN CANCER MANAGEMENT

Biomarker development follows a known pathway, from technical and clinical validation to the assessment of clinical utility. Such distinction is also relevant for LBs, and in particular when the health economic benefits are to be discussed (22). Khoury et al present an adapted model for evidence generation distinguishing between four types of evaluation domains (23). The assessment of technical validity of an assay includes the estimation of common metrics such as sensitivity and specificity, and analyzes the performance of the assay against a known test standard. Clinical validity then correlates the test findings to a disease process or disease state, for instance in diagnostic or prognostic studies. Third, clinical utility broadly assesses whether the implementation of a test will change health outcomes, whether the process of delivering those outcomes improves as well as the resources utilized to achieve these health outcomes can be reduce. Finally, ethical and legal considerations of diagnostic test use are considered. Broadly speaking, health economists are mainly interested in the clinical utility of diagnostic testing, the added value for society and thus implicitly the additional life-years of QALYs gained. However, in addition to changes in health outcomes or changes in the care process to deliver those outcomes, a discussion amongst health economists revealed that other value components from diagnostic testing may also be relevant for society such as planning value, sometimes referred to as value-of-knowing, and psychic value (24).

To follow the biomarker development pathway, multiple clinical trials were performed to address the clinical validity or prognostic value of CTCs. Different studies showed that the presence of >5 CTCs was shown to have prognostic value in metastatic breast cancer (mBC) and metastatic castration resistant prostate cancer (mCRPC) whereas the presence of >3 CTCs was found to be prognostic in colorectal cancer in terms of overall survival before initiation of a new line of therapy and could predict

the response to therapy already after the first cycles of therapy (25). In addition to several prognostic studies, also the predictive value of CTC enumeration has been investigated to guide decisions on the start of systemic therapy. For instance, in a phase II study, Krebs *et al* found that colorectal cancer patients with elevated baseline CTC counts (≥ 3 CTCs/7.5 ml) could benefit from an intensive chemotherapy regimen, unlike patients with low CTC counts, which can be concluded to be one of the first studies showing predictive validity of CTCs (26). The interpretation of CTC results is still subject of study, and the use of cut-offs (less or more than 5 CTC) is an oversimplification, and complicates the evaluation of change in CTC as a response marker as the number of CTCs are low (< 10) anyway in the majority of cases (27)(28).

Evaluation of CTCs as a response marker was evaluated by a 30% decline of CTCs (initial 5 CTCs/7.5 mL) and was associated with treatment response in mCRPC as early as three weeks within the start of therapy (29). This was confirmed in a recent systematic review presenting the evidence for use of CTC number as a response measure in metastatic prostate cancer. They defined several measures of response based on CTC counts at baseline and at 13 weeks, and concluded that CTC0 (change from > 1 CTC at baseline to 0 CTC at 13 weeks) and CTC conversion (change from > 5 CTC at baseline to < 4 CTC at week 13) had the highest discriminative power for overall survival (30).

These findings imply that CTCs potentially allow the determination of treatment response 7-8 weeks earlier than using standard RECIST criteria and Prostate Specific Antigen (PSA) with a substantial cost-saving and prevention of toxicity. Although the clinical evidence for this hypothesis is subject of the CTC-STOP trial, an early stage simulation model populated with existing clinical trial data confirmed that CTC enumeration may guide early discontinuation and treatment switching reducing cost of over-treatment (31).

Another example of the predictive validity of CTCs, is the expression of the AR-V7 splice variant on CTCs. Recent studies have shown that the expression of the AR-V7 splice variant on CTCs impacts clinical management because they may guide the selection of either hormonal treatment (AHT, Abiraterone or Enzalutamide both around 28,000 US\$ per patient 2) or chemotherapy in naïve patients (e.g. Docetaxel at 9,000 US\$ or Cabazitaxel at 65,000 US\$ per patient). This stratification may result in substantial cost savings in patients with AHT drug resistance (32). Although these examples provide preliminary evidence of health economic benefits, they also require further research on their potential to gain more efficient use of resources.

CLINICAL VALIDATION AND THE CLINICAL UTILITY OF LIQUID BIOPSIES IN BREAST CANCER

We now focus on (metastatic) breast cancer, where a recent paper was published reviewing the use of LBs in breast cancer (BC), mostly on the enumeration of tumor derived particles or the characterization of the most common gene specific abnormalities (33).

In primary breast cancer, several studies provide evidence on the prognostic value of LBs. The German SUCCESS trial showed the independent prognostic relevance of CTCs before and after adjuvant chemotherapy (34). Also, a recently published pooled analysis of 3,173 patients confirmed that the presence of CTCs was an independent predictor of poor disease-free survival, overall survival, and BC specific survival (18).

The first study in metastatic breast cancer report on 177 mBC patients in which CTC were enumerated before, and at 4 discrete time-intervals after the initiation of therapy (35-37).

From this study it was concluded that CTC detection before the start of treatment is highly predictive for overall survival and may be used for patient stratification (36). Also repeated assessment of CTCs in mBC show that elevated levels of CTCs over time are associated with progression-free and overall survival (35). These findings were confirmed in a pooled analysis from data generated from 1944 patients included in 17 different European sites (37). The same study also showed that serum tumor markers did not add any significance to the model.

6

Another key question to address is whether liquid biopsies add value compared to routine clinical assessment of progression using CT imaging. Budd et al performed CT scans in 138 metastatic Breast Cancer patients before and after initiation of therapy, and compared the results with CTC counts obtained at baseline and 4 weeks after the start of therapy. Both, CT imaging and CTCs are then used to determine progression. They found that there is concordance between radiology and CTCs in about 76% of the patients after four weeks in the assessment of either stable disease/partial response or progressive disease. In addition, they also correlated radiology and CTC assessment at the first follow up with overall survival for all patients and for patients receiving either first or second line systemic therapy. Reported hazard ratios are 4.08 for prediction of response to first-line therapy using CTCs and 5.37 for radiology. Hazard ratios for prediction of response for second line therapy are 2.79 and 1.44 for CTCs and radiology respectively. From there results, Budd et al conclude that CTCs are more reproducible than radiologic response, and that CTCs are an early

and reliable indication of disease status and suggest CTCs are a superior surrogate endpoint (38).

As discussed in the previous section on biomarker development and validation, the ability of LBs to change clinical management ultimately is required to being used in clinical practice. This implies that new prospective randomized controlled trials have to be designed that include a therapeutic decision based on the assessment of CTCs with known predictive validity. Several studies that do evaluate clinical utility are in the final stages of data collection and analysis and results will be available soon (39).

One of the first studies evaluating if CTCs do change clinical management and outcomes, was the SWOG S0500 trial. The SWOG0500 trial was a randomized controlled design, including patients with mBC after previous lines of treatment such as with hormone therapy or selected targeted agents (40). CTC counts were available from included patients, and patients initially having >5 CTCs started a first line of chemotherapy underwent a repeated follow-up examination of CTCs at 22 days. Patients initially having <5 CTCs were treated at the physician's discretion, but were not followed for CTC changes. Patients persistently having >5 CTCs at the follow-up test at 22 days, were then randomized to either maintain therapy or change therapy. The SWOG0500 trial has not been able to find a difference in survival for patients stratified based on CTC counts. The SWOG0500 has been criticized because the negative findings were merely a failure of the study design than a failure of the CTC test (41). In particular, patients included in the SWOG0500 had advanced disease and were not likely to have any therapeutic benefit at all at the time of inclusion.

While the SWOG0500 actually guided the clinical decision to switch or maintain therapy by randomizing a subgroup of patients based on CTCs, clinical trials with such LB guided treatment decisions are still scarce but essential to establish real-world clinical utility. An example is the DETECT trial, which aims for improvements in survival by phenotyping CTCs to guide targeted therapy (42). The DETECT trials have included the largest mBC population that is currently investigated. More important, the therapeutic decision is left by the discretion of the physician, based on the characterization of discordance between the primary tumor, metastases and LBs. The premise of the DETECT IV is that patients with a HER2-negative primary tumor can develop HER2-positive CTC during disease progression, requiring a different therapeutic approach. Such discordance has been shown in about 19% of patients with metastatic breast cancer and does impact clinical management (43). Dagogo-Jack and Shaw only recently published a review showing both spatial (between solid

tumor and metastatic sites) and temporal tumor heterogeneity (tumor dynamics over time) and others have suggested liquid biopsies may be the appropriate platform for determining the full extent of tumor heterogeneity (44) and also help identifying new molecular targets in case of drug resistance (45).

Finally, the French STIC trial is designed to address the potential of using CTC counts to prescribe either hormonal or chemotherapy as first-line treatment in hormone-receptor positive mBC (39). The hypothesis of the trial is that the CTC based approach is non-inferior to the physicians' choice in terms of survival, while a reduction of intensive chemotherapy regimens may be possible by prescribing chemotherapy only to those patients with high CTC counts. Although the final results are to be published, a presentation at ESMO (2016) of the inclusion and randomization results in the STIC trial suggests that treatment decisions changed in 38% of the patients based on CTCs. Final results of the trial will be available in 2018.

POTENTIAL HEALTH ECONOMIC BENEFITS OF LIQUID BIOPSIES IN CANCER MANAGEMENT

Personalized medicine undoubtedly creates health economic potential as treatment decisions can be better targeted, thereby avoiding overtreatment and improving health outcomes. However, several authors also identified the many challenges associated with health economic modeling of personalized medicine applications, and this is no different for LBs, where clinical utility remains an essential requirement. Utility in this respect refers to the ability of a diagnostic test to change clinical management and/or to improve clinical outcomes (46). This definition also evokes another question, i.e. how to distinguish between predictive validity and utility. In health economics, the utility of a test refers to the benefits and risks of (not) taking the test anticipating future health consequences (47). But before a test be considered for its clinical utility, it should at least demonstrate predictive validity. I.e. the test should be able to distinguish a favorable from a non-favorable outcome if treatment commences. But more importantly, to evaluate clinical utility in the real-world, clinical decisions should be made according to a test result and deviations from that (for whatever reason) will directly influence the utility of a test. In other words, if a specific biomarker assay or LB indicate either non-response or therapy resistance the actual utility of a test only can be shown if clinical decisions are made to either discontinue or switch treatment. This physician attitude or preference towards the use of LBs is often neglected, yet plays a key-role in translating LBs to clinical practice.

There are several potential applications of LBs with an opportunity to improve clinical outcomes and control healthcare costs, in addition to those applications where it is

simply not possible to implement solid tissue biopsies. Screening and early detection of cancer have been suggested, e.g. by detecting hypermethylated DNA in urine or blood (48) or circulating proteins and cell-free DNA (6). Unfortunately, clinical research on LBs for screening is scarce and there is not a lot of evidence for use in screening to date. Instead, at least three other potentially valuable applications are hypothesized impacting clinical management (49), i.e. 1: improved tumor staging by providing evidence on micro-metastasis and thereby guiding systemic therapy, 2: response monitoring, e.g. a change in CTC count as an early sign of progression or stable disease and 3: treatment targeting by determining a complete profile of targets for targeted drugs, tumor heterogeneity and mechanisms of resistance following single-cell sequencing and genomic profiling. In particular, the value of LBs for single-cell or ctDNA sequencing to determine tumor heterogeneity is challenging and promising but requires further (health economic) research.

CHALLENGES WITH THE REIMBURSEMENT OF LIQUID BIOPSIES

While the evidence for using LBs in cancer management is emerging, more clinical research and more evidence on health economic benefits is required. As illustrated, several studies are ongoing to evaluate the potential of using LBs in clinical practice in an attempt to address whether additional diagnostic information from LBs changes clinical management and thereby health outcomes. Such evidence is also essential to populate and update health economic decision models, as currently existing basic predictive models still have substantial uncertainty in estimating the health economic potential of LBs.

However, while clinical trials are ongoing, health economic modelling may already support the prioritization and design of LBs trials. Early stage health economic models are proposed to prioritize further development and to identify biomarkers that potentially can make it to the clinic (50) (51). Early stage models for LBs are developed for mCRCP, supporting an early switch to Cabazitaxel following Docetaxel treatment (31), and useful to identify model parameters that do change outcome. Furthermore, early models may help distinguishing between and prioritizing different types of LBs, e.g. CTCs, ct-DNA or td-EV.

The health economic evaluation of LBs is not a panacea for its ultimate use. In order to implement molecular (companion) diagnostic testing in clinical practice, several challenges have to be faced (52). First, implementing a (companion) diagnostic to stratify patients after full approval and reimbursement of a cancer medicine is challenging, because a companion diagnostic may influence the market share of a

pharmaceutical (53-55). Second, evidence development for diagnostics is difficult as multiple combinations of test outcomes cannot be easily investigated in a clinical trial (56,57). According to Towse and Garrison, “we need to be conscious about the difficulty of developing the evidence base for diagnostics in personalized medicine”. And, although this certainly applies to LBs, evidence development for biomarkers is difficult in general. There have been many biomarkers that have been published and claimed to be therapeutically useful, but few become part of the clinical decision-making process due to technical, validation and market access issues (57). Schneider *et al* request policy makers to be aware of this implementation barrier, and urge for defining specific requirements for the introduction of biomarkers (57).

EXPERT COMMENTARY

A significant increase in clinical research on liquid biopsies, covering all aspects from isolation to validation of tumor markers in blood, can be observed in the last decade. Also, several expert commentaries in medical oncology journals (58) (59-61) as well as the inclusion of liquid biopsies in the Cleveland top-10 medical innovations (62), suggest a promising future and a rapid change of cancer management using liquid biopsies, and ct-DNA in particular. Nevertheless, the actual clinical use of liquid biopsies is only changing slowly. Part of this can be explained by the known challenges of bringing personalized medicine to the clinic.

6

Awareness amongst medical oncologists about increasing drug prices and associated financial toxicity has also brought a new perspective for the further development of liquid biopsies as they may improve cost-effective drug prescription. For this to happen, health economists may find ways to share expertise in the process of early discovery, clinical validation and the (early) assessment of health economic benefits (63). In particular, health economists’ expertise is warranted in the evaluation of the clinical and economic consequences of implementing LBs using (early stage) decision models, the prioritization of LBs evidence development using a value-of-information framework, and a more detailed analysis of the cost of implementing biomarker assays and LBs. Previous authors have already identified the complexity of health economic modelling in personalized medicine (64) (65), and this certainly applies to the use of LBs with the complexity and uncertainties associated with the serial assessment of LBs, and the large amount of targeted agents and immunotherapy used as 1st, 2nd or 3rd line therapies in different sequences.

FIVE-YEAR VIEW

As illustrated, the use of LBs is emerging and basic scientists and clinical researchers have put considerable effort in the identification and analysis to further improve cancer management. Liquid biopsies may change cancer management substantially, allowing a fully personalized treatment strategy based on the molecular characterization of LBs. While most of the current studies focus at the identification and technical and clinical validation of LBs, there is a need for health economists to be involved and to identify and prioritize further development of LBs and to smoothen the path to implementation and reimbursement of LBs in the management of cancer patients. This will maximize efforts, specify where LBs do have additional impact and ultimately guarantee adoption in the healthcare system within the next 5 years.

In this view, we do see the potential of LBs for clinical management and encourage the design of clinical trials providing the evidence on clinical utility required for trial based economic evaluations to characterize therapeutic potential and health economic benefits of using LBs in clinical management. But, we also emphasize that the clinical utility of LBs is not only about the predictive validity of a test. Physician attitudes or preferences towards the use of LBs for patient management plays a key-role in the process of translating LBs into clinical practice, and thus in the attempts to achieve clinical and societal benefits.

KEY ISSUES

A significant research effort is invested in the isolation, molecular characterization and validation of various blood-based tumor markers, known as liquid biopsies.

Clinical research in the last 2 decades have shown prognostic and predictive value of CTCs to determine response and overall survival in breast cancer, prostate cancer and colorectal cancer.

Studies exploring clinical utility of liquid biopsies to change cancer management are underway and results are expected to support implementation and reimbursement of liquid biopsies

Liquid biopsies have health economic potential if used to serially monitor treatment response to inform decisions to discontinue inactive treatment, or to switch treatment to agents that target other molecular mechanisms in case of resistance.

DISCLOSURES

Financial disclosures: The authors declare no conflicts of interest and/or financial support

Keywords: Liquid biopsies, cancer, circulating tumor cells, ctDNA, health economics, biomarker, diagnostic

REFERENCES

* Of interest

** Of considerable interest

1. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *The Lancet Oncology*. 2013 Nov;**14**(12):1165–74.

2. Danzon P, Towse A, Mestre-Ferrandiz J. Value-based differential pricing: efficient prices for drugs in a global context. *Health Econ*. 2015 Mar;**24**(3):294–301.

3. van Harten W, IJzerman MJ. Responsible pricing in value-based assessment of cancer drugs: real-world data are an inevitable addition to select meaningful new cancer treatments. *Ecanermedicalscience*. 2017;**11**:ed71.

4. Del Paggio JC, Azariah B, Sullivan R, Hopman WM, James FV, Roshni S, et al. Do Contemporary Randomized Controlled Trials Meet ESMO Thresholds for Meaningful Clinical Benefit? *Annals of Oncology*. 2017 Jan **1**;**28**(1):157–62.

5. Phillips KA, Ann Sakowski J, Trosman J, Douglas MP, Liang S-Y, Neumann P. The economic value of personalized medicine tests: What we know and what we need to know. *Genetics in Medicine*. Nature Publishing Group; 2014 Jan **1**;**16**(3):251–7.

6. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L,

et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science. American Association for the Advancement of Science*; 2018 Jan **18**;**1**:eaar3247–10.

7. Bardelli A, Pantel K. Liquid Biopsies, What We Do Not Know (Yet). *Cancer Cell*. 2017 Feb **13**;**31**(2):172–9.

8. Onstenk W, Gratama JW, Foekens JA, Sleijfer S. Towards a personalized breast cancer treatment approach guided by circulating tumor cell (CTC) characteristics. *Cancer Treat Rev*. 2013 Nov;**39**(7):691–700.

9. Ashworth TR. A Case of Cancer in Which Cells Similar to Those in the Tumors Were Seen in the Blood after Death. *Australasian Medical Journal*. **14**:146–9.

10. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin Cancer Res. American Association for Cancer Research*; 2004 Oct **15**;**10**(20):6897–904.

11. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med. Massachusetts Medical Society*; 2004 Aug **19**;**351**(8):781–91.

12. Cohen SJ, Punt CJA, Iannotti N, Saidman BH, Sabbath KD, Gabrail NY, et al. Relationship of circulating tumor cells to tumor response, progression-free survival, and

overall survival in patients with metastatic colorectal cancer. *J Clin Oncol*. 2008 Jul 1;26(19):3213–21.

13. de Bono JS, Scher HI, Montgomery RB, Parker C, Miller MC, Tissing H, et al. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clin Cancer Res. American Association for Cancer Research*; 2008 Oct 1;14(19):6302–9.

14. Lv Q, Gong L, Zhang T, Ye J, Chai L, Ni C, et al. Prognostic value of circulating tumor cells in metastatic breast cancer: a systemic review and meta-analysis. *Clin Transl Oncol. Springer Milan*; 2016 Mar;18(3):322–30.

15. Goldkorn A, Ely B, Quinn DI, Tangen CM, Fink LM, Xu T, et al. Circulating Tumor Cell Counts Are Prognostic of Overall Survival in SWOG S0421: A Phase III Trial of Docetaxel With or Without Atrasentan for Metastatic Castration-Resistant Prostate Cancer. *Journal of Clinical Oncology*. 2014 Mar 10;32(11):1136–42.

16. Hardingham JE, Grover P, Winter M, Hewett PJ, Price TJ, Thierry B. Detection and Clinical Significance of Circulating Tumor Cells in Colorectal Cancer--20 Years of Progress. *Mol Med*. 2015 Oct 27;21 Suppl 1(Suppl 1):S25–31.

17. Syrigos K, Fiste O, Charpidou A, Grapsa D. Circulating Tumor Cells Count as a Predictor of Survival in Lung Cancer. *Critical Reviews in Oncology/Hematology*. 2018 Mar 7.

18. Janni WJ, Rack B, Terstappen LWMM, Pierga J-Y, Taran F-A, Fehm T, et al. Pooled Analysis of the Prognostic Relevance of Circulating Tumor Cells in Primary Breast Cancer. *Clin Cancer Res*. 2016 May 15;22(10):2583–93.

19. van Dalum G, Stam G-J, Scholten LFA, Mastboom WJB, Vermes I, Tibbe AGJ, et al. Importance of circulating tumor cells in newly diagnosed colorectal cancer. *Int J*

Oncol. Spandidos Publications; 2015 Mar;46(3):1361–8.

20. Lianidou ES. Gene expression profiling and DNA methylation analyses of CTCs. Pantel K, Alix-Panabières C, editors. *Mol Oncol*. 2016 Mar 1;10(3):431–42.

21. Bernabé R, Hickson N, Wallace A, Blackhall FH. What do we need to make circulating tumour DNA (ctDNA) a routine diagnostic test in lung cancer? *Eur J Cancer*. 2017 Jun 10;81:66–73.

22. Thariani R, Veenstra DL, Carlson JJ, Garrison LP, Ramsey S. Paying for personalized care: Cancer biomarkers and comparative effectiveness. *Mol Oncol*. 2012 Apr;6(2):260–6.

23. Khoury MJ, Berg A, Coates R, Evans J, Teutsch SM, Bradley LA. The Evidence Dilemma In Genomic Medicine. *Health affairs*. 2008 Nov 1;27(6):1600–11.

24. Lee DW, Neumann PJ, Rizzo JA. Understanding the medical and nonmedical value of diagnostic testing. *Value Health*. 2010 Mar;13(2):310–4.

25. Miller MC, Doyle GV, Terstappen LWMM. Significance of Circulating Tumor Cells Detected by the CellSearch System in Patients with Metastatic Breast Colorectal and Prostate Cancer. *J Oncol*. 2010;2010:617421.

26. Krebs MG, Renehan AG, Backen A, Gollins S, Chau I, Hasan J, et al. Circulating Tumor Cell Enumeration in a Phase II Trial of a Four-Drug Regimen in Advanced Colorectal Cancer. *Clin Colorectal Cancer*. 2015 Jun;14(2):115–22.e1–2.

27. Coumans FAW, Ligthart ST, Uhr JW, Terstappen LWMM. Challenges in the enumeration and phenotyping of CTC. *Clin Cancer Res. American Association for Cancer Research*; 2012 Oct 15;18(20):5711–8.

28. Coumans FAW, Ligthart ST, Terstappen LWMM. Interpretation of Changes in Circulating Tumor Cell Counts. *Transl Oncol*. 2012 Dec 1;5(6):486–IN4.
29. Lorente D, Olmos D, Mateo J, Bianchini D, Seed G, Fleisher M, et al. Decline in Circulating Tumor Cell Count and Treatment Outcome in Advanced Prostate Cancer. *Eur Urol*. 2016 Dec;70(6):985–92.
30. Heller G, McCormack R, Kheoh T, Molina A, Smith MR, Dreicer R, et al. Circulating Tumor Cell Number as a Response Measure of Prolonged Survival for Metastatic Castration-Resistant Prostate Cancer: A Comparison With Prostate-Specific Antigen Across Five Randomized Phase III Clinical Trials. *J Clin Oncol*. American Society of Clinical Oncology; 2017 Dec 22;36(6):JCO2017752998–580.
31. Degeling K, Schivo S, Mehra N, Koffijberg H, Langerak R, de Bono JS, et al. Comparison of Timed Automata with Discrete Event Simulation for Modeling of Biomarker-Based Treatment Decisions: An Illustration for Metastatic Castration-Resistant Prostate Cancer. *Value in Health*. 2017 Jul;in press(10):1411–9.
32. Markowski MC, Frick KD, Eshleman JR, Luo J, Antonarakis ES. Cost-Savings Analysis of AR-V7 Testing in Patients With Metastatic Castration-Resistant Prostate Cancer Eligible for Treatment With Abiraterone or Enzalutamide. *Prostate*. 2016 Dec;76(16):1484–90.
33. Berghuis AMS, Koffijberg H, Prakash J, Terstappen LWMM, IJzerman MJ. Detecting Blood-Based Biomarkers in Metastatic Breast Cancer: A Systematic Review of Their Current Status and Clinical Utility. *Int J Mol Sci*. 2017 Feb 9;18(2).
34. Rack B, Schindlbeck C, Jückstock J, Andergassen U, Hepp P, Zwingers T, et al. Circulating tumor cells predict survival in early average-to-high risk breast cancer patients. *JNCI Journal of the National Cancer Institute*. 2014 May 15;106(5):793.
35. Hayes DF, Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Miller MC, et al. Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival. *Clin Cancer Res*. 2006 Jul 15;12(14 Pt 1):4218–24.
36. Cristofanilli M, Hayes DF, Budd GT, Ellis MJ, Stopeck A, Reuben JM, et al. Circulating tumor cells: A novel prognostic factor for newly diagnosed metastatic breast cancer. *Journal of Clinical Oncology*. 2005 Dec 1;23(7):1420–30.
37. Bidard F-C, Peeters DJ, Fehm T, Nolé F, Gisbert-Criado R, Mavroudis D, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *The Lancet Oncology*. 2014 Apr;15(4):406–14.
38. Budd GT, Cristofanilli M, Ellis MJ, Stopeck A, Borden E, Miller MC, et al. Circulating tumor cells versus imaging--predicting overall survival in metastatic breast cancer. *Clin Cancer Res*. 2006 Nov 1;12(21):6403–9.
39. Bidard F-C, Fehm T, Ignatiadis M, Smerage JB, Alix-Panabières C, Janni W, et al. Clinical application of circulating tumor cells in breast cancer: overview of the current interventional trials. *Cancer Metastasis Rev*. 2013 Jun;32(1-2):179–88.
40. Smerage JB, Barlow WE, Hortobagyi GN, Winer EP, Leyland-Jones B, Srkalovic G, et al. Circulating Tumor Cells and Response to Chemotherapy in Metastatic Breast Cancer: SWOG S0500. *Journal of Clinical Oncology*. 2014 Oct 13;32(31):3483–90.
41. Bidard F-C, Pierga J-Y. Clinical Utility of Circulating Tumor Cells in Metastatic Breast Cancer. *Journal of Clinical Oncology*. 2015 May 10;33(14):1622–2.

42. Schramm A, Friedl TWP, Schochter F, Scholz C, de Gregorio N, Huober J, et al. Therapeutic intervention based on circulating tumor cell phenotype in metastatic breast cancer: concept of the DETECT study program. *Arch Gynecol Obstet*. 2016 Sep 9;293:271–81.
43. De Gregorio A, Friedl TWP, Huober J, Scholz C, de Gregorio N, Rack B, et al. Discordance in Human Epidermal Growth Factor Receptor 2 (HER2) Phenotype Between Primary Tumor and Circulating Tumor Cells in Women With HER2-Negative Metastatic Breast Cancer. *JCO Precision Oncology*. 2017 Sep;(1):1–12.
44. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol*. Nature Publishing Group; 2018 Feb;15(2):81–94.
45. Parikh AR, Corcoran RB. Monitoring resistance through liquid biopsy. *Annals of Oncology*. 2018 Jan 1;29(1):8–11.
46. Grosse SD, Khoury MJ. What is the clinical utility of genetic testing? *Genetics in Medicine*. 2006 Jul;8(7):448–50.
47. Burke W. Clinical validity and clinical utility of genetic tests. *Curr Protoc Hum Genet*. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2009 Jan;Chapter 9:Unit9.15–9.15.7.
48. Ibáñez de Cáceres I, Cairns P. Methylated DNA sequences for early cancer detection, molecular classification and chemotherapy response prediction. *Clin Transl Oncol*. 2007 Jul;9(7):429–37.
49. Cabel L, Proudhon C, Gortais H, Loirat D, Coussy F, Pierga J-Y, et al. Circulating tumor cells: clinical validity and utility. *Int J Clin Oncol*. 2017 Jun;22(3):421–30.
50. Miquel-Cases A, Schouten PC, Steuten LMG, Retèl VP, Linn SC, van Harten WH. (Very) Early technology assessment and translation of predictive biomarkers in breast cancer. *Cancer Treat Rev*. 2017 Jan;52:117–27.
51. IJzerman MJ, Koffijberg H, Fenwick E, Krahn M. Emerging Use of Early Health Technology Assessment in Medical Product Development: A Scoping Review of the Literature. *Pharmacoeconomics*. 2017 Apr 21;35(7):727–40.
52. Phillips KA, Liang S-Y, van Bebbber S, Canpers Research Group. Challenges to the translation of genomic information into clinical practice and health policy: Utilization, preferences and economic value. *Curr Opin Mol Ther*. 2008 Jun;10(3):260–6.
53. Payne K, Annemans L. Reflections on market access for personalized medicine: recommendations for europe. *Value Health*. 2013 Sep;16(6 Suppl):S32–8.
54. Meckley LM, Neumann PJ. Personalized medicine: factors influencing reimbursement. *Health Policy*. 2010 Feb;94(2):91–100.
55. Ramsey SD, Veenstra D, Tunis SR, Garrison L, Crowley JJ, Baker LH. How Comparative Effectiveness Research Can Help Advance “Personalized Medicine” In Cancer Treatment. *Health affairs*. 2011 Dec 5;30(12):2259–68.
56. Towse A, Garrison LP. Economic incentives for evidence generation: promoting an efficient path to personalized medicine. *Value Health*. 2013 Sep;16(6 Suppl):S39–43.
57. Schneider D, Bianchini G, Horgan D, Michiels S, Witjes W, Hills R, et al. Establishing the Evidence Bar for Molecular Diagnostics in Personalised Cancer Care. *Public Health Genomics*. 2015;18(6):349–58.
58. Wauters E, Vansteenkiste J. Will liquid biopsies become our fluid transition to personalized immunotherapy? *Annals of Oncology*. 2018 Jan 1;29(1):11–3.

59. Bardia A, Haber DA. Solidifying liquid biopsies: can circulating tumor cell monitoring guide treatment selection in breast cancer? *J Clin Oncol*. 2014 Nov 1;32(31):3470–1.

60. Tarazona N, Cervantes A. Liquid biopsy: another tool towards tailored therapy in colorectal cancer. *Annals of Oncology*. 2018 Jan 1;29(1):7–8.

61. Williams PM, Conley BA. Clinical Application of Liquid Biopsies. *JAMA Oncol*. American Medical Association; 2016 Aug 1;2(8):1003–5.

62. Friedrich MJ. Going With the Flow: The Promise and Challenge of Liquid Biopsies. *JAMA : the journal of the American Medical Association*. American Medical Association; 2017 Sep 26;:1095–7.

63. Rogowski WH, John J, IJzerman MJ. Translational Health Economics. In: *World Scientific Handbook of Global Health Economics and Public Policy*. 3rd ed. World Scientific Publishing Company; 2016.

64. Annemans L, Redekop K, Payne K. Current methodological issues in the economic assessment of personalized medicine. *Value Health*. 2013 Sep;16(6 Suppl):S20–6.

65. Degeling K, Koffijberg H, IJzerman MJ. A systematic review and checklist presenting the main challenges for health economic modeling in personalized medicine: towards implementing patient-level models. *Expert Rev Pharmacoecon Outcomes Res*. Taylor & Francis; 2017 Feb;17(1):17–25.

CHAPTER 7

AN ONLINE TOOL FOR THE RAPID ASSESSMENT OF THE HEALTH ECONOMIC IMPACT OF USING CTCs TO GUIDE SYSTEMIC THERAPY IN METASTATIC BREAST CANCER

A.M.S. BERGHUIS, M.J. IJZERMAN, F. BERGER, J.Y. PIERGA,
W. JACOT, C. ALIX-PANABIÈRES, F.C. BIDARD, H. KOFFIJBERG

PENDING SUBMISSION

ABSTRACT

INTRODUCTION: The clinical validity and utility of biomarkers that use a count or expression threshold to classify patients in subgroups, directly depends on this threshold. As it is infeasible to investigate the clinical and economic impact of a range of different thresholds in clinical trials, simulation models can support to determine which threshold potentially is most valuable. This study presents an online tool that enables exploring the health economic impact of using different circulating tumor cell (CTC) count threshold values to guide endocrine- and chemotherapy in metastatic breast cancer patients.

METHODS: An online tool including a decision tree comparing the CTC assay with usual care to guide the choice between either chemotherapy or endocrine therapy in metastatic breast cancer was developed. An exponential distribution was used to reflect the CTC count distribution in a clinical trial population and Weibull distributions were used to reflect overall survival (OS) as a function of these CTC counts. For any given CTC threshold (1, 5 or 8 CTCs per 7,5 mL), modelled outcomes were the overall survival (OS), quality-adjusted life years (QALYs) and the incremental cost-effectiveness ratio (ICER). Parameter values were based on data derived from the STIC METABREAST trial. A hypothetical scenario analysis was performed to demonstrate the potential impact of using different thresholds and survival distributions.

RESULTS: Benefits of guiding treatment by CTC counts obviously depends on the benefits of chemo- and endocrine therapy. Health outcomes were similar in the CTC-guided strategy compared to usual care, when using a threshold of either 1, 5 or 8 CTCs/7.5mL, with incremental QALYs of -0.042, -0.039 and 0.036, respectively. The scenario analysis demonstrated that health outcomes of using CTC-guidance were beneficial when it was assumed that survival with chemotherapy is substantially better in patients with low CTC counts, with incremental QALYs of 0.248, 0.237 and 0.226 for the different thresholds, respectively.

CONCLUSION: Online tools, such as presented here for assessing the impact of CTC-guided treatment in ER+ HER2- breast cancer patients, can provide insights in the clinical utility and health economic outcomes beyond the clinical validity. They may therefore be considered valuable instruments for exploring uncertainty, discussing its impact, and for deliberation on the potential value and design of new clinical trials.

BACKGROUND

Biomarkers are frequently used to classify patients in subgroups and subsequently guide treatment decisions [1]. To be able to classify patients based on the expression of a particular biomarker, thresholds for counts (such as in CTCs) or expression (like PD-L1 expression to guide immunotherapy) are necessary, which are initially derived from observational studies in which the distribution of count values observed in specific populations is compared to clinical outcomes. Following, evidence on the clinical validity of using the particular biomarker with these thresholds is derived from prognostic studies [2]. While the prognostic validity of multiple diagnostic thresholds is frequently evaluated, the clinical utility of these particular biomarkers with different thresholds is typically not assessed and validated in a randomized clinical trial (RCT) [3]. Although the impact of applying a specific biomarker threshold value can be assessed in a clinical trial, identifying which threshold value would be optimal is not possible [4]. However, the threshold applied may considerably influence the clinical benefit of using a particular biomarker, as it determines which subgroups patients are classified into and subsequently, which treatment (if any) they should receive [5]. Therefore, the threshold that will be used to evaluate biomarker impact should be carefully selected. Testing multiple thresholds simultaneously in one study would directly result in a multiplication of the required sample size for evaluation [6]. In the simplest case, when exploring three different threshold values, this would already double the costs of any clinical trial and require twice as many patients to be included (i.e. the trial would have four arms instead of two arms when investigating just one threshold value). When the difference in outcomes between usual care and a single biomarker threshold would be larger than the difference in outcomes between the different applied threshold values, it would again require a larger sample size to reach statistical significance levels if differences between these thresholds also are to be demonstrated. Clearly, clinical trials are not well suited for the process of exploring the impact of using different threshold values.

Before such new biomarkers will be implemented in particular clinical practice settings, their health and economic impact should be demonstrated [7]. Several exploratory modelling methods have been developed to compare the potential health economic benefits of biomarkers with each other from early developmental stages onwards [8, 9]. Exploratory simulation models may be used as an alternative to RCTs to explore the clinical outcomes of hypothetical trial arms based on biomarker expression thresholds that guide treatment decisions. In addition, such models may also provide insight in the health economic outcomes of using particular biomarker expression thresholds

and their potential impact on clinical management. Following, the exploration of the impact of multiple thresholds of biomarker expression may indicate which thresholds potentially are valuable to further evaluate, potentially in an RCT. However, frequently these models are only suitable to evaluate a particular biomarker in a predefined setting [10]. None of these models currently support the option to rapidly and easily update the model parameter values that were used to assess the outcomes of the biomarker based strategies.

Integrating exploratory simulation models in online tools not only improves the transparency of these models, it also provides the opportunity to directly test the impact of using different parameter values and to rapidly update parameter values whenever new evidence becomes available. In addition, these tools have an intuitive graphical interface, which enables easy communication of the model structure and analysis to non-model experts, provides the opportunity to interactively present model outcomes and increases the accessibility of the model if these can be made publicly available [11].

Liquid biopsies are one of the promising diagnostics currently being developed and validated, that may provide important diagnostic (biomarker) information regarding treatment decisions and effectiveness [12]. Substantial effort has been invested in studies evaluating circulating tumor DNA, micro RNA, circulating tumor cells (CTCs) and extracellular vesicles. In metastatic breast cancer, the type of liquid biopsies that are most frequently investigated are CTCs [13]. Extensive research in several developmental phases has been performed and recently the first randomized controlled clinical trials (RCTs) were performed [14]. The STIC METABREAST trial (STIC trial) demonstrated that prescribing either chemo- or endocrine therapy based on a CTC count threshold of 5 CTCs per 7.5 mL of blood is non-inferior to usual clinical decision making [15]. CTCs can be used as a (companion) diagnostic tool by either counting the total number of CTCs in a given amount of blood or by characterizing these cells [16]. Studies aiming to evaluate whether CTCs are potentially clinically valuable in guiding treatment decisions or changing clinical management, typically use a threshold of 5 CTCs per 7.5 mL of blood [17, 18].

In this study, an online tool was developed to explore the impact of using different threshold values for the CTC count used to classify patients into subgroups and, hence, guide treatment. In the tool, the case study from the STIC trial was used, in which CTCs were used to guide systemic therapy in metastatic estrogen receptor (ER) positive, human epidermal growth factor receptor 2 (HER2) negative breast cancer patients.

With this tool rapid assessment of the health economic impact of using different CTC count thresholds is feasible, based on user-defined evidence and distributions.

METHODS

CASE STUDY DESCRIPTION

To demonstrate the value of exploring the potential health economic impact of using different biomarker threshold values for patient classification in an online tool, the case study of the STIC trial was used. In the STIC trial, metastatic ER positive, HER2 negative breast cancer patients were included for whom two diagnostic strategies were defined. The first strategy represented the diagnostics and outcomes as currently used in usual care, the second strategy represented a similar pathway but additionally applied the CTC test, and used a CTC count threshold to guide treatment decisions. In the base-case analysis, a threshold of 5 CTCs per 7.5 mL of blood was used. For patients with ≥ 5 CTCs per 7.5 mL, chemotherapy was prescribed, whereas patients with < 5 CTCs were prescribed endocrine therapy. A simple structured decision tree was developed to estimate the short term incremental costs and health effects, of using the CTC-guided strategy compared to usual care.

PREDICTION OF SURVIVAL IN SUBGROUPS

To reflect the incidence of patients who have a particular number of CTCs, an exponential distribution was defined. Using this distribution, the proportion of patients having more or less CTCs than the threshold can be calculated. To reflect the overall-survival (OS) associated with having a particular number of CTCs, two Weibull distributions were defined, one for survival following chemotherapy and one for survival following endocrine therapy. The shapes of all distributions can be updated in the online tool. Based on these distributions, four estimates of the average OS can be derived, which are:

- OS of patients receiving chemotherapy with less CTCs than the threshold
- OS of patients receiving chemotherapy with more CTCs than the threshold
- OS of patients receiving endocrine therapy with less CTCs than the threshold
- OS of patients receiving endocrine therapy with more CTCs than the threshold

The average utility, i.e. the health-related quality-of-life score, of patients with metastatic breast cancer was derived from the STIC trial and was used to convert the OS to the total number of Quality Adjusted Life Years (QALYs). The average costs of treatment with chemotherapy and endocrine therapy were also derived from the STIC trial.

The CTC count was classified as positive when the number of CTCs was equal to, or higher than the threshold value, whereas the test was classified negative when patients had less CTCs than the threshold value. The sensitivity and specificity of the CTC test were used to calculate the percentage of patients that were diagnosed true positive, false positive, true negative or false negative. In the decision tree it was assumed that patients diagnosed positive for CTCs, i.e. had a number of CTCs equal to or above the threshold value, would always receive chemotherapy, whereas patients would always receive endocrine therapy if they have a number of CTCs below the threshold value. From the decision tree, based on the abovementioned estimations of survival, QALYs and costs, the incremental cost-effectiveness ratio (ICER) was calculated.

TOOL

The full model was implemented in an online tool using the Shiny package in R (version 3.5.1). The online tool is available via https://berghuisams.shinyapps.io/CTC_Thresholds/. The online tool contains a tool guide which explains where and which outcomes are presented and how these should be interpreted.

In the tool, three steps were implemented. First, a user-defined CTC count distribution and a user-defined survival distribution relating the CTC count to survival time of individual patients, can be updated. The incidence distribution can be updated by changing the rate of the exponential distribution. The shape and the scale (or asymptote) parameter from the Weibull distributions can be updated to change the survival distribution. Here, real-time visualization supports changing the distribution parameters in such a way that the resulting distribution curves match the expectations of the user. Second, based on the user defined disutility of chemotherapy, the survival and QALY outcomes for both arms of the decision tree are calculated. Third, the cost estimates for chemotherapy, endocrine therapy and the CTC test can be defined by the user. Finally, the cost-effectiveness of additionally using the CTC test to guide therapy, compared to not using this test, is calculated and this calculation is automatically updated when any of the aforementioned evidence parameters is modified by the user.

SCENARIO ANALYSIS

To further illustrate the use of the tool, a scenario analysis using different values for the CTC threshold, different shapes for the CTC incidence, different shapes for both survival curves and different costs of the CTC test is performed. CTC threshold values that were evaluated in the scenarios were 1, 5 and 8 CTCs per 7.5mL where the value

of 1 and of 8 CTCs were arbitrarily selected to be lower and higher, respectively, than the commonly used threshold of 5 CTCs. These scenarios demonstrate that potential future and hypothetical scenarios can be evaluated within such a tool.

In addition to the base-case scenario in which the survival distributions were derived from the STIC trial, two hypothetical distributions of survival as a function of these CTC counts were defined for patients on chemotherapy and endocrine therapy. In the first hypothetical scenario (adjusted curve 1), survival with endocrine therapy was assumed to be higher than survival with chemotherapy when patients have less than approximately 30 CTCs per 7.5mL of blood. If patients have more than 30 CTCs, survival for patients on chemotherapy was assumed to be higher than for patients on endocrine therapy, thereby implying that chemotherapy only provides a survival benefit for those patients that have more than 30 CTCs. In the second hypothetical scenario (adjusted curve 2), it was assumed that survival for patients on chemotherapy is higher than for patients on endocrine therapy when patients have less than 40 CTCs. Survival for patients with more than approximately 40 CTCs was assumed to be equal for patients receiving chemotherapy or endocrine therapy.

The results of the scenario analysis were presented in an incremental cost-effectiveness plane to demonstrate the impact of updating these parameter values.

RESULTS

In figure 1, a decision tree is presented that demonstrates the principle of using a biomarker with a particular threshold to guide treatment, by using the example of the STIC trial. Outcomes that were assigned to the end point of the decision tree, were assumed to be equal for patient groups with similar characteristics (i.e. patient group 1, 2, 3 and 4 as indicated in figure 1). Patients in group 1 were diagnosed true positive for CTCs, i.e. have a number of CTCs which is equal to or above the threshold value, and received chemotherapy, whereas patients in group 3 were diagnosed false positive for CTCs and also received chemotherapy. Patients in group 4 were true negative for CTCs and received hormone therapy, while patients in group 2 were diagnosed false negative and also received hormone therapy.

In the base-case analysis, the sensitivity of the CTC test was set to 0.8 whereas the specificity was set to 0.7. The utility in general for metastatic breast cancer patients was derived from the STIC trial and calculated to be 0.65 for patients receiving endocrine therapy and 0.60 for patients receiving chemotherapy (disutility of 0.05 for chemotherapy). The incidence of having ≥ 5 CTCs per 7.5mL of blood was 35%

in the STIC trial. In the usual care arm of the STIC trial, 27% of the patients received chemotherapy, whereas 73% received endocrine therapy. Costs for all patient groups were based on the costs of either chemotherapy (€ 29,768) or endocrine therapy (€ 13,993) and the costs of the additional CTC test (€ 450).

Figure 1: decision tree of guiding chemotherapy and endocrine therapy by using the additional CTC test or usual care without CTC test. Inc. = Incidence.

Table 1 presents the survival and QALY outcomes that were used as endpoints for all patient groups as presented in figure 1, based on the user defined distributions for survival following chemotherapy and following endocrine therapy.

Table 1: Outcomes applied to the different patient groups in the decision tree

Patient group	OS (in years)	QALY	Cost (in €)
1: CTC+ and chemotherapy	2.048	0.864	30,264
2: CTC + and endocrine therapy	1.988	0.926	15,182
3: CTC- and chemotherapy	2.374	1.039	30,826
4: CTC- and endocrine therapy	2.227	1.085	13,007
5: Usual care chemotherapy	2.215	0.913	29,768
6: Usual care endocrine therapy	2.129	1.033	13,993

The outcomes of the decision tree are presented in table 2, resulting in an estimated ICER of € 86,865 per QALY.

Table 2: Outcomes of the decision tree for the CTC guided strategy and usual care

Decision Tree Arm	Survival	QALY	Cost (in €)
Usual Care	2.152	1.001	18,252
CTC Guided	2.189	1.003	21,466
Difference	0.037	0.001	3,214

SCENARIO ANALYSIS

In table 3, the outcomes and the parameter values for the 18 hypothetical scenarios are presented. In these scenarios, the impact of multiple CTC count thresholds, different survival curves and potentially lower costs (€ 150 versus € 450) of the CTC test were evaluated.

Table 3: Parameter values of the investigated scenarios

Scenario	CTC threshold	Survival curve	Cost CTC test	QALY difference	Cost difference
1	1	Base-case	€ 450	-0.042	€ 8,740
2	5	Base-case	€ 450	-0.039	€ 8,464
3	8	Base-case	€ 450	-0.036	€ 8,263
4	1	Adjusted curve 1	€ 450	-0.008	€ 8,740
5	5	Adjusted curve 1	€ 450	0.002	€ 8,464
6	8	Adjusted curve 1	€ 450	0.006	€ 8,263
7	1	Adjusted curve 2	€ 450	0.248	€ 8,740
8	5	Adjusted curve 2	€ 450	0.237	€ 8,464
9	8	Adjusted curve 2	€ 450	0.226	€ 8,263
10	1	Base-case	€ 150	-0.042	€ 8,440
11	5	Base-case	€ 150	-0.039	€ 8,164
12	8	Base-case	€ 150	-0.036	€ 7,963
13	1	Adjusted curve 1	€ 150	-0.008	€ 8,440
14	5	Adjusted curve 1	€ 150	0.002	€ 8,164
15	8	Adjusted curve 1	€ 150	0.006	€ 7,963
16	1	Adjusted curve 2	€ 150	0.248	€ 8,440
17	5	Adjusted curve 2	€ 150	0.237	€ 8,164
18	8	Adjusted curve 2	€ 150	0.226	€ 7,963

In figure 2, the specific shapes of the base-case and hypothetical survival curves and the parameter values that were used to update these distributions are presented.

Figure 2: Survival distributions used in the scenario analysis

In figure 3, the QALY and cost differences between the CTC guided strategy and the usual care strategy are presented. The exact numbers of these point-estimates were presented in table 3. In those scenarios in which the second adjusted survival

curve was used (in which it was assumed that chemotherapy has a survival benefit for patients with less than approximately 40 CTCs), it was demonstrated that health outcomes are better for patients when the CTC-guided strategy is used and costs are higher than the usual care strategy. For those patients in which the first survival curve was used (in which it was assumed that chemotherapy only has a survival benefit for patients with more than approximately 30 CTCs), it was demonstrated that health outcomes are similar when using the CTC-guided strategy compared to usual care, while the costs are higher. When using the survival distributions derived from the STIC trial (base-case analysis), it was demonstrated that health outcomes are slightly worse when using the CTC-guided strategy compared to usual care, whereas the costs are higher.

Figure 3: Cost-effectiveness plane of the results of the scenario analysis

DISCUSSION

Many diagnostic tests require a threshold to inform treatment- or other patient management decisions. Although relevant information may be lost when dichotomizing continuous test results, this is often necessary to support decisions which are binary by nature (e.g. treatment with either chemotherapy or endocrine therapy). While the impact of using specific threshold values can to some extent be explored in prognostic cohort studies, it is more difficult to design prospective RCTs to compare their impact. Nonetheless, RCTs are essential in building the evidence base required for successful implementation. Simulation of health and economic outcomes to guide the design and execution of future RCTs may therefore support efficient use of resources by

increasing the chance that such trials can actually demonstrate (clinical) meaningful differences [19].

In this study it was demonstrated that it is feasible to use an online decision support tool to explore the impact of different CTC count thresholds and their accompanying survival distributions on health and economic outcomes. Creating such tools is relatively inexpensive compared with performing large, time consuming RCTs but they can obviously never produce similar evidence, and should therefore not be interpreted as a replacement of these trials but as complementary analyses. In addition to standard cost-effectiveness analysis, other quantitative methods, such as value of information analysis, and budget impact analysis may be implemented in such tools.

This would provide further insights into key parameters in the analysis, their uncertainty, and potential policy implications, which can in turn facilitate further discussion between clinicians, analysts, policy makers, and funding agencies on guiding and prioritizing and future study designs. In this tool, it was assumed that a positive test result was followed by the prescription of chemotherapy. Such tools may, however, also include multiple or other treatment strategies, which enables to model the impact of using such diagnostics when clinicians deviate from the recommended treatment based on the test results.

However, there are currently no guidelines or strict regulations on the responsibilities of the developers or end users of these tools. Many modern software tools (e.g. software packages in the R environment for statistical computing) are provided without any warranty. Furthermore, the transparency of programming code or tools, and the efforts spent on their validation are often unclear, although open source modelling initiatives and new coding frameworks are now appearing in literature [20-22]. Though the developers may have intended to build tools for explorative purposes only, they will in general not be able to fully prohibit end users to derive outcomes for settings for which the tool was not validated, nor from prohibiting them to assign too much value to the (inherently uncertain) outcomes of such an exploratory analysis. In addition, the accuracy of these tools is dependent on the setting to which the tool is applicable. Solely for explorative research purposes, the extent to which parameters may be changed is flexible and not necessarily solely based on evidence, whereas parameter values in decision tools that directly support clinical decision making should be extremely accurate and evidence-based. Furthermore, the importance of the different aspects of outcome representations and visualization within these tools

could likely be tailored to the expertise and interests of different stakeholders and may therefore deviate from some users' expectations.

In the current analysis, an exponential distribution and two Weibull distributions were used to reflect the CTC counts and survival distributions. Even though these distributions are rather flexible, when fitting these distributions to the STIC trial data, these do still not perfectly reflect the survival and CTC count distributions for these patients. Consequently, when these distributions are used instead of the real STIC data, health economic outcomes can, and have demonstrated to be, different from the health economic outcomes resulting when using the full empirical dataset. Furthermore, ideally the selected distribution type should also allow to reflect that prescribing chemotherapy or endocrine therapy is equally good in terms of survival when treating patients with a CTC count below the threshold, whereas chemotherapy might have additional survival benefit for patients with a CTC count above the threshold. This would likely require combining different distributions for these different trajectories, and would therefore also require more real-world survival data to fit these distributions to, which is usually not available. Within the STIC trial, only in a few patients high numbers of CTCs were detected. Currently these were aggregated in the subgroup of patients having over 100 CTCs, even though in some of these patients just over 100 CTCs were detected, whereas in others almost 3000 CTCs were detected.

No substantial differences in health or economic outcomes were demonstrated when using a threshold of 1, 5 or 8 CTCs per 7.5 mL blood. However, when handling a threshold of 5 CTCs per 7.5mL, the probability of deriving a false negative test result is substantially higher for patients with 6 or 7 CTCs than for patients with >1000 CTCs. These patient-level differences were not taken into account in the current analysis, but may have influenced the final outcome measures.

CONCLUSIONS

Online tools to assess the impact of using different biomarker thresholds, enable a simple and visualized exploration of the impact that key parameters have in health economic analyses. Even though such tools will never replace RCTs or other empirical studies, they are valuable to facilitate discussions between clinicians and other experts involved. Since developing and using such tools is relatively easy and cheap there is no reason to further explore their merits and potential applications. In this study, we demonstrated the value of a decision tool using a case study on a CTC test in metastatic breast cancer patients. The current exploration indicated that using CTC detection in

these patients may be of particular interest for those subgroups of patients with low numbers of CTCs when it was assumed that chemotherapy could provide substantial survival benefits for these patients. Though the current tool provided insights on the impact of using different thresholds of CTC counts, such tools can be easily adapted to give insights in the impact on health and economic outcomes associated with the use of other biomarkers in different diseases.

References

1. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol*. 2016;34(10):1134-50.
2. Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. *Personalized medicine*. 2010;7(1):33-47.
3. Subtil F, Rabilloud M. Estimating the optimal threshold for a diagnostic biomarker in case of complex biomarker distributions. *BMC medical informatics and decision making*. 2014;14:53.
4. van Giessen A, de Wit GA, Moons KGM, Dorresteijn JAN, Koffijberg H. An alternative approach identified optimal risk thresholds for treatment indication: an illustration in coronary heart disease. *Journal of clinical epidemiology*. 2018;94:122-31.
5. Diao G, Dong J, Zeng D, Ke C, Rong A, Ibrahim JG. Biomarker threshold adaptive designs for survival endpoints. *Journal of biopharmaceutical statistics*. 2018;28(6):1038-54.
6. Jiang W, Freidlin B, Simon R. Biomarker-adaptive threshold design: a procedure for evaluating treatment with possible biomarker-defined subset effect. *J Natl Cancer Inst*. 2007;99(13):1036-43.
7. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovsk F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II-An ISPOR Good Research Practices Task Force report. *Value Health*. 2015;18(2):161-72.
8. de Graaf G, Postmus D, Westerink J, Buskens E. The early economic evaluation of novel biomarkers to accelerate their translation into clinical applications. *Cost Effectiveness and Resource Allocation*. 2018;16(1):23.
9. Kluytmans A, Deinum J, Jenniskens K, van Herwaarden AE, Gloerich J, van Gool AJ, et al. Clinical biomarker innovation: when is it worthwhile? *Clinical chemistry and laboratory medicine*. 2019.
10. Postmus D, de Graaf G, Hillege HL, Steyerberg EW, Buskens E. A method for the early health technology assessment of novel biomarker measurement in primary prevention programs. *Statistics in medicine*. 2012;31(23):2733-44.
11. Carlson JJ, Walton SM, Basu A, Chapman RH, Campbell JD, McQueen RB, et al. Achieving Appropriate Model Transparency: Challenges and Potential Solutions for Making Value-Based Decisions in the United States. *PharmacoEconomics*. 2019.
12. Ravelli A, Reuben JM, Lanza F, Anfossi S, Cappelletti MR, Zanotti L, et al. Breast cancer circulating biomarkers: advantages, drawbacks, and new insights. *Tumour Biol*. 2015;36(9):6653-65.
13. Berghuis AM, Koffijberg H, Prakash J, Terstappen LW, MJ IJ. Detecting Blood-Based Biomarkers in Metastatic Breast Cancer: A Systematic Review of Their Current Status and Clinical Utility. *Int J Mol Sci*. 2017;18(2).
14. Bidard FC, Fehm T, Ignatiadis M, Smerage JB, Alix-Panabieres C, Janni W, et al. Clinical application of circulating tumor cells in breast cancer: overview of the current interventional trials. *Cancer metastasis reviews*. 2013;32(1-2):179-88.

15. Bidard F-C, Jacot W, Dureau S, Brain E, Bachelot T, Bourgeois H, Goncalves A, Ladoire S, Naman H, Dalenc F, Gligorov J, Espie M, Levy C, Ferrero J-M, Loirat D, Cottu P, Dieras V, Simondi C, Berger F, Alix-Panabieres C, Pierga J-Y. Clinical utility of circulating tumor cell count as a tool to chose between first line hormone therapy and chemotherapy for ER+ HER2- metastatic breast cancer: Results of the phase III STIC CTC trial [abstract]. Proceedings of the 2018 San Antonio Breast Cancer Symposium; 2018 Dec 4-8; San Antonio, TX. Philadelphia (PA): AACR; Cancer Res 2019;79(4 Suppl):Abstract nr GS3-07. 2018.
16. Cabel L, Proudhon C, Gortais H, Loirat D, Coussy F, Pierga JY, et al. Circulating tumor cells: clinical validity and utility. *Int J Clin Oncol*. 2017;22(3):421-30.
17. IJzerman MJ, Berghuis AMS, de Bono JS, Terstappen L. Health economic impact of liquid biopsies in cancer management. *Expert Rev Pharmacoecon Outcomes Res*. 2018:1-7.
18. Bidard FC, Peeters DJ, Fehm T, Nole F, Gisbert-Criado R, Mavroudis D, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *The Lancet Oncology*. 2014;15(4):406-14.
19. Jenniskens K, Lagerweij GR, Naaktgeboren CA, Hooft L, Moons KGM, Poldervaart JM, et al. Decision analytic modeling was useful to assess the impact of a prediction model on health outcomes before a randomized trial. *Journal of clinical epidemiology*. 2019;115:106-15.
20. Jansen JP, Incerti D, Linthicum MT. Developing Open-Source Models for the US Health System: Practical Experiences and Challenges to Date with the Open-Source Value Project. *Pharmacoeconomics*. 2019.
21. Incerti D, Curtis JR, Shafrin J, Lakdawalla DN, Jansen JP. A Flexible Open-Source Decision Model for Value Assessment of Biologic Treatment for Rheumatoid Arthritis. *Pharmacoeconomics*. 2019;37(6):829-43.
22. Alarid-Escudero F, Krijkamp EM, Pechlivanoglou P, Jalal H, Kao SZ, Yang A, et al. A Need for Change! A Coding Framework for Improving Transparency in Decision Modeling. *Pharmacoeconomics*. 2019.

CHAPTER 8

HEALTH ECONOMIC IMPACT OF USING CTCS TO GUIDE SYSTEMIC THERAPY IN METASTATIC ER+ HER2- BREAST CANCER PATIENTS RESULTS FROM A RANDOMIZED CONTROLLED MULTICENTER TRIAL

A.M.S. BERGHUIS, H. KOFFIJBERG, F. BERGER, C. ALIX-PANABIÈRES,
W. JACOT, J.Y. PIERGA, F.C. BIDARD, M.J. IJZERMAN

PENDING SUBMISSION

ABSTRACT

BACKGROUND: In metastatic Estrogen Receptor (ER) positive, Human Epidermal Growth Factor Receptor 2 (HER2) negative breast cancer patients, the decision to prescribe endocrine- (in low risk patients) or chemotherapy (in high risk patients) as first line metastatic treatment is based on the absence of visceral crisis or adverse prognostic characteristics which cannot be objectively measured and result in practice variation and over-treatment. Recently, the STIC METABREAST trial demonstrated that using the number of circulating tumor cells (CTCs) to standardize the current decision is non-inferior to current clinical practice in terms of Progression Free Survival (PFS). This study assessed the cost-utility of standardizing the prescription of endocrine- or chemotherapy by using a CTC count threshold (5 CTCs/7.5mL) compared to current clinical practice.

METHODS: Patient-level data (n=755) were derived from the prospective, multicenter (n=17), randomized controlled STIC METABREAST trial. Utilities were calculated by mapping the scores of the quality of life questionnaire (QLQ-C30) to EQ-5D utilities, and used to calculate Quality-Adjusted Life Years (QALY) over a 2-year time horizon. The incremental costs, effects and the incremental cost-effectiveness ratio (ICER) were calculated for the whole study population, the clinically low-risk subgroup in which treatment was escalated from endocrine- to chemotherapy and the clinically high-risk subgroup in which treatment was de-escalated from chemo- to endocrine therapy. A two-stage bootstrap approach was used to assess the uncertainty in outcomes.

RESULTS: Health outcomes were similar and costs were higher in the CTC guided arm (1.011 QALYs; €19,382) compared to the usual care arm (1.000 QALYs; €18,250), resulting in an ICER of €102,909/QALY. Using CTC detection in the high risk subgroup has favorable cost-effectiveness (ICER €-353,444) and unfavorable cost-effectiveness in the low risk subgroup (ICER €570,285). Results from the bootstrap (n=10,000) showed a 31% probability that CTC-guidance is cost-effective when considering a threshold of €20,000/QALY in the full population, 5% in the low risk subgroup and 92% in the high risk subgroup, respectively.

CONCLUSION: This trial-based health economic analysis demonstrated that CTC guided selection of endocrine and chemotherapy in metastatic ER+ HER2- breast cancer patients is unlikely to be cost-effective in the full patient population on the short term. Subgroup analyses demonstrated that CTC-guidance to de-escalate from chemotherapy to endocrine therapy in patients initially classified as high risk by physicians, is likely cost-effective.

BACKGROUND

The most frequently diagnosed form of cancer among women is breast cancer (1). The last decades, substantial progress was made in the diagnostics and treatment of breast cancer. For several subtypes of breast cancer that express a particular protein or receptor, targeted therapies have become available which substantially improve survival for these patients (2). However, metastatic breast cancer still largely is an incurable condition, with a median survival of only 2 years after detection of the metastasis (3, 4). To extend survival for these patients, it is important to classify the type of tumor and select the optimal treatment as early as possible (5, 6). Currently, characterization and staging of breast cancer tumors is done by using histopathological test results from tissue biopsies, imaging modalities or combinations of those. For most subtypes of breast cancer, clinical guidelines already include recommendations for the most effective treatment strategy (7, 8).

In metastatic estrogen receptor (ER) positive, human epidermal growth factor receptor 2 (HER2) negative breast cancer patients, physicians can prescribe either chemotherapy or endocrine therapy. Whether patients are eligible for either one of those therapies is predominantly based on the absence of visceral crisis or adverse prognostic factors, which currently can not be objectively measured and therefore, may cause clinical practice variation (9). In order to decrease potential undesirable practice variation or overtreatment that is likely to arise due to subjective judgements, standardized objective criteria to guide these treatment decisions should be developed. Liquid biopsies are one of the innovations developed last decades that have shown to have prognostic value for breast cancer patients and may therefore support and improve the standardization of these treatment decisions (10-12). A substantial number of liquid biopsies has yet been investigated, such as circulating tumor DNA, circulating tumor RNA, circulating proteins or circulating tumor cells (CTCs). The most frequently investigated type of liquid biopsies in metastatic breast cancer are CTCs. Several studies have already shown the prognostic validity of CTCs, whereas the clinical role of CTCs as a diagnostic strategy to decide on treatment decisions is currently evaluated in several trials with different aims (13, 14).

The STIC CTC METABREAST trial (STIC trial) investigated whether guiding endocrine- or chemotherapy by using the number of CTCs as a threshold, is non-inferior to guiding this decision based on patient- and tumor characteristics and the physicians' expert opinion (15). Recently, the results of the STIC trial demonstrated non-inferiority in using CTCs to guide chemo- or endocrine therapy in metastatic ER+

HER2- breast cancer patients (16). In addition, it was demonstrated that patients that were classified as high risk, either clinically driven or CTC driven (≥ 5 CTCs per 7.5mL of blood), that receive chemotherapy have a significantly longer progression free survival (PFS) and overall survival (OS) than patients that receive endocrine therapy. Usually, the aim of performing a non-inferiority trial is to demonstrate that especially toxic treatment can be safely de-escalated. The STIC trial additionally demonstrated, however, that the CTC driven strategy may also result in longer PFS when escalating treatment from endocrine- to chemotherapy when ≥ 5 CTCs were found in patients that were initially clinically identified as low risk.

Cancer related healthcare expenditures currently already account for 4-6% of the total healthcare expenditures in the European Union and are still growing (17). Apart from the PFS benefits that were demonstrated in the STIC trial, improved and standardized allocation of systemic treatment, may also support sustainable cancer care. To inform decisions regarding the implementation of new technologies, an empirical cost-effectiveness analysis based on data collected alongside a clinical trial should be performed (18). Along the STIC trial, data on health economic outcome measures were collected. This study evaluates the health economic impact of standardizing treatment decisions on endocrine and chemotherapy by using CTC guidance in ER positive HER2 negative breast cancer patients in a cost-utility analysis.

METHODS

DATA

Patient-level data on metastatic ER+ HER2- breast cancer patients were derived from the randomized controlled STIC trial. In this prospective, multicenter trial, patients diagnosed with metastatic disease eligible for their first line of metastatic treatment were included and followed for 2 years. For all patients included in the study, the treating physician initially determined whether these patients were eligible for either chemotherapy or endocrine therapy based on the absence of visceral crisis or other adverse prognostic factors. Subsequently, for all patients a CTC test was performed to determine the number of CTCs in 7.5 mL of blood. Following, these patients were randomized over two arms: A) usual care strategy: the physician prescribes endocrine therapy or chemotherapy based on patient- and tumor characteristics, and B) CTC guided strategy: the physician prescribes endocrine therapy when the number of CTCs is < 5 and prescribes chemotherapy when the number of CTCs is ≥ 5 . The number of CTCs of patients in arm A was not disclosed to the treating physician. In figure 1, an overview of the patients included in each subgroup in both arms is presented.

Arm	Arm A (Usual Care)				Arm B (CTC-guided)			
Physician's Choice	Endocrine therapy		Chemotherapy		Endocrine therapy		Chemotherapy	
No. of CTCs	< 5 CTCs	≥ 5 CTCs	< 5 CTCs	≥ 5 CTCs	< 5 CTCs	≥ 5 CTCs	< 5 CTCs	≥ 5 CTCs
Therapy received	Endo	Endo	Chem	Chem	Endo	Chem	Endo	Chem
Subgroups	1 Clinically low risk		2 Clinically high risk		3 Clinically low risk		4 Clinically high risk	

Figure 1: Overview of the treatment allocation decisions of patients included in the STIC Trial

Endo = endocrine therapy, Chem = chemotherapy

In the subgroup analysis, both clinically low risk and both clinically high risk subgroups will be compared. In the clinically low risk subgroup, the health economic consequences of using CTCs to standardize the decision to escalate from endocrine therapy to chemotherapy will be evaluated. In the clinically high risk subgroup, the health economic consequences of using CTCs to standardize the decision to de-escalate from chemotherapy to endocrine therapy will be evaluated.

EVALUATION POINTS

Clinical patient- and tumor characteristics were evaluated at inclusion, every 2 months in the first year and every 4 months in the second year, resulting in a total of 9 time points at which data were collected. In addition, data on the number and type of imaging tests performed, toxic reactions, hospitalization and productivity losses were collected by using a customized evaluation form (designed for the STIC trial) at each of these time points. Data on quality of life were collected using the Quality of Life Questionnaire C30 (QLQ-C30), developed by the European Organization for Research and Treatment of Cancer (EORTC) (19). These questionnaires were completed at inclusion, 2 months, 6 months, 12 months and 24 months.

OUTCOME MEASURES

A cost-utility analysis of using CTCs to standardize the decision to guide either endocrine- or chemotherapy, compared to usual care, was performed. The main outcome measure was the incremental cost-effectiveness ratio (ICER), which is based on the incremental costs divided by the incremental health outcomes in terms of

quality adjusted life years (QALYs) of the CTC guided strategy compared with usual care. Both, cost and effect outcome measures were discounted at a rate of 3% for the follow-up time. As the trial was designed to be a non-inferiority trial, confidence intervals (CIs) for the health economic outcome measures were presented (20).

HEALTH OUTCOMES

The QLQ-C30 incorporates 30 questions classified in 5 functional and 3 symptom scales. As this questionnaire does not allow immediate calculation of health-related quality of life scores (i.e. utility values) a previously published mapping algorithm was used to convert questionnaire answers to utility values (21). When answers on multiple QLQ-C30 questions are missing, the EORTC recommends to impute this data (19). Imputation of missing values was performed in questionnaires with 15 or more questions filled in answers. Multiple imputation was performed in R (version 3.5.1) using the mice package, and was conducted separately for each point in time at which patients should have completed the QLQ-C30 (22). For every time point at which the QLQ-C30 administered, 10 datasets were imputed. In each imputation, all question answers were used as predictors to missing answers. For patients with full answers, and patients for whom these 10 datasets could be imputed, the QLQ-C30 scale scores were mapped to utilities using the response mapping model of Longworth et. al (21, 23, 24). The final utility value at each time point, was defined as the average of the utilities calculated from the 10 imputed datasets. For patients with less than 15 questions answered in a questionnaire, the QLQ-C30 answers were not imputed, and no utility could be calculated from the answers. Therefore, missing utilities were imputed directly for these patients based on the utility derived from the next or previous questionnaire, the arm in which patients were included, whether or not they received chemotherapy, whether any toxicities were experienced in the period before the particular evaluation, progression free survival (PFS) and overall survival (OS) time, the number of CTCs and the age of the patients.

8

QALYs were calculated by integrating the utility values of patients over the time period in which they experienced these utility values, in a stepwise approach. For example, if a patient was alive for the whole period between the first and the second completion of the QLQ-C30, the average utility over that period was multiplied with the length of that period. If a patient died within a specific period, the average of the utility from the last completed QLQ-C30 and the utility of being dead (utility: 0) was multiplied with the length of the period the patient was still alive after this last completed QLQ-C30.

COSTS

List prices were derived from the French Healthcare Authority and were used to allocate costs to the imaging tests used and therapies prescribed during each evaluation period. Drug costs for endocrine therapy were calculated by multiplying the list price for a particular drug with the usually prescribed dose of that particular drug in a standard treatment cycle or time period. When no specific treatment switch or stopping date was available, it was assumed that the treatment was continued until progression. Total drug costs were calculated by multiplying the drug costs with the number of treatment cycles the drug was prescribed.

In France, the costs of all hospital stays are registered to diagnosis-treatment-combinations, for which full treatment prices are available independent on the ward in which the patient stayed in the hospital. For those patients in whom the cause of unplanned hospital stays during the evaluation period were unknown, Dutch list prices were used to estimate the cost of the hospital based on the length of stay at the particular ward in the hospital. Specifically the Dutch costs of hospital stays were chosen to base these estimations on, as the Dutch healthcare system has a similar diagnosis-treatment-combination coding system, but also provides an overview of the costs of hospital stays in several departments of the hospital. As some of these costs were derived from reference prices from different years, these costs were adjusted to 2019 costs by using Dutch consumer price index levels (25).

Total costs calculated per patient included the cost of treatment, CTC detection (including all materials necessary for the analysis of the results), imaging and hospital monitoring (consultations, medical procedures, possible complications or toxicities and hospitalizations).

SENSITIVITY ANALYSIS

In the STIC trial, multiple centers in France were included until the final number of patients needed to obtain sufficient statistical power was reached. To reflect the uncertainty in the health and economic outcomes, caused by variation on patient-level and between centers, a two-stage bootstrap was performed. In each bootstrap sample, institutes were sampled until at least the number of patients included in the original trial were included. Following, for each sampled institute, the same number of patients as initially included in the trial from that particular institute, were randomly sampled with replacement. For each sampled dataset, the incremental costs, QALYs and the corresponding ICER of the usual care arm compared to the CTC

guided arm were calculated. The bootstrap was repeated 10,000 times. In addition, for all subgroups, i.e. the full trial population, the clinically low risk patients and the clinically high risk patients, analyzed in the bootstraps, a cost-effectiveness acceptability curve was created.

RESULTS

In the STIC trial patients with metastatic ER+ HER2- breast cancer from multiple centers (n=17) in France were included (n=755) and randomized over the usual care arm (n=378) and the CTC guided arm (n=377).

The number of patients alive for whom 15 or more questions were unanswered in the QLQ-C30 questionnaires increases over time, with a total number of n=72 (9.54 %) for the first questionnaire (timepoint 1), n=186 (24.64%) for the second questionnaire (timepoint 2), n=288 (38.15%) for the third questionnaire (timepoint 3), n=421 (55.76%) for the fourth questionnaire (timepoint 4) and n=571 (75.63%) for the fifth questionnaire (timepoint 5), respectively.

The exact costs that were allocated to hospital stays, treatments prescribed, imaging techniques used and management of (potential) toxic reactions, are presented in appendix A.

In table 1, the number of patients, mean survival outcomes, QALYs and costs are presented for both trial arms. Guiding chemo- or endocrine therapy based on CTCs (arm B) results in slightly higher average QALYs (difference of 0.011) and higher costs (difference of € 1,132). These health-economic outcomes result in an ICER of € 102,909 per QALY gained, which exceeds the applied cost-effectiveness threshold of € 20,000 per QALY. The relatively wide 95% CIs, as presented in table 1, imply that there may be substantial patient heterogeneity leading to large differences in outcomes in terms of QALY and costs between individuals, given that the number of patients per subgroup is substantial.

Table 1: Health and economic outcomes with confidence intervals per arm and per included subgroup

Arm	No. of patients	Mean OS (months)	Mean PFS (months)	QALYs			Cost (€)		
				Mean	CI lower limit	CI upper limit	Mean	CI lower limit	CI upper limit
Arm A	378	25.8	14.67	1.000	0.952	1.048	18250	15847	20653
Arm B	377	26.0	16.18	1.011	0.961	1.061	19382	17160	21603

In table 2, the number of patients, mean survival outcomes, QALYs and costs are presented for the clinically relevant subgroups that were presented in figure 1.

Table 2: Health and economic outcomes with confidence intervals per arm and per included subgroup

Subgroup *	No. of patients	Mean OS (Months)	Mean PFS (Months)	QALYs			Cost (€)		
				Mean	CI lower limit	CI upper limit	Mean	CI lower limit	CI upper limit
Low risk (1)	275	25.5	15.12	1.032	0.978	1.087	13945	11288	16601
High risk (2)	103	26.6	13.47	0.915	0.819	1.01	29745	25175	34316
Low risk (3)	277	26.5	17.32	1.039	0.982	1.097	17937	15587	20286
High risk (4)	100	24.8	13.02	0.933	0.83	1.035	23383	18169	28598

* Subgroups were defined in figure 1

The clinical results of the STIC trial have indicated that only using CTCs to escalate treatment from endocrine therapy to chemotherapy in patients whom were classified as low risk patients by the physician, is beneficial in terms of PFS (mean PFS: 17.3 months in the CTC guided arm, 15.1 months in the physician guided arm). The QALY difference between these groups is 0.007 (1.039 CTC guided arm/subgroup 3 versus 1.032 in the usual care arm/subgroup 1), the cost difference is € 3,992, resulting in an ICER of € 570,285 per QALY gained.

When assuming that CTCs will only be used in those patients that were classified as high risk patients by the physician and subsequently prescribed chemotherapy, treatment was de-escalated in patients in whom <5 CTCs were detected. The results of using CTCs de-escalate treatment in those patients with <5 CTCs compared to the physicians choice (which is prescribing chemotherapy) are comparable in terms of PFS (mean PFS: 13.0 months in the CTC guided arm /subgroup 4 versus 13.5 months in the CTC guided arm/subgroup 2). The QALY difference between these groups is 0.018 (0.982 in the CTC guided arm/subgroup 4 versus 0.978 in the CTC guided arm/subgroup 2), whereas the cost difference between these groups is € -6,362, resulting in an ICER of € -353,444 per QALY gained.

The calculated ICERs are, however, uninformative given the low differences in QALY outcomes and the patient-level uncertainty between these QALY outcomes.

SENSITIVITY ANALYSIS

To estimate the uncertainty in the calculated health and cost outcomes, and ICER, two-stage bootstrapping with 10,000 replications was performed. The results of the bootstrap illustrate that the substantial heterogeneity in outcomes between patients translates to substantial uncertainty in overall outcomes, and are presented in figure 2. When using CTCs to either escalate or de-escalate treatment, as in the original STIC trial, 31.41% of the bootstrap samples can be considered cost-effective when considering a willingness to pay (WTP) threshold of € 20,000 per QALY.

Figure 2: Incremental costs and QALYs for all bootstrap samples in the full trial population (arm A versus arm B) and in the clinically high- (subgroup 4 versus 2) and clinically low risk group (subgroup 3 versus 1) (n=10,000).

When using CTCs to escalate treatment for those patients that were initially classified as clinically low risk and subsequently would have received endocrine therapy, 5.16% of the bootstrap samples can be considered cost-effective when considering a WTP threshold of € 20,000 per QALY. When using CTCs to de-escalate treatment for those patients that were initially classified as clinically high risk and subsequently would have received chemotherapy, 92.10% of the bootstrap samples can be considered cost-effective when considering a WTP threshold of € 20,000 per QALY.

Figure 3: Cost-effectiveness Acceptability Curves for the full trial group, the clinically low- and the clinically high risk subgroup.

Figure 3 visualizes how the probability that CTC-guide treatment is cost-effective compared with usual care changes when the willingness-to-pay threshold is varied, in a cost-effectiveness acceptability curve. Using CTC detection to select those patients in which treatment can be de-escalated (patients initially classified as being high risk patients), has the highest probability of being cost-effective for all willingness to pay thresholds in the range of € 0 to € 175,000.

Discussion

Though the STIC trial demonstrated non-inferiority in terms of PFS of using the CTC guided strategy to standardize the decision to prescribe either endocrine- or chemotherapy compared to usual care in particular subgroups, health-economic outcome measures demonstrated substantial uncertainty. The incremental QALY difference between the CTC guided arm and the usual care arm was found to be rather small and uncertain, as demonstrated by the large CIs.

The sensitivity analysis showed that incremental QALYs range from -0.2 to 0.2, while the incremental costs range from € -20,000 to € 10,000. Consequently, this cost-utility

analysis of the full patient population included in the STIC trial indicates that using CTCs to guide chemo- or endocrine therapy is not cost-effective. In addition, as this is designed to be a non-inferiority study with only marginal clinical benefits, the ICER is relatively uninformative. Furthermore, as the follow-up of the STIC trial was only 2 years, this censored time horizon may have been too short to support a positive outcome in terms of the traditionally estimated ICER. First, the disutility of escalating from endocrine therapy to chemotherapy when ≥ 5 CTCs were detected, may not outweigh the survival benefits. Second, the utility gain derived when de-escalating treatment from chemotherapy to endocrine therapy when < 5 CTCs were detected, may not outweigh the additional costs of using CTCs to guide this treatment. Though survival and QALY outcomes beyond this 2 year time horizon may be extrapolated by using statistical modelling techniques, collecting data on the actual long term survival and QALY results for the included patients will take some years.

From a clinical perspective, new diagnostics are more likely to be used as guidance strategies to escalate treatment than to de-escalate treatment, mainly due to potential safety concerns. The clinical outcomes of the STIC trial demonstrated that escalating treatment from endocrine- to chemotherapy when ≥ 5 CTCs were detected in the subgroup of clinically low risk patients appeared to be more beneficial than de-escalating treatment from chemo- to endocrine therapy when < 5 CTCs were detected in the subgroup of clinically high risk patients in terms of their PFS. These benefits are in line with the potential setting in which new diagnostics are potentially most likely to be used. However, from a health economic perspective, it was demonstrated that de-escalating treatment from chemo- to endocrine therapy when < 5 CTCs were detected is most likely to be cost-effective. However, it is unlikely that particular de-escalation strategies will be implemented in clinical practice. Currently, the cost-effectiveness of escalating treatment when ≥ 5 CTCs were detected in patients that were classified initially as clinically low risk patients was unfavorable, but may be due to the short time horizon. As using CTCs in this subgroup seems to be most promising from a clinical perspective, collecting long term outcomes over an extended time horizon in both subgroups may be of particular interest for future studies.

Several methods were selected to calculate the health economic outcomes. The QLQ-C30 questionnaire was mapped to utilities using the algorithm of Longworth et. al. (21). Using a different algorithm could influence the patient-level QALYs that were calculated and subsequently the health economic outcomes. Furthermore, in the analysis of the full trial population, list prices were used to calculate patient-level costs. Therefore, the uncertainty in outcomes assessed with the bootstrap approach

could not include any potential uncertainty in the costs of specific treatments prescribed, imaging modalities used and monitoring strategies.

In the STIC trial, a CTC count threshold of 5 CTCs per 7.5 mL blood was used to either prescribe chemotherapy (≥ 5 CTCs) or endocrine therapy (<5 CTCs) in the CTC guided trial arms, which is a frequently used but still arbitrary threshold. Future research modelling potential survival outcomes when using different CTC thresholds should identify whether a different threshold might result in more favorable cost-effectiveness outcomes.

CONCLUSION

Using CTC counts to guide endocrine and chemotherapy currently appears unlikely to be cost-effective on the short term when used in all metastatic ER+ HER2- breast cancer patients. Incremental differences in QALYs were relatively small between the CTC guided trial arms and between the separately analyzed subgroups. Using CTCs to de-escalate from chemotherapy to endocrine therapy in the clinically high risk subgroup seems most likely to be cost-effective, as the CTC guidance has a high chance of saving costs without affecting health outcomes. The STIC trial also demonstrated that escalating treatment from endocrine to chemotherapy is beneficial from a clinical perspective. However, from a health economic perspective it is unlikely that this strategy will be cost-effective when evaluating short term health economic outcomes, as the additional prescription of chemotherapy will be more costly. Consequently, it is of particular interest for future research to collect and evaluate long term health and economic outcomes of this particular subgroup.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(1):7-30.
2. Schramm A, De Gregorio N, Widschwendter P, Fink V, Huober J. Targeted Therapies in HER2-Positive Breast Cancer - a Systematic Review. *Breast Care (Basel)*. 2015;10(3):173-8.
3. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359-86.
4. Köhne CH, Ufen M-P, Novopashenny I, Wischneswky M, Wolters R, Regierer AC, et al. Metastatic breast cancer: are we treating the same patients as in the past? *Annals of Oncology*. 2013;25(1):95-100.
5. Polyak K. Heterogeneity in breast cancer. *J Clin Invest*. 2011;121(10):3786-8.
6. Reinert T, Barrios CH. Overall survival and progression-free survival with endocrine therapy for hormone receptor-positive, HER2-negative advanced breast cancer: review. *Ther Adv Med Oncol*. 2017;9(11):693-709.
7. NABON. Breast cancer guideline. 2012.
8. Mueller C, Haymond A, Davis JB, Williams A, Espina V. Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev Proteomics*. 2018;15(2):131-52.
9. Cardoso F, Senkus E, Costa A, Papadopoulos E, Aapro M, Andre F, et al. 4th ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 4)dagger. *Ann Oncol*. 2018;29(8):1634-57.
10. Chan CWH, Law BMH, So WKW, Chow KM, Waye MMY. Novel Strategies on Personalized Medicine for Breast Cancer Treatment: An Update. *Int J Mol Sci*. 2017;18(11).
11. Nalejska E, Maczynska E, Lewandowska MA. Prognostic and predictive biomarkers: tools in personalized oncology. *Molecular diagnosis & therapy*. 2014;18(3):273-84.
12. Ravelli A, Reuben JM, Lanza F, Anfossi S, Cappelletti MR, Zanotti L, et al. Breast cancer circulating biomarkers: advantages, drawbacks, and new insights. *Tumour Biol*. 2015;36(9):6653-65.
13. Berghuis AM, Koffijberg H, Prakash J, Terstappen LW, MJ IJ. Detecting Blood-Based Biomarkers in Metastatic Breast Cancer: A Systematic Review of Their Current Status and Clinical Utility. *Int J Mol Sci*. 2017;18(2).
14. Lampignano R, Berghuis AMS, Fehm T, Neubauer H. Chapter 28 - Role of Circulating Tumor Cells in Breast Cancer Patients. In: Dammacco F, Silvestris F, editors. *Oncogenomics*: Academic Press; 2019. p. 383-92.
15. Bidard FC, Fehm T, Ignatiadis M, Smerage JB, Alix-Panabieres C, Janni W, et al. Clinical application of circulating tumor cells in breast cancer: overview of the current interventional trials. *Cancer metastasis reviews*. 2013;32(1-2):179-88.
16. Bidard F-CJW, Dureau S, Brain E, Bachelot T, Bourgeois H, Goncalves A, Ladoire S, Naman H, Dalenc F, Gligorov J, Espie M, Levy C, Ferrero J-M, Loirat D, Cottu P, Dieras V, Simondi C, Berger F, Alix-Panabieres C, Pierga J-Y. Clinical utility of circulating tumor cell count as a tool to chose between first line hormone therapy and chemotherapy for ER+ HER2- metastatic breast cancer: Results of the phase III STIC CTC trial [abstract]. *Proceedings of the 2018 San Antonio Breast Cancer Symposium*; 2018 Dec

- 4-8; San Antonio, TX. Philadelphia (PA): AACR; Cancer Res 2019;79(4 Suppl):Abstract nr GS3-07. 2018.
17. Jonsson B, Hofmarcher T, Lindgren P, Wilking N. The cost and burden of cancer in the European Union 1995-2014. *Eur J Cancer*. 2016;66:162-70.
18. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II-An ISPOR Good Research Practices Task Force report. *Value Health*. 2015;18(2):161-72.
19. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*. 1993;85:365-76.
20. Bosmans JE, de Bruijne MC, van Hout HP, Hermens ML, Ader HJ, van Tulder MW. Practical guidelines for economic evaluations alongside equivalence trials. *Value Health*. 2008;11(2):251-8.
21. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. *Value Health*. 2013;16(1):202-10.
22. van Buuren S, Groothuis-Oudshoorn CGM. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*. 2011;45(3).
23. Longworth L, Yang Y, Young T, Mulhern B, Hernandez Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18(9):1-224.
24. Woodcock F, Doble B, Consortium C. Mapping the EORTC-QLQ-C30 to the EQ-5D-3L: An Assessment of Existing and Newly Developed Algorithms. *Med Decis Making*. 2018;38(8):954-67.
25. Statistiek CBvd. Consumentenprijzen; prijsindex 2015 = 100. 2019.

APPENDIX A; COST ALLOCATIONS

Table 1: Costs allocated to treatment, hospital stays, imaging and diagnostics

Type of costs	Costs accounted	Unit	Source
Alkeran	€ 402.67	Cycle	French Healthcare Authority (list prices)
Capecitabin	€ 402.67	Cycle	
Carboplatin	€ 402.67	Cycle	
Cyclophosphamide	€ 402.67	Cycle	
Docetaxel	€ 402.67	Cycle	
Doxorubicin	€ 402.67	Cycle	
Doxo liposomal	€ 402.67 + € 1320.90	Cycle	
Epirubicin	€ 402.67	Cycle	
Eribulin	€ 402.67 + € 1504 .00	Cycle	
Fluorouracil	€ 402.67	Cycle	
Methothrexate	€ 402.67	Cycle	
Paclitaxel	€ 402.67	Cycle	
Vinorelbin	€ 402.67	Cycle	
Other chemo	€ 402.67	Cycle	
Consult	€ 28.00	One appointment	
Tamoxifen	€ 3.34	Month	
Anti Aromastase :Letrozole	€ 41.71	Month	
LH-RH antagonist: Zoladex implant	€ 109.49	Month	
Fulvestrant	€ 926.68	Month	
Bevacizumab	€ 402.67 + € 2130.00	Cycle	
Everolimus	€ 402.67	Cycle	
Lapatinib	€ 402.67		
Other targeted therapy: Palbociclib	€ 2860.44	Month	
Aranesp	€ 211.00	Cycle	
Eprex	€ 143.78	Cycle	
Neorocormon	€ 820.60	Cycle	
Lenograstim	€ 419.80	Cycle	
Filgrastim	€ 560.10	Cycle	
Pegfilgrastim	€ 596.92	Cycle	

Intensive care department	€ 2,748.63	Day	Dutch Healthcare
Conventional hospital department	€ 1,473.44	Day	Authority (list prices)
Daycare Hospital	€ 370.92	Day	
Home treatment	€ 139.11	Day	
Home care	€ 84.63	Day	
X-ray	€ 21.28	Scan	French Healthcare
Echo	€ 52.45	Scan	Authority (list prices)
Scintigraphy	€ 168.71	Scan	
PET scan	€ 1,089.54	Scan	
CT scan	€ 183.54	Scan	
MRI	€ 240.09	Scan	
Radiotherapy	€ 4,386.44	Full treatment	
Cost CTC Test	€ 450.00	Test	Real costs (derived from one of the centers participating in the STIC trial)

CHAPTER 9

GENERAL DISCUSSION

DISCUSSION

With the ongoing pressure on healthcare budgets and the high cost of anti-cancer treatment, there is a growing need to rationalize healthcare spending in relation to the additional value achieved. Health technology assessment (HTA) is the discipline of transparently and systematically evaluating the clinical and economical consequences of new strategies in healthcare. HTA typically focuses on the societal perspective by explicitly assessing the balance between the added value of a new treatment or technology and its incremental cost, in comparison with usual care. While traditional HTA focuses foremost on evaluating cost-effectiveness [1], recent value frameworks, launched by the European Society of Medical Oncology (ESMO) and by the American Society of Clinical Oncology (ASCO) focus on improving clinical decisions on drug treatment. The ESMO Magnitude of Clinical Benefit Scale (ESMO-MCBS) aims to rank cancer drugs according to the magnitude of their clinically meaningful benefit, which is mainly determined by weighing survival gain against adverse events associated with the particular drug [2]. ASCO also developed a Value Framework (ASCO-VF) that facilitates discussions between clinicians and patients to make the most valuable treatment decisions [3]. Other initiatives focus on developing new diagnostic strategies to tailor targeted treatment, improve clinical outcomes and reduce costs [4]. Though these value frameworks sometimes consider costs, their main purpose is to inform clinical decisions.

The main aim of this thesis was to address the value of molecular diagnostics and biomarkers to improve the clinical outcomes and the cost-effectiveness of breast cancer treatment. In the first part of this thesis we have studied several aspects of usual breast cancer care. This included the availability of data on the coverage, the administration and cost of drug treatment and the performance of traditional molecular diagnostics in clinical practice. In the second part, we have further investigated the potential of liquid biopsies by evaluating the health economic impact of circulating tumor cells (CTCs), which are currently the most frequently investigated type of liquid biopsies. All together, these studies identified several challenges regarding various aspects of addressing the health economic potential of using molecular biomarkers derived from solid tissue material or liquid biopsies in breast cancer management.

RESEARCH INTO BIOMARKERS

Personalized medicine becomes more and more complex with the many new (combinations of) biomarkers, potentially serving as targets for either new or already

existing treatments. The evaluation of tumor heterogeneity and mechanisms of therapy resistance frequently requires a combination of biomarkers being investigated to inform clinical decisions on the prescription of therapies [5-7]. Nonetheless, most research on biomarkers still focuses on single biomarkers, as demonstrated for liquid biopsies in chapter 6 [8]. Despite the ease of detecting multiple biomarkers on the same tumor sample concurrently, in almost 70% of all studies on liquid biopsies in metastatic breast cancer the research focus was on one single biomarker, which were mainly circulating tumor cells (CTCs). Focusing on single (companion) biomarkers may be cheaper but is noteworthy in this context, as one specific platform is often able to detect multiple biomarkers simultaneously using the same tumor or blood sample.

Although several biomarkers are currently available, many are not (yet) used in clinical practice [9]. Furthermore, the study discussed in chapter 5 has shown that there is emphasis on the discovery of new biomarkers or innovative diagnostic methods, which are usually classified as proof-of-concept studies. However, while it was concluded that the number of studies reporting on clinical utility is limited, publication bias could not be excluded, as evidence demonstrating ineffectiveness of detecting particular (combinations of) biomarkers is usually not published [10]. Consequently, the number of performed validation studies may have been underestimated. Nonetheless, many published biomarker studies following proof-of-concept studies frequently lack to sufficiently focus on the validation of the particular diagnostic technologies [8].

One of the potential reasons underlying the limited number of validation studies may be that it is technically and ethically challenging to frequently obtain new or different tumor- or patient materials by repeatedly taking biopsies. Therefore, studies may regularly end up including specific subgroups of patients for whom the biomarker under investigation may yield most benefit. Especially in relatively small subgroups of patients, validating technologies and generalizing the derived results is complex. A potential solution to overcome these difficulties may be to use biobanks to store patient materials and subsequently facilitate the inclusion of more and diversified samples in new studies [11, 12]. Although the National Cancer Institute (NCI) has already established an overview of best practices for biobanks, there are no detailed regulations (yet) on the establishment, maintenance, standardization in the information recorded, or financial challenges and responsibilities of such biobanks [13]. To be able to overcome such challenges, additional regulations should first be established [14, 15].

Using new technologies that are able to detect multiple gene aberrations significantly increases the number of potentially suitable therapeutic options. Tumor heterogeneity may be discovered over time or between multiple lesions which even further complicates treatment decisions. While tumor heterogeneity is thought of as being one of the major complexities in the treatment of cancer, currently single measurements for ER and HER2 are performed for the majority of patients with breast cancer [16]. Measurements of single biomarkers at one timepoint do, by definition, not allow to capture potential tumor heterogeneity over time between the receptors tested on the tumor needle biopsy and the tumor resection. Furthermore, performing single tests to detect a similar number of biomarkers sequentially instead of simultaneously, may increase the time to diagnosis [17]. Nonetheless, given the rapid developments in next-generation sequencing (NGS) and whole genome sequencing (WGS) and the fact that single biomarker tests are currently not repeated in the majority of patients, it seems to be more valuable to focus on determining the impact of NGS, WGS or composite biomarkers (combined measurements for multiple biomarkers) instead of determining the clinical utility of single biomarkers.

Even though single biomarker tests are relatively inexpensive, no detailed clinical guidelines describe for which patients these tests should be performed and whether performing multiple tests is valuable for particular patient groups. Recently performed studies focusing on liquid biopsies are also mainly focusing on detecting the impact of single biomarkers [8]. As single biomarker tests are currently not repeated in the majority of patients, this may imply that there is either no reason to believe that heterogeneity may arise in the majority of breast cancer patients, or that it is currently unknown whether or how the corresponding treatment decisions should change if heterogeneity would be discovered.

FROM VALIDATION TO CLINICAL PRACTICE

Even though it is challenging to demonstrate the potential value of diagnostics in validation studies, it is mandatory to perform such studies to ensure that using the diagnostic in clinical practice is sufficiently reliable to make decisions regarding patient outcomes, safety and drug dosing [18, 19]. Randomized controlled trials (RCTs) still are the gold standard to validate pharmaceutical products, as these eliminate or reduce selection bias and the likelihood of spurious results due to confounding. However, RCTs may not be the most appropriate research design for evaluating diagnostics, as these do not necessarily account for user skills or other intra-user-variability aspects, which may directly influence the technology's impact.

Designing and performing RCTs to evaluate the value of (new) diagnostic technology may be even more challenging when the test results substantially impact the use of particular treatments [20]. Physicians may have different opinions on the value of the technology under investigation, which may cause them to be reluctant to base actual treatment decisions on outcomes of new diagnostics in early observational trial phases. In addition, there currently is no consensus on the strength of new evidence that should be generated to conclusively demonstrate the performance of new diagnostic technology, or validate them [21].

In many diagnostic studies, thresholds are to some extent arbitrarily selected to distinguish patient groups. The same applies to research on CTCs, for which typically a threshold of 1 or 5 CTCs per 7.5 mL of blood is being used. In a single RCT, typically only one threshold for a single marker or subset of markers is evaluated. To generate evidence on the clinical utility of multiple thresholds simultaneously, i.e. use a threshold of 1, 3 and 5 CTCs per 7.5 mL of blood to guide treatment, a large number of patients should be included in a single substantially large RCT or in multiple studies, as every threshold value considered requires an additional trial arm. However, financial resources are often scarce and restrict the number and size of studies that can be performed. As it is not feasible to validate all potentially valuable biomarkers in time consuming and expensive RCTs, a delay in implementing new diagnostics in clinical practice is caused. Even though the FDA already established a biomarker qualification program and provided guidance on drug development tools to improve on the timeliness of implementing new drugs in clinical practice [22], more innovative approaches in evaluating diagnostics and the extent to which technology should be fully validated may further improve the timeliness of innovations in healthcare.

Though the technology used to detect CTCs was validated and CTCs have been demonstrated to have prognostic value in cohort studies, these are not (yet) used in clinical practice. One reason is that, even though the prognostic value of CTCs has been demonstrated, their clinical utility to improve outcomes remains unclear. While most evidence on the clinical utility is usually generated following technology development, the uptake of technology is better when the potential clinical utility is already maximized during the development process [23, 24]. However, the interpretation and evaluation of clinical utility is being complicated by the existence of multiple definitions, as it is sometimes referred to as being the relevance and usefulness of a technology [23], while other times being described as the extent to which use of the test results in improved health outcomes [25]. Consequently, the definition of clinical utility evaluated in research frequently does not align with the

definition of clinical utility used to justify the implementation of new technology in clinical practice. To ensure that future studies on CTCs, liquid biopsies or any other molecular biomarkers can adequately address the clinical utility, it is of importance that uniform and generally accepted definitions of clinical utility will be generated. However, despite the differences in currently used definitions, all include the comparison with usual care. Regardless of the particular definition of clinical utility that is applied, assessing the clinical utility of new biomarkers will always at least require an adequate assessment of the performance of traditional diagnostics.

THE VALUE OF REAL WORLD DATA TO EVALUATE MOLECULAR DIAGNOSTICS

A less expensive and less time consuming alternative to evaluating the performance of diagnostic technology in prospective RCTs, would be to use real world data to produce RCT-like equivalent patient groups [26]. Real world data is most often referred to as data that is collected outside the context of RCTs in observational studies or routine clinical practice [27, 28]. Registries collect data from routine clinical practice for each individual patient and may for example be used to extract real world data to evaluate the performance of diagnostics, costs or the effectiveness of a treatment strategy in different subgroups of patients. Such registries can for example be hospital databases, insurance records or patient registries. As data from registries have been previously collected and are easily accessible, it is relatively cheap and fast to use such registries to evaluate clinical practice [29]. More specifically, data extracted from registries enables to determine the clinical utility of already implemented molecular diagnostics. Furthermore, while RCTs are usually not able to capture intra-user and inter-user variability, this may be assessed when using RWD. In addition, registries offer the potential to address questions for rare diseases or small subgroups of patients, as data of substantial numbers of patients can be collected over time [30]. Furthermore, information collected in registries is often standardized, which simplifies data analyses and improves the process of linking different databases [31, 32]. Subsequently, such comprehensive linked databases may be used to investigate even more complex research questions [33].

In the Netherlands, several highly standardized registries of the different specialisms that collaborate in breast cancer management have already been established. The nationwide network and registry of histo- and cytopathology in the Netherlands (PALGA) can be linked to the Netherlands Cancer Registry (NCR), which contains treatment and survival information, but it can also be linked to various hospital registries in which data on costs or resource use within that particular hospital are

collected [34, 35]. In chapter 3 and chapter 4 of this thesis it was demonstrated that it is feasible to use linked data extracted from the NCR and PALGA to evaluate the outcomes of traditional molecular diagnostics in breast cancer management [36].

Another advantage of using registries are valuable to evaluate the performance of diagnostics that are already embedded in clinical practice, is that the impact of technologies on a larger patient population than investigated in a clinical trial can be addressed. As registries collect data from daily clinical practice, these can subsequently be used to assess the quality or diagnostic performance of the current state of healthcare. Therefore, registries are a valuable source of information that enable to frequently monitor the quality of care by evaluating the performance of current diagnostic pathways. Nonetheless, such evaluations of the performance of traditionally used diagnostics can also be very valuable in the context of innovative research investigating the clinical utility of new biomarkers or other diagnostic technology. Although the safety of new diagnostics should always be evaluated in prospective settings, earlier evaluation of new diagnostics in clinical practice instead of in isolated RCT settings would enable real world registry platforms to further monitor the performance or quality of these diagnostics, which may be especially valuable for low risk biomarkers or diagnostics that are mainly intended for exploratory use.

Real world settings may generate valuable evidence that might support the validation of new technologies, as currently no highly regulated technology is allowed to be used in clinical practice until it is extensively validated and its cost-effectiveness is assessed. Shifting from research settings such as highly controlled RCTs to determine the clinical validity or clinical utility towards real world clinical settings may also provide more insight in inter- and intra-user variability and consequently enables to address the real health economic performance of using new technology in clinical practice. In addition, as several highly standardized registries are already available in the Netherlands, such a shift may also enable better use of the resources that are already present. Future research should focus on developing new methods to evaluate the validity and clinical utility of new technology, potentially by using real world data registries, to improve on the timeliness of innovations in healthcare.

Although several of challenges concerning the establishment, linking and use of registries remain, the existence of such registries has already proven several times to be valuable for multiple research purposes worldwide. Nonetheless, establishing collaborations between national and international registries may even further improve the quality of the data that is being collected [33, 37, 38].

CHALLENGES OF USING REAL WORLD DATA

Though real world data is valuable to evaluate current practice, there are several remaining challenges that should be addressed. As the initial aim of establishing a registry usually is not to address research questions, potential clinically meaningful patient characteristics may not have been collected in the data stored in these registries. In addition, it is not obligatory for all types and settings of diseases to use information systems that are linked to data registries. Furthermore, standardizing the data over the years and between different registries is not straightforward [30, 39]. The guidelines that are used in clinical practice are likely to change over time, resulting in different subsets of the data. Similarly, protocols, modules or programs that are used to collect the data that are stored within these registries are also evolving and improving. However, as new guidelines and new data collection modules are not necessarily implemented simultaneously, data collected at different time points may differ (in level of detail and in structure) and subsequently become less standardized. Consequently, in certain registries some relevant data may be missing or may be of lower quality than desired.

While using data from registries to address specific predefined research questions, knowledge on the structure of the data and potential difficulties in the standardization of the data is subsequently gathered. In order to further improve the quality of the data collected in these registries, it is important to establish solid and persistent collaborations with the institutions managing such registries or databases. Furthermore, to enhance the quality and continuity of data collection in these registries, it is important that these institutions (remain to) receive substantial funding in order to do so [40]. In addition, to optimize the use of these registries, more attention should be paid to the visibility of those registries, or to the institutions managing those.

EARLY STAGE HEALTH ECONOMIC EVALUATION OF DIAGNOSTICS

Critical to routine implementation of new diagnostic tests is a demonstration of their clinical utility but also their cost-effectiveness. Common, trial-based, cost-effectiveness studies assess costs and quality adjusted life years (QALYs) as outcomes of interest. One of the first RCTs evaluating the clinical performance of using CTCs to guide treatment was the STIC METABREAST trial, that used CTCs to guide chemotherapy and hormone therapy in metastatic HER2 negative, ER positive breast cancer. The cost-effectiveness analysis of the STIC METABREAST trial was presented in chapter 7. While the trial has not been able to demonstrate favorable cost-effectiveness

outcomes in the base-case analysis, the bootstrap demonstrated that there is substantial uncertainty in the difference in health benefits and costs of the CTC based strategy compared to usual care. One of the potential causes of this uncertainty in health economic outcomes is that the study evaluates the performance of using CTCs on a short-term. In general, evaluating the clinical utility of new diagnostics that have an impact on invasive treatment by performing an RCT with short-term follow-up, will not reveal the total impact in terms of health outcomes and costs. In this context, health economic models can be used to estimate the total and long term impact, but these typically require assumptions on survival and may consequently also yield uncertain outcomes.

To calculate cost-effectiveness outcome measures, an estimation of the cost of using a particular technology should also be derived. Particularly when the technology that is being evaluated has an impact on the prescription of an expensive drug, it is important that the drug costs are incorporated in the analyses [41]. Therefore, estimates of the average usage of the drug in milligrams, its mode of administration, and the corresponding costs are necessary. Currently, these costs are not available in a standardized way in databases that are typically used to identify costs within scientific research [42]. It would be valuable and less time consuming to use data from a database that contains generic cost estimates per specific type of administration and a particular number of treatment cycles in cost-effectiveness research. Furthermore, even if standardized administration information would be available for each drug, addressing the exact cost of particular treatments remains challenging, as the only prices available for all drugs are the list prices. As hospitals may buy larger or multiple packages of particular drugs in order to be eligible for potential discounts, the actual cost price (from a hospital perspective) may be substantially lower than the list prices [43]. This further complicates the estimation of the cost-effectiveness of new diagnostic technologies.

THE POTENTIAL OF DECISION TOOLS IN NEW BIOMARKER DISCOVERY

Early health economic methods can be used to support decisions on the prioritization of research on new diagnostic technology, as these enable to identify room for improvement in current healthcare pathways, and provide insight in the health economic potential during the development [44]. In such analyses aggregated outcome measures are usually used, such as the average survival for a particular patient population [45]. Assumptions on aggregated outcome measures may influence the final outcomes of the analyses, and also diminish the applicability of the

analysis to other scenarios, i.e. when it is assumed in the model that patients receive a particular therapy if a particular biomarker expressed positive, the analysis cannot easily be updated to evaluate the consequences of prescribing a different therapy to these patients. The use of aggregated evidence derived from real world data instead of making assumptions based on aggregated evidence derived from clinical trials may be a valuable addition to model based health economic evaluations.

To be able to rapidly update model parameter values as soon as new evidence becomes available, and to update parameter values to reflect different scenario's, online decision tools may be used. These tools enable to get insight in the potential impact of different parameter values on the calculated outcome measures. Such decision tools are also suitable to adjust the outcomes to the user experiences or stakeholder preferences, insights, or simply a different clinical setting and may consequently facilitate discussion between policy makers, clinicians and technology developers on the optimal use of new technologies. These discussions may increase the extent to which the technology potentially will be aligned with the target setting before implementation, as user preferences may be better understood by the different stakeholders. In addition, when providing such tools to stakeholders such as the government or technology developers, these may yield valuable insights, especially in prioritization mechanisms. Selecting which studies should be performed is challenging in biomarker research, as investigating a particular technology or analyzing a particular biomarker may yield information which is also relevant to a different set of biomarkers or the prescription of a therapy that is based on the particular biomarker expression [46]. As it will never be feasible to test newly developed biomarkers in all potential settings, it might be valuable to estimate the potential health economic outcomes by using these types of decision tools [47]. In addition, especially in research on liquid biopsies, thresholds of biomarker expression are being used to classify whether the particular biomarker expressions are either present or absent. Assuming that these tools are based on the best available evidence, they may be suitable to facilitate the prioritization of the clinical research agenda when considering different thresholds of the expression a particular biomarker to distinguish different patient subgroups [48].

Although the main aim of decision tools is to make decision models available online, these tools may serve multiple additional purposes. Such online decision tools enable to improve on the transparency in complex model codes while not directly requiring advanced software knowledge [49]. In addition, these tools may create a more user friendly interface that may facilitate discussion amongst stakeholders in early

research phases, as multiple scenarios can be explored by easily updating parameter values. Creating the structure of the online tools based on decision models that are already developed will fundamentally not require major efforts. However, there still seems to be a lack of incentive to share models online, as model developers tend to label their models proprietary following the substantial amount of time and money invested in developing those models. Consequently, only a minority of the decision models are being fully presented online nowadays. In addition, simply making models available by means of an online tool presumably requires much more regulation on the responsibilities of the developers and users than currently present.

CONCLUSIONS

This thesis has demonstrated that the results of the traditional diagnostics used to determine tumor characteristics are often concordant between the tumor needle biopsy and tumor resection. However, one emerging development is the use of less invasive methods rather than the traditional biopsies to repeatedly assess potential tumor heterogeneity. While liquid biopsies are thought to be promising, this thesis demonstrates the majority of studies only poorly evaluate clinical utility of biomarkers in breast cancer. However, demonstrating the clinical utility is required before diagnostic tests will be implemented in clinical practice. The only study to date focusing on determining the clinical utility of new biomarkers in breast cancer, is the STIC METABREAST trial, of which the clinical outcomes have shown that using CTCs to guide hormone- and chemotherapy is non-inferior to current clinical practice in patients with metastatic breast cancer. Though the STIC METABREAST trial did not demonstrate favorable outcomes regarding the cost-effectiveness of using CTCs to guide hormone- and chemotherapy, this may be partly explained by the fact that only short term follow-up data could be collected. In addition, a predetermined threshold of 5 CTCs per 7.5 mL of blood was used to distinguish the different patient groups, which may not yield optimal health economic outcomes.

In future research on biomarkers, online decision tools can provide insight in the potential impact of new diagnostic tests and the impact of changing test cut-off values on the cost-effectiveness of using the biomarker guided strategy compared to usual care. Such decision tools provide insight in the potential cost-effectiveness and may therefore be used to prioritize biomarker development.

REFERENCES

1. Banta HD. PERSPECTIVE: SOME CONCLUSIONS FROM MY LIFE IN HEALTH TECHNOLOGY ASSESSMENT. *International journal of technology assessment in health care*. 2018;34(2):131-3.
2. Cherny NI, Sullivan R, Dafni U, Kerst JM, Sobrero A, Zielinski C, 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(8):1547-73.
3. Schnipper LE, Davidson NE, Wollins DS, Tyne C, Blayney DW, Blum D, et al. American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options. *J Clin Oncol*. 2015;33(23):2563-77.
4. Katz SJ, Jaggi R, Morrow M. Reducing Overtreatment of Cancer With Precision Medicine: Just What the Doctor Ordered. *Jama*. 2018;319(11):1091-2.
5. Aumeeruddy MZ, Mahomoodally MF. Combating breast cancer using combination therapy with 3 phytochemicals: Piperine, sulforaphane, and thymoquinone. *Cancer*. 2019;125(10):1600-11.
6. Bayat Mokhtari R, Homayouni TS, Baluch N, Morgatskaya E, Kumar S, Das B, et al. Combination therapy in combating cancer. *Oncotarget*. 2017;8(23):38022-43.
7. Leary M, Heerboth S, Lapinska K, Sarkar S. Sensitization of Drug Resistant Cancer Cells: A Matter of Combination Therapy. *Cancers*. 2018;10(12).
8. Berghuis AMS, Koffijberg H, Prakash J, Terstappen LWMM, IJzerman MJ. Detecting Blood-Based Biomarkers in Metastatic Breast Cancer: A Systematic Review of Their Current Status and Clinical Utility. *Int J Mol Sci*. 2017;18(2).
9. Frangiannis NG. Biomarkers: hopes and challenges in the path from discovery to clinical practice. *Translational research : the journal of laboratory and clinical medicine*. 2012;159(4):197-204.
10. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Annals of the New York Academy of Sciences*. 1993;703:135-46; discussion 46-8.
11. Roberts JN, Karvonen C, Graham K, Weinfeld M, Joy AA, Koebel M, et al. Biobanking in the Twenty-First Century: Driving Population Metrics into Biobanking Quality. *Adv Exp Med Biol*. 2015;864:95-114.
12. Somiari SB, Somiari RI. The Future of Biobanking: A Conceptual Look at How Biobanks Can Respond to the Growing Human Biospecimen Needs of Researchers. *Adv Exp Med Biol*. 2015;864:11-27.
13. Hallmans G, Vaught JB. Best practices for establishing a biobank. *Methods Mol Biol*. 2011;675:241-60.
14. Schmitt S, Kynast K, Schirmacher P, Herpel E. Challenges for quality management in implementation, maintenance, and sustainability of research tissue biobanks. *Virchows Archiv : an international journal of pathology*. 2016;468(1):93-9.
15. Quinlan PR, Groves M, Jordan LB, Stobart H, Purdie CA, Thompson AM. The Informatics Challenges Facing Biobanks: A Perspective from a United Kingdom Biobanking Network. *Biopreservation and biobanking*. 2015;13(5):363-70.
16. Berghuis AMS, van Deurzen CHM, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ. Real-world data on discordance between estrogen, progesterone, and HER2 receptor expression on diagnostic tumor

- biopsy versus tumor resection material. *Breast Cancer Res Treat.* 2019.
17. Rossing M, Sorensen CS, Ejlersen B, Nielsen FC. Whole genome sequencing of breast cancer. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica.* 2019;127(5):303-15.
18. Medical devices must be carefully validated. *Nature Biomedical Engineering.* 2018;2(9):625-6.
19. Kaza M, Karazniewicz-Lada M, Kosicka K, Siemiatkowska A, Rudzki PJ. Bioanalytical method validation: new FDA guidance vs. EMA guideline. Better or worse? *Journal of pharmaceutical and biomedical analysis.* 2019;165:381-5.
20. Ferrante di Ruffano L, Dinnes J, Sitch AJ, Hyde C, Deeks JJ. Test-treatment RCTs are susceptible to bias: a review of the methodological quality of randomized trials that evaluate diagnostic tests. *BMC medical research methodology.* 2017;17(1):35.
21. Akhmetov I, Bubnov RV. Assessing value of innovative molecular diagnostic tests in the concept of predictive, preventive, and personalized medicine. *The EPMA journal.* 2015;6:19.
22. Amur S, LaVange L, Zineh I, Buckman-Garner S, Woodcock J. Biomarker Qualification: Toward a Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. *Clin Pharmacol Ther.* 2015;98(1):34-46.
23. Lesko LJ, Zineh I, Huang SM. What is clinical utility and why should we care? *Clin Pharmacol Ther.* 2010;88(6):729-33.
24. Lesko LJ, Zineh I. DNA, drugs and chariots: on a decade of pharmacogenomics at the US FDA. *Pharmacogenomics.* 2010;11(4):507-12.
25. Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. *Clin Chem.* 2012;58(12):1636-43.
26. Gugiu PC, Gugiu MR. A critical appraisal of standard guidelines for grading levels of evidence. *Evaluation & the health professions.* 2010;33(3):233-55.
27. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. *Value Health.* 2017;20(7):858-65.
28. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, et al. Good Practices for Real-World Data Studies of Treatment and/or Comparative Effectiveness: Recommendations from the Joint ISPOR-ISPE Special Task Force on Real-World Evidence in Health Care Decision Making. *Value Health.* 2017;20(8):1003-8.
29. Jarow JP, LaVange L, Woodcock J. Multidimensional Evidence Generation and FDA Regulatory Decision Making: Defining and Using "Real-World" Data. *Jama.* 2017;318(8):703-4.
30. Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, et al. A checklist for retrospective database studies--report of the ISPOR Task Force on Retrospective Databases. *Value Health.* 2003;6(2):90-7.
31. Kennedy L, Craig AM. Global registries for measuring pharmaco-economic and quality-of-life outcomes: focus on design and data collection, analysis and interpretation. *Pharmacoeconomics.* 2004;22(9):551-68.
32. Larsson S, Lawyer P, Garellick G, Lindahl B, Lundstrom M. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value. *Health affairs (Project Hope).* 2012;31(1):220-7.

33. Kurian AW, Mitani A, Desai M, Yu PP, Seto T, Weber SC, et al. Breast cancer treatment across health care systems: linking electronic medical records and state registry data to enable outcomes research. *Cancer*. 2014;120(1):103-11.
34. van der Sanden GA, Coebergh JW, Schouten LJ, Visser O, van Leeuwen FE. Cancer incidence in The Netherlands in 1989 and 1990: first results of the nationwide Netherlands cancer registry. Coordinating Committee for Regional Cancer Registries. *Eur J Cancer*. 1995;31a(11):1822-9.
35. Casparie M, Tiebosch ATMG, Burger G, Blauwgeers H, van de Pol A, van Krieken JHJM, et al. Pathology databanking and biobanking in the Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cellular Oncology*. 2007;29:19-24.
36. Berghuis AMS, van Deurzen CHM, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ. Real-world data on discordance between estrogen, progesterone, and HER2 receptor expression on diagnostic tumor biopsy versus tumor resection material. *Breast Cancer Res Treat*. 2019.
37. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet (London, England)*. 2018;391(10125):1023-75.
38. Varella L, Eziokwu AS, Jia X, Kruse M, Moore HCF, Budd GT, et al. Real-world clinical outcomes and toxicity in metastatic breast cancer patients treated with palbociclib and endocrine therapy. *Breast Cancer Res Treat*. 2019;176(2):429-34.
39. Justo N, Espinoza MA, Ratto B, Nicholson M, Rosselli D, Ovcinnikova O, et al. Real-World Evidence in Healthcare Decision Making: Global Trends and Case Studies From Latin America. *Value Health*. 2019;22(6):739-49.
40. Wurzer P, Hundeshagen G, Cambiaso-Daniel J, Fischer S, Hoflehner H, Spendel S, et al. Lessons Learned From Breast Implant Registries: A Systematic Review. *Annals of plastic surgery*. 2019.
41. Mushlin AI, Ruchlin HS, Callahan MA. Costeffectiveness of diagnostic tests. *Lancet (London, England)*. 2001;358(9290):1353-5.
42. Berghuis AMS, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ. Evidence on the cost of breast cancer drugs is required for rational decision making. *Ecancermedicalscience*. 2018;12:825.
43. van Harten WH, Wind A, de Paoli P, Saghatchian M, Oberst S. Actual costs of cancer drugs in 15 European countries. *The Lancet Oncology*. 2016;17(1):18-20.
44. IJzerman MJ, Koffijberg H, Fenwick E, Krahn M. Emerging Use of Early Health Technology Assessment in Medical Product Development: A Scoping Review of the Literature. *PharmacoEconomics*. 2017;35(7):727-40.
45. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. *Value Health*. 2012;15(6):835-42.
46. Miquel-Cases A, Retel VP, van Harten WH, Steuten LM. Decisions on Further Research for Predictive Biomarkers of High-Dose Alkylating Chemotherapy in Triple-Negative Breast Cancer: A Value of Information Analysis. *Value Health*. 2016;19(4):419-30.

47. de Graaf G, Postmus D, Westerink J, Buskens E. The early economic evaluation of novel biomarkers to accelerate their translation into clinical applications. *Cost Effectiveness and Resource Allocation*. 2018;16(1):23.

48. Helissey C, Berger F, Cottu P, Dieras V, Mignot L, Servois V, et al. Circulating tumor cell thresholds and survival scores in advanced metastatic breast cancer: the observational step of the CirCe01 phase III trial. *Cancer letters*. 2015;360(2):213-8.

49. Sampson CJ, Arnold R, Bryan S, Clarke P, Ekins S, Hatswell A, et al. Transparency in Decision Modelling: What, Why, Who and How? *PharmacoEconomics*. 2019.

SUMMARY

FOUNDATION OF THESIS

Significant improvement in the delivery of breast cancer care has been made in the last decades. Several treatments that substantially improve patients' clinical outcomes became available on the market. Whether patients will respond to a particular treatment, is usually dependent on several tumor characteristics. Further improving the ability to identify tumor characteristics and thereby extending personalized medicine in oncology, may support further improvements in breast cancer care and potentially reduce over- and under treatment or decrease costs.

Several initiatives have focused on the use of prognostic or predictive molecular biomarkers that may guide the treatment of breast cancer patients. In current clinical practice, tissue biopsies are used to determine particular biomarker expression and characterize the tumor. However, obtaining these tissue biopsies is very invasive and consequently may not be suitable to repeatedly assess biomarker expression. Furthermore, due to potential sampling errors and tumor heterogeneity, new technologies to repeatedly assess the tumor expression and better capture the full tumor profile may be valuable. An emerging and innovative field aiming to develop a new technique that is less invasive is focusing on the use of liquid biopsies.

Several studies have yet evaluated the prognostic and predictive value of different types of liquid biopsies. Such innovative techniques can, however, only be successfully implemented in clinical practice when their value compared to usual care has been demonstrated in clinical trials. Furthermore, as diagnostics may influence the prescription of drugs, it is important that the cost-effectiveness of such new clinical pathways is addressed. In this thesis, the health economic value and the clinical utility of using molecular biomarkers in breast cancer management have been addressed.

SUMMARY

To be able to evaluate the clinical utility of new diagnostics, their performance and health economic impact should be compared to the performance of usual care. Therefore, in the first part of this thesis, several components of usual breast cancer care were evaluated.

In chapter 2, evidence on the reimbursement and costs of breast cancer drugs was synthesized for the Netherlands. It was demonstrated which breast cancer drugs are available in the Netherlands and the extent to which generalized prescription is available for these drugs. Though health insurance covers almost all breast

cancer drugs that can be used in the Netherlands, an important finding was that off-patent generics, such as paclitaxel and doxorubicin, still seem to be relatively costly. However, these findings were based on list prices, which may deviate from the actual prices paid for these drugs. Furthermore, these list prices are established for particular forms and strengths of the drug substance. In health economic analyses, the costs of using particular drugs are frequently used and derived from platforms that are easily accessible for health economists. Therefore, it is preferred to have dosing information, such as the form and strength of the drug substance that should be prescribed, available in a standardized way on a platform that is easily accessible for health economists. However, such prescription information is frequently not available in the preferred standardized way for health economic research.

The prescription of drugs in usual care is mainly dependent on the expression of several receptors that can be tested from histologic material derived by tumor needle biopsies or the full tumor resection. Tests that are regularly performed on these specimens are, for example, the immunohistochemistry and in situ hybridizations, which may provide information on the estrogen (ER), progesterone (PR) and human epidermal growth factor receptor 2 (HER2) status.

In chapters 3 and 4, the frequency of discordance in the ER, PR and HER2 status between the tumor needle biopsy and the tumor resection was evaluated. In addition, the extent to which such potential discordance influences the prescription of particular types of adjuvant treatment was addressed. The studies performed have demonstrated that in only a small subgroup of patients, the ER, PR and HER2 status are determined on both, tumor needle biopsy and tumor resection material. When these tests were performed on both tumor materials, the test results are concordant in the majority of patients. For those patients with concordant test results, the treatment that is prescribed usually matches the expected type of treatment based on these receptor expressions according to the guideline. For a minority of patients, a positive expression of ER or HER2 was not followed by the expected prescription of endocrine- or targeted therapy.

Even though the tests that influence the current treatment prescriptions seem to generally have a good performance, these can only be performed on tumor needle biopsy or tumor resection material, which are both very invasive to derive. To be able to frequently check for potential mutations, without having to use these invasive tests, new technologies may be needed. Liquid biopsies are an emerging and innovative field of new technology which aim to offer this potential. A substantial number of

studies have been performed, which evaluated the performance and potential of several biomarkers, cells or other contents of fluids that may be derived from the human body. It was, however, unknown to which extent these liquid biopsies have been investigated and to which developmental research phase these developments were currently.

In chapter 5, a systematic review was presented in which it was evaluated what type of liquid biopsies have been investigated in metastatic breast cancer patients so far. In addition, for each type of liquid biopsy, or specific biomarker, the research phase in which the studies were performed was determined. The study has demonstrated that most attention has been paid to research investigating circulating tumor cells. Furthermore, research on circulating tumor cells also reached the latest phase in technology development. In general, it was found that studies on liquid biopsies mainly investigated their clinical potential whilst their health economic impact frequently remained unaddressed.

In chapter 6, we have assessed the health economic potential and remaining challenges of liquid biopsies, in general, in an expert review. We have discussed the increasing number of studies and the potential change in clinical management that these liquid biopsies may facilitate. Though liquid biopsies seem to be very promising, their health economic potential is not (yet) clearly addressed. To address the health economic potential from an early stage, health economic expertise is necessary. As it is time-consuming to repeatedly update health economic outcomes when new evidence becomes available during the development process, health economic model structures can be implemented in online tools that enable to easily and frequently update model parameters.

The clinical validity and utility of biomarkers are frequently dependent on a predefined threshold that is used to classify patients in multiple subgroups based on the expression of the particular biomarker. As evaluating all possible thresholds in clinical trials is not suitable, in chapter 7, an online tool was developed to provide insight into the clinical utility and health economic outcomes beyond the clinical validity of using multiple thresholds for circulating tumor cells. It was demonstrated that it is feasible to use such online tools to explore the uncertainty and discuss the impact of other potential scenario's which circulating tumor cells may be used. Results derived from such tools may facilitate the discussions to prioritize future research.

However, before any of those diagnostics will be used in clinical practice, randomized clinical trials are necessary to demonstrate their (cost-) effectiveness. The clinical

results of the STIC METABREAST trial have demonstrated that using circulating tumor cells to guide endocrine- and chemotherapy is non-inferior to clinical practice for metastatic breast cancer patients. However, the health economic analysis in chapter 8 showed that it is unlikely that using circulating tumor cells is cost-effective on the short term. Even though a substantial number of studies have yet demonstrated the value of liquid biopsies and circulating tumor cells, more attention should especially be paid to their health economic potential. In addition, the measures by which diagnostics are currently evaluated may have to be improved before the real health economic potential of such diagnostics can be demonstrated.

CONCLUSIONS

In this thesis, it was demonstrated that traditional diagnostics are less frequently performed on tumor needle biopsy and tumor resection material than expected. When performed on both materials, the results of the traditional diagnostics used to determine tumor characteristics are usually concordant. As these traditional diagnostics are invasive and therefore not suitable to frequently obtain tumor materials, new methods are being developed to assess potential tumor heterogeneity. While liquid biopsies are thought to be promising, the majority of studies focusing on liquid biopsies is not sufficiently focusing on evaluating their clinical utility, which is necessary before such new methods will be implemented in clinical practice. To date, the clinical utility of new liquid biopsies is only evaluated in the STIC METABREAST trial. Though the clinical outcomes have shown that using circulating tumor cells to guide either chemotherapy or endocrine therapy in metastatic breast cancer is non-inferior, the health economic outcomes have demonstrated that this strategy is not cost-effective when considering short term outcomes.

SAMENVATTING

TOTSTANDKOMING PROEFSCHRIFT

Er is enorme vooruitgang geboekt in de behandeling van borstkanker de laatste decennia. Verschillende behandelingen die de klinische patiënt uitkomsten significant verbeteren zijn op de markt gekomen. Of een patiënt goed reageert op een bepaalde behandeling is afhankelijk van bepaalde tumor karakteristieken. Het verder verbeteren van de mogelijkheden om dergelijke tumor karakteristieken te identificeren en daarmee gepersonaliseerde oncologische zorg nog verder te verbeteren, zorgt er voor dat ook de totale borstkankerszorg verbetert en dat mogelijk overbehandeling, onderbehandeling en kosten kunnen worden gereduceerd.

Verschillende projecten hebben ingezoomd op het gebruik van prognostische of predictieve moleculaire biomarkers die de behandeling van borstkanker patiënten zouden kunnen sturen. In de huidige klinische praktijk, worden biopten genomen om de expressie van bepaalde biomarkers te bepalen en zo de tumor te karakteriseren. Echter, het nemen van deze biopten is erg invasief en om die reden zijn deze dan ook niet bruikbaar om herhaaldelijk biomarker expressie te bepalen. Daarnaast kan het zo zijn dat de tumor expressie niet goed wordt bepaald door ‘sampling errors’ of tumor-heterogeniteit. Nieuwe technologieën die mogelijk beter in staat zijn om herhaaldelijk tumor expressie te bepalen en het volledige tumor-profiel in kaart te brengen zouden om die reden waardevol kunnen zijn. Een opkomend innovatief veld dat een nieuwe technologie probeert te ontwikkelen die minder invasief is, focust voornamelijk op het gebruik van de zogenoemde ‘liquid biopsies’, ofwel vloeibare biopten zoals bijvoorbeeld bloedtesten.

Verschillende studies hebben reeds gekeken naar de prognostische en predictieve waarde van die bloedtesten. Dergelijke innovatieve technieken kunnen echter alleen maar succesvol in de klinische praktijk worden geïmplementeerd, wanneer de waarde van deze testen ten opzichte van de huidige praktijk is aangetoond in klinische trials. Daarnaast is het mogelijk dat diagnostiek invloed kan hebben op het voorschrijven van bepaalde medicijnen. Het is daarom erg van belang dat ook de kosten-effectiviteit van dergelijke nieuwe klinische paden wordt geadresseerd. In dit proefschrift is de gezondheidseconomische waarde en de klinische utiliteit van het gebruik van moleculaire biomarkers in borstkanker zorg geadresseerd.

SAMENVATTING

Om de klinische utiliteit van nieuwe diagnostiek te kunnen bepalen, moet de ‘performance’ en gezondheidseconomische impact in ieder geval worden vergeleken

met de huidige klinische praktijk. In het eerste deel van dit proefschrift zijn daarom verschillende componenten van de huidige borstkankerzorg bekeken. In hoofdstuk 2 hebben we informatie over de vergoedingen en kosten van borstkanker medicatie samengevat voor Nederland. In dit onderzoek is bekeken welke borstkanker medicatie allemaal beschikbaar is in Nederland. Daarnaast is geëvalueerd in welke mate algemene informatie over het voorschrijven van deze medicatie beschikbaar is. De zorgverzekeraars dekken in Nederland bijna alle borstkanker medicatie die gebruikt mag worden. Een belangrijke bevinding uit dit onderzoek is echter dat ook generieke medicatie zoals paclitaxel en doxorubicine, waarvoor het patent reeds is verlopen, nog steeds erg kostbaar lijken.

Echter, deze bevindingen zijn gebaseerd op lijst-prijzen (maximum-prijzen die gehanteerd mogen worden voor medicatie, bepaald door de overheid) en deze kunnen afwijken van de daadwerkelijke prijzen die betaald worden voor deze medicatie. Daarnaast zijn die lijstprijzen beschikbaar voor bepaalde verpakkingen van de substantie van het medicijn. In gezondheidseconomische analyses worden de kosten van het gebruik van een bepaald medicijn vaak meegenomen. Deze worden verkregen vanuit een platform dat toegankelijk is voor gezondheidseconomen. Het is daarom wenselijk dat doseringsinformatie ook op een gestandaardiseerde manier beschikbaar is op eenzelfde soort platform. Echter blijkt dergelijke informatie over de dosering nu vaak nog niet beschikbaar te zijn in een gestandaardiseerde vorm die gewenst is voor gezondheidseconomisch onderzoek.

Het voorschrijven van medicijnen is in de huidige klinische praktijk voornamelijk nog afhankelijk van de expressie van verschillende receptoren, of biomarkers, die momenteel bepaald worden door testen te doen met histologisch (weefsel-) materiaal dat verkregen wordt door een biopsie of de volledige tumor-resectie (de volledige tumor verkregen na chirurgie). Testen die regelmatig op weefsel gedaan worden, zijn bijvoorbeeld de immunohistochemie en de in situ hybridisatie. Deze testen geven informatie over de expressie van oestrogeen (ER), progesteron (PR) en humaan epidermale groeifactor receptor 2 (HER2).

In hoofdstuk 3 en 4 is de frequentie van discordantie in de ER, PR en HER2 status tussen het tumor biopsie en de tumor resectie geëvalueerd. Daarnaast is er bepaald in welke mate mogelijke discordantie in deze receptoren het voorschrijven van bepaalde typen adjuvante behandeling (behandeling na chirurgie) beïnvloedt. De onderzoeken die gepresenteerd zijn in deze hoofdstukken laten zien dat er enkel in kleine subgroepen van patiënten een test wordt gedaan voor ER, PR en HER2 op

zowel het biopt als de tumor-resectie. Wanneer deze testen zijn gedaan op beide materialen, zijn de test-resultaten concordant voor de meerderheid van patiënten. Voor patiënten met concordante test-resultaten werd ook gezien dat de behandeling die voorgeschreven is vaak overeenkomt met het type behandeling dat verwacht kan worden op basis van de receptor-expressie volgens de richtlijn. Voor een kleine minderheid van patiënten, werd een positieve expressie van ER of HER2 niet opgevolgd door het voorschrijven van een endocriene of doelgerichte therapie.

Ondanks het feit dat de testen die mogelijk invloed hebben op de behandeling die wordt voorgeschreven een relatief goede performance hebben, kunnen deze enkel worden uitgevoerd wanneer men de beschikking heeft over weefsel verkregen uit een biopt of de tumor resectie, welke beide dus erg invasief zijn om te verkrijgen. Om frequenter te kunnen onderzoeken of er mogelijk mutaties zijn opgetreden in de tumor, zonder dat deze invasieve testen gedaan hoeven worden, is er nieuwe technologie nodig. De liquid biopsies zijn een opkomend en innovatief veld dat de mogelijkheid biedt om frequenter te testen. Een substantieel aantal studies is reeds uitgevoerd, die hebben gekeken naar de performance van verschillende biomarkers, cellen of andere deeltjes die verkregen kunnen worden uit vloeistoffen in het lichaam. Het was echter onbekend in welke mate deze liquid biopsies onderzocht zijn en in welke onderzoeksfase zij zich momenteel bevinden.

In hoofdstuk 5 is er een systematische review gepresenteerd waarin is geëvalueerd welk type liquid biopsies er tot op heden onderzocht zijn in gemetastaseerde borstkanker. Daarnaast is er voor elk specifiek type liquid biopsy, of specifieke biomarker, bepaald in welke onderzoeksfase het onderzoek naar die biomarker zich bevond. Deze studie heeft laten zien dat de meeste aandacht is besteed aan circulerende tumorcellen. Daarnaast is aangetoond dat het onderzoek naar circulerende tumorcellen zich ook in de verste fase in het gehele ontwikkelings-proces bevond. Over het algemeen hebben studies zich voornamelijk gericht op het onderzoeken van de klinische mogelijkheden van dergelijke biomarkers. Het gezondheidseconomische potentieel van dergelijke biomarkers bleef daarbij vaak onderbelicht.

In hoofdstuk 6 hebben we gekeken naar het gezondheidseconomische potentieel en verdere uitdagingen van liquid biopsies in het algemeen, in een zogenoemde expert review. We hebben daarin het toenemende aantal studies naar deze liquid biopsies en hun potentieel om het klinische management te veranderen bediscussieerd. Ook al lijken liquid biopsies erg veelbelovend, hun gezondheidseconomische potentieel is dus nog niet goed genoeg aangetoond. Om dat potentieel al in een vroeg stadium

in de ontwikkeling te laten zien, is er gezondheidseconomische expertise nodig. Het is echter wel erg tijdrovend om herhaaldelijk gezondheidseconomische uitkomsten te berekenen wanneer er nieuw bewijs beschikbaar komt gedurende het proces van het ontwikkelen van een dergelijke test. Gezondheidseconomische modelstructuren waarmee deze uitkomsten berekend worden kunnen daarom worden geïmplementeerd in online tools. Op die manier kunnen de parameters van dergelijke modellen makkelijk en herhaaldelijk worden aangepast.

De klinische validiteit en utiliteit van biomarkers zijn vaak afhankelijk van een vooraf bepaalde afkapwaarde die gebruikt wordt om patiënten op basis van de expressie van een bepaalde biomarker in meerdere subgroepen in te delen. Het is niet mogelijk om al deze afkapwaarden in een klinische trial te onderzoeken. In het onderzoek dat in hoofdstuk 7 is gepresenteerd is er een online tool gepresenteerd waarmee het mogelijk is om deze afkapwaarden aan te passen voor circulerende tumorcellen. Op deze manier kon voor meerdere waarden inzicht gegeven worden in de mogelijke klinische utiliteit en gezondheidseconomische uitkomsten. Dit onderzoek heeft laten zien dat het haalbaar is om dergelijke online tools te gebruiken om de onzekerheid in uitkomsten te bekijken en om de impact van mogelijke scenario's te bediscussiëren. Dergelijke inzichten kunnen input zijn voor vervolgonderzoek.

Voordat nieuwe diagnostiek in de praktijk geïmplementeerd kan worden is het echter noodzakelijk om de kosten-effectiviteit van deze diagnostiek aan te tonen door middel van klinische trials. De resultaten van de STIC METABREAST trial hebben laten zien dat het gebruik van circulerende tumorcellen om de behandelbeslissing rondom endocriene- of chemotherapie voor patiënten met gemetastaseerde borstkanker te sturen, niet inferieur is aan de huidige klinische praktijk. De gezondheidseconomische analyse heeft echter laten zien dat het gebruiken van circulerende tumorcellen niet kosten-effectief is op korte termijn. Ook al zijn er al een aantal studies die reeds hebben gekeken naar de waarde van de liquid biopsies en de circulerende tumorcellen, er moet nog steeds meer aandacht worden besteed aan het gezondheidseconomische potentieel. Daarnaast zijn de maten waarmee diagnostiek momenteel wordt geëvalueerd wellicht minder bruikbaar om het gezondheidseconomische potentieel van dergelijke nieuwe diagnostiek te evalueren.

CONCLUSIES

In dit proefschrift is aangetoond dat traditionele diagnostiek minder frequent dan verwacht wordt uitgevoerd op zowel tumor biopsies als de volledige tumor-resectie. Wanneer deze testen op beide materialen worden uitgevoerd zijn de resultaten van

de testen vaak concordant. Gezien de invasieve aard van de traditionele methoden, zijn deze niet passend om frequent tumor-materiaal te verkrijgen. Nieuwe methoden worden daarom ontwikkeld om mogelijke tumor-heterogeniteit in kaart te brengen. Hoewel liquid biopsies veelbelovend lijken om dit te doen, focust de meerderheid van studies naar liquid biopsies onvoldoende op de klinische utiliteit van deze liquid biopsies. Echter, het adresseren hiervan is nodig alvorens dergelijke methoden in de klinische praktijk kunnen worden geïmplementeerd. De klinische utiliteit van liquid biopsies is geëvalueerd in een enkele trial, de STIC METABREAST trial. Hoewel de klinische resultaten lieten zien dat het sturen van de behandeling van patiënten met gemetastaseerde borstkanker met endocriene- of chemotherapie niet inferieur is aan de huidige praktijk, zijn de gezondheidseconomische resultaten nog erg onzeker.

ACKNOWLEDGEMENTS

DANKWOORD

Gedurende mijn promotietraject heb ik enorm veel mensen leren kennen van wie ik onwijs veel heb mogen leren. Uiteraard hebben zij voor een groot deel bijgedragen aan mijn ontwikkeling betreffende de inhoud van dit proefschrift. Daarnaast heb ik ook heel erg veel over mijzelf als persoon, als collega, als vriendin, als zus, dochter of als ander familielid geleerd. Ook al is een deel van mijn familie- en vriendenkring helemaal niet thuis in wetenschappelijk onderzoek of de inhoud en structuur van een promotie-traject, toch heeft iedereen altijd de tijd en moeite genomen om naar me te luisteren als ik in al mijn enthousiasme weer een (waarschijnlijk vaak veel te ingewikkeld) technisch verhaal begon te vertellen.

Iedereen om mij heen, zowel binnen als buiten mijn werkomgeving, heeft er een steentje aan bijgedragen dat ik mijn promotietraject met zo enorm veel plezier heb kunnen doorlopen de afgelopen jaren. Voordat ik verder ga schrijven aan dit dankwoord wil ik eerst iedereen ontzettend bedanken voor jullie steun de afgelopen jaren!

Oh ja, even een kleine aanwijzing voor de leken onder ons... Het dankwoord is datgene wat iedereen als eerste (en misschien wel als enige) leest.

Maarten, tijdens het eerste college dat ik in de Master Health Sciences in 2015 van jou kreeg, was ik er eindelijk écht van overtuigd dat ik de juiste richting had gekozen. Omdat je zo ontzettend veel weet van allerlei aspecten binnen jouw vakgebied, maar daar ook nog eens op een enorm enthousiaste en inspirerende manier over kan vertellen, ben je voor mij eigenlijk altijd een groot voorbeeld geweest vanaf dat moment. Je hebt me vele kansen en mogelijkheden geboden om mijzelf niet alleen op het gebied van onderzoek doen, maar ook op het gebied van andere vaardigheden te kunnen ontwikkelen. Het heeft mij enorm veel motivatie gegeven dat jij er eigenlijk altijd wel op vertrouwd dat het goed kwam. Op verschillende (internationale) meetings mocht ik dan ook zelf onze onderzoeken presenteren en dat voelde telkens weer als een grote eer. Jouw netwerk is enorm en ik ben je er enorm dankbaar voor dat je me de kans hebt gegeven om met allerlei mensen uit dat netwerk in contact te komen.

Erik, ik heb jou leren kennen tijdens mijn afstuderen voor de Master Health Sciences. Je hebt toen mijn afstudeeropdracht begeleid die bestond uit het maken van een vroeg gezondheidseconomisch model. Hoewel ik er toen al van overtuigd was dat jij alles wist van dergelijke modellen en het maken daarvan, is dat de afgelopen jaren eigenlijk alleen maar meer en meer bevestigd. Ik kan me nog steeds eigenlijk niet voorstellen dat er iemand bestaat die me daar beter bij zou hebben kunnen begeleiden dan dat jij dat deed. Buiten dat je daarbij enorm goed kan helpen, heeft het mij enorm geholpen dat jij altijd vanuit die modellen een vertaalslag naar de klinische praktijk kon maken. Jouw kennis

en kunde hebben mij wat dat betreft enorm gemotiveerd! Eigenlijk had dit natuurlijk een volledig “Erik-bedankt-boek” moeten zijn, maar dat kon in de budgetten nét niet helemaal uit. Bedankt dat ik ‘jouw patiënt’ heb mogen zijn de afgelopen jaren en dat ik je de dansvloer op heb mogen slepen bij ISPOR!

Oh ja, en ook bedankt voor alle verrassingsstellages of stickerpatronen die je op mijn bureau hebt gebouwd als ik een dag niet op kantoor was geweest, haha!

Maarten en Erik, ik had me echt geen betere begeleiders kunnen wensen de afgelopen jaren. Ik heb enorm met jullie gelachen maar ook echt ontzettend veel van jullie geleerd. Daar waar ik na mijn master het gevoel had dat ik eigenlijk nog steeds van alles net niet genoeg wist, durf ik nu zelfverzekerd de universiteit te verlaten. Dit allemaal heb ik echt voornamelijk aan jullie en aan alle kansen die ik van jullie heb gekregen te danken!

Dear members of the doctoral commission, thank you very much for reading and approving this thesis and for taking the time to be present at the defense ceremony.

Dear co-authors, thank you very much for all of your time and the support that you provided the last couple of years. You all have been a huge inspiration to me to even further improve on our research. Thank you for your trust and your cooperation.

Leon, Stefan en Carolien, gedurende mijn promotietraject heb ik veel contact met jullie gehad over allerlei vraagstukken en artikelen. Enorm bedankt voor alle ondersteuning die jullie de afgelopen jaren hebben geboden. Ook al zijn jullie niet officieel betrokken geweest bij mijn promotie-traject, jullie kennis, enthousiasme en toewijding zijn echt een enorme stimulans voor mij geweest.

Koen, wat hebben we enorm veel gelachen en plezier gemaakt in Raffa 5155. We moeten trouwens zeker nog eens onderzoek doen naar de invloed van borrels op de kwaliteit van onze prestaties in de ochtend, naar het effect van een Spaanse commentator op het doelsaldo bij tafelvoetbal, naar het aantal surflessen dat je als beginner nodig hebt als je al op een rolkarretje op kantoor hebt geoefend, het aantal kleiduiven dat je kan raken als je een expert bent met een elastiekjespistool of de gemiddelde gewichtstoename als je een paar jaar bij ons op kantoor hebt gezeten. Zoals jij ook al eens hebt benadrukt, zijn we samen inderdaad echt de hele wereld over gereisd, en hebben we enorm veel met elkaar meegemaakt. Bedankt voor alle motivatie die je me ook hebt gegeven, maar misschien nog wel meer bedankt voor alle positiviteit die je altijd weer met een mega glimlach wist mee te brengen en voor alle andere random uitspraken waar ik zo hard om heb moeten lachen!

Annelies, wat ben je toch een heerlijk mens! Enorm bedankt dat je een van mijn paranymfen wil zijn, maar bovenal bedankt voor al die ontzettend gezellige momentjes de afgelopen jaren. Ook al heb ik je de afgelopen tijd enorm moeten missen bij de vakgroep, jouw steun en toeverlaat is er zeker niet minder door geworden en dat waardeer ik enorm. Bedankt voor al je geduld en je luisterend oor als ik weer eens in al mijn enthousiasme of frustratie vreselijk aan het ratelen was, en natuurlijk voor het sponsoren bij Starbucks als ik weer eens mijn portemonnee was vergeten!

Deon, thank you so much for being one of my paranymphs, but even more for all the fun moments that we had in the office, during conferences, birthdays or just at a random day when we decided that it was time for some fun and drinks. Also many thanks for your support in the organization of the summer school the last years. Of course, also special thanks for being my (or actually HTSR's) dancing queen at the conferences!

Melissa, bedankt voor alle gezellige momenten die ik met jou heb mogen delen. Als we weer eens uit eten gingen, waren wij uiteraard weer degenen die als laatsten aan het genieten waren van de hapjes, drankjes en de dansjes! En onwijs bedankt voor het lenen van jouw man, **Kevin**, voor het opmaken van dit boekje. Het ziet er fantastisch uit!

Karin, allereerst enorm bedankt voor het vertrouwen dat je in me hebt gehad. Ik heb je mogen ondersteunen bij het vak Medische Statistiek terwijl mijn eigen eerste tentamen statistiek – weliswaar zonder te leren – een drama was om na te kijken en dat wist jij nog al te goed! We hebben samen een aantal keer een R-meeting bezocht en cursussen in R mogen geven. Enorm bedankt voor jouw loyaliteit, gezelligheid, enthousiasme, interesse en alle kansen die je me hebt geboden.

Michelle, bedankt voor al je gezelligheid de afgelopen jaren. We hebben (zeker afgelopen zomer) enkele dagen met zijn tweeën de vakgroep draaiende gehouden en regelmatig enorm gelachen. Dat beperkte zich uiteraard niet tot de momenten dat de andere collega's er niet waren, ook aan de lunchtafel hebben regelmatig tranen over onze wangen gestroomd. Daarnaast hebben we samen ook nog eens aan een onderzoek mogen werken en zijn we op meerdere congressen geweest en hebben we vooral in Baltimore samen heel gezond ontbeten iedere dag. Love die Subway. Bedankt voor alles!

Evita, het was weliswaar kort maar krachtig, maar toch bedankt voor alle gezelligheid die je hebt meegebracht in die korte tijd! Oh ja, en ook bedankt voor die mooie paarden die je hebt getekend op het bord in Raffa 5155.

Alle andere (oud-) collega's van HTSR, bedankt voor alles de afgelopen jaren. Bedankt voor alle gezelligheid, maar bovenal voor alle waardevolle gesprekken en discussies!

Everybody from the Medical Cell BioPhysics (MCBP) department, thank you for your patience in explaining several biological principles in the beginning of my PhD trajectory. You all inspired me with your dedication to the work you all do. Several of you have also helped me out during the summer school, special thanks for that!

Esther, bedankt voor het enthousiasme waarmee je ons telkens weer hebt geholpen met de data die uit jullie PALGA database kwam. Jouw toewijding en interesse in de onderzoeken die gedaan worden met PALGA data zijn echt bewonderenswaardig en hebben mij enorm gemotiveerd. Bedankt daarvoor.

Everyone involved in the ISPOR Student Network, thank you for the amazing times that we had during the ISPOR conferences. Special thanks to Laura, Jason, Blythe and William, you are all such wonderful people and you definitely made “My ISPORs” into the best events I have attended the last years. Thank you all for providing me the opportunity to work with you and thanks for all fun times we had as well.

CuriousU colleagues, thank you very much for all the fantastic days that we had together during the summer schools the last couple of years! I really enjoyed being part of this amazing summer school. Though I haven’t been participating that much the last year because I had to finish writing this thesis, you definitely made my summers at the university worth it! Really appreciate it, thanks guys!

Rianne, dank voor je betrokkenheid, je interesse, je vertrouwen, je steun en je gezelligheid. CuriousU was zonder jou voor mij niet hetzelfde geweest!

Mijn lieve vriendjes & vriendinnetjes, Carolien, Jamila, Judith, Nienke, Anne, Lizzy en Melanie, enorm bedankt voor alles de afgelopen jaren! Bedankt voor jullie steun, voor jullie vertrouwen, voor jullie geduld, jullie luisterend oor, jullie gekkigheid en nog veel meer van dat. Altijd als ik even iets leuks wilde doen, even toe was aan een beetje ontspanning of gewoon een goed gesprek wilde voeren stonden jullie voor me klaar. Jullie hebben er altijd voor gezorgd dat ik weer fris en fruitig (oké, met uitzondering van de zondagen dat we wellicht een paar borreltjes beter hadden kunnen overslaan), vol goede moed en gemotiveerd aan een nieuwe week kon beginnen. Zonder jullie was dit me allemaal echt niet gelukt.

Maarten, iedereen weet inmiddels wel dat jij eigenlijk meer een soort van extra broertje dan een vriend voor me bent. Hoewel we niet meer bij elkaar om de hoek wonen stond je wel altijd direct voor me klaar. Enorm bedankt daarvoor, en natuurlijk voor alle gezelligheid!

Mijn stalgenootjes en andere paardenvriendjes, ontzettend bedankt voor alle gezellige momenten de afgelopen jaren. Mijn paarden hebben er samen met jullie voor gezorgd dat ik iedere dag en in de weekenden weer even een aantal uurtjes lekker kon ontspannen.

Papa, Mama, bedankt voor al jullie steun de afgelopen jaren, niet alleen tijdens mijn PhD maar ook tijdens mijn studie daarvoor hebben jullie alle vertrouwen in mij gehad! Dat heeft mij enorm gemotiveerd en zonder jullie had ik dit allemaal écht niet kunnen bereiken.

Mike en Roxan, bedankt dat jullie ook elke keer maar weer mijn welbekende werkverhaal hebben aangehoord tijdens gelegenheden met familie en vrienden. Ook bedankt voor al het geduld dat jullie met me hebben gehad de afgelopen jaren, alle vertrouwen dat ook jullie in me hebben gehad, maar nog meer voor alle plezier dat ik met jullie heb mogen beleven! Jullie zijn fantastisch!

Papa, bedankt dat je altijd zo trots op me bent geweest! Ik vond het altijd fantastisch om te horen dat jij weer aan vrienden of kennissen hebt proberen uit te leggen waar ik mee bezig ben. Bedankt voor je onverminderde vertrouwen de afgelopen jaren! Babette, ook jij bedankt voor je oneindige interesse en enorme lading positiviteit de afgelopen jaren! Daarnaast wil ik je ook graag bedanken voor het vroegtijdig lezen en corrigeren van de Nederlandse teksten in dit proefschrift.

Mama, je hebt altijd alle vertrouwen in me gehad en dat heeft me enorm gemotiveerd! Je bent altijd geïnteresseerd geweest in het werk dat ik doe en dat waardeer ik enorm! Ook onwijs bedankt dat je me altijd in alles onvoorwaardelijk hebt gesteund! En dat ik altijd bij jullie kon aanschuiven natuurlijk, haha. Huub, ook jij enorm bedankt voor je onverminderde interesse en goede hulp de afgelopen jaren!

Joep, Noa en Aviva, ook jullie als bonusbroertje en bonuszusjes bedankt voor alle gezelligheid de afgelopen jaren! Jullie hebben vaak, net als Mike en Rox, het hele verhaal moeten aanhoren over mijn eeuwig durende ingewikkelde onderzoeken. Maar ook met jullie heb ik samen vele leuke dingen mogen beleven en daar ben ik jullie enorm dankbaar voor!

Mijn hele familie, van alle kanten, familie kun je niet kiezen, maar ik ben wel heel erg dankbaar dat jullie mijn familie zijn!

Oma Marianne, of eigenlijk Oma Holten zoals ik dat vroeger altijd zei, bedankt dat jij er ook altijd was. Na het werk kon ik bij jou altijd even binnen-vallen voor een kopje koffie, een broodje, even kletsen, en allerlei andere gezellige dingen. Bedankt voor je onverminderde interesse en onvoorwaardelijke steun.

Opa en Alexandra, bedankt voor jullie oneindige interesse en support de afgelopen jaren.

Alice, Lynn, Remco, jullie ook enorm bedankt voor jullie oneindige interesse en steun de afgelopen jaren! En uiteraard nog veel meer voor alle gezellige momenten die we met elkaar hebben mogen meemaken tijdens familiedagen, verjaardagen, kerstdagen en de rest natuurlijk.

Oma Greet en Opa Jan Berghuis (in memoriam), helaas hebben we recentelijk ook afscheid moeten nemen van onze lieve oma Greet, en van opa al een behoorlijk aantal jaren geleden. Oma, je grapte altijd nog dat het tijd werd dat ik eens een echte baan ging zoeken, wat zou je trots zijn om te zien dat dat nu echt gelukt is. Bedankt voor je oneindige interesse!

Lianne, Sjors, Bas, Dorien, Bruno, Milan, ook jullie wil ik enorm bedanken voor jullie oneindige interesse en steun de afgelopen jaren! In het bijzonder wil ik jullie nog even bedanken voor jullie bijdrage aan het 30-seconds spel met Kerst, wat hebben we toen ook ontzettend gelachen!

Ook enorme dank aan alle anderen die ik in dit dankwoord misschien wel vergeten ben, maar die ik de afgelopen jaren wel regelmatig om mij heen heb gehad of heb mogen leren kennen. Ik heb me enorm gesteund gevoeld en altijd het idee gehad dat iedereen er volledig op vertrouwde dat dit allemaal wel goed kwam, en dat deed het dan ook!

Dank jullie wel!

CURRICULUM VITAE

CURRICULUM VITAE

Anne Margreet Sofie Berghuis was born on the 4th of February 1992 in Deventer, the Netherlands. After completing her secondary education, at the Etty Hillesum Lyceum (Deventer) high school in 2010, she started as a bachelor student in the Health Sciences program at the University of Twente in Enschede, the Netherlands. She finished this program in 2014 and started her master Health Sciences at the University of Twente directly after. In the master, she specialized in the track: Health Technology and Innovation. In

parallel, she took a course on quantitative methods for operations management. In 2015 she wrote her master thesis on: "Using Circulating Tumor Cells in the early staging phase of breast cancer" and obtained her master degree Cum Laude after defending this thesis.

From October till December 2015 she worked as a Junior Researcher at the department of Health Technology and Services Research at the University of Twente. The main aim of the project she was working on during this period, was to further extend the work that she was working on in her master thesis. During this period, the research from the master thesis was presented at the ISPOR conference in Milan.

In March 2016 she started as a PhD candidate at the department of Health Technology and Services Research at the University of Twente. Her PhD project mainly focused on the clinical utility of molecular biomarkers in breast cancer management. The results of this PhD project are described in this dissertation. For one of the projects in which she used the Dutch Pathology Registry (PALGA), she was awarded the PALGA prize, for best research using this database.

Besides the PhD project, Sofie was involved in several side activities within the International Society for Health Economics and Outcomes Research (ISPOR). From 2017-2018 she was involved in the Student Network Chapter Development Committee. From 2018-2019 she was the president for the ISPOR University of Twente Student Chapter and co-chairing the ISPOR Student Network Education Committee. For her contributions in this term (2018-2019) she received an ISPOR Distinguished Service Award.

Sofie will continue her career at GlaxoSmithKline in Zeist, the Netherlands. She will start working as a Market Access Manager.

LIST OF PUBLICATIONS

THESIS PUBLICATIONS

PUBLISHED

Detecting blood-based biomarkers in metastatic breast cancer: A systematic review of their current status and clinical utility. Berghuis AMS, Koffijberg H, Prakash J, Terstappen LWMM & IJzerman MJ (2017). International journal of molecular sciences, 18(2), [363]. DOI: 10.3390/ijms18020363

Evidence on the cost of breast cancer drugs is required for rational decision making. Berghuis AMS, Koffijberg H, Terstappen LWMM, Sleijfer S & IJzerman MJ (2018). Ecancermedicalscience, 12, [825]. DOI: 10.3332/ecancer.2018.825

Health economic impact of liquid biopsies in cancer management. IJzerman MJ, Berghuis AMS, de Bono JS & Terstappen LWMM (2018). Expert review of pharmacoeconomics & outcomes research, 1-7. DOI: 10.1080/14737167.2018.1505505

Real-world data on discordance between estrogen, progesterone and HER2 receptor expression on diagnostic tumor biopsy versus tumor resection material. Berghuis AMS, van Deurzen CHM, Koffijberg H, Terstappen LWMM, Sleijfer S, IJzerman MJ (2018). Breast Cancer Research and Treatment. 18, 2019.

SUBMITTED

Treatment decisions of adjuvant systemic therapy following repeated assessment of estrogen and the human epidermal growth factor receptor 2 on tumor biopsy and resection material in invasive primary breast cancer patients. Berghuis AMS, Koffijberg H, van Deurzen CHM, Terstappen LWMM, Sleijfer S, IJzerman MJ

PENDING SUBMISSION

An online tool for the Rapid assessment of the health economic impact of using CTCs to guide systemic therapy in metastatic breast cancer. Berghuis AMS, IJzerman MJ, Berger F, Pierga JY, Jacot W, Alix-Panabières C, Bidard FC, Koffijberg H

Health economic impact of using CTCs to guide systemic therapy in metastatic ER+ HER2- breast cancer patients – Results from a randomized controlled multicenter trial. Berghuis AMS, Koffijberg H, Berger F, Alix-Panabières C, Jacot W, Pierga JY, Bidard FC, IJzerman MJ

NON-THESIS PUBLICATIONS

PUBLISHED

Oncogenomics. From basic research to precision medicine. Chapter 28: Role of Circulating Tumor Cells in Breast Cancer Patients. Lampignano R, Berghuis AMS, Fehm T & Neubauer H, Elsevier (2018).

Cost-effectiveness of procalcitonin testing to guide antibiotic treatment duration in critically ill patients: results from a randomised controlled multicentre trial in the Netherlands. Kip MMA, van Oers JA, Shajiei A, Beishuizen A, Berghuis AMS, Girbes AR, de Jong E, de Lange DW, Nijsten MWN, IJzerman MJ, Koffijberg H & Kusters R. 13 Nov 2018 Critical care (London, England). 22, 1, 10 p., 293

EpCAMhigh and EpCAMlow circulating tumour cells in metastatic prostate and breast cancer patients. Mutter-de Wit S, Manicone M, Rossi E, Lampignano R, Berghuis AMS, Andree KC, IJzerman MJ & Terstappen LWMM. 2 Nov 2018 Oncotarget. 9, 86, p. 35705-35716 12 p.

How to Agree on a CTC: Evaluating the Consensus in Circulating Tumor Cell Scoring. Zeune LL, de Wit S, Berghuis AMS, IJzerman MJ, Terstappen LWMM & Brune C. Cytometry A. 2018 Sep 24. doi: 10.1002/cyto.a.23576

IN PROGRESS

Radioactive seed versus wire-guided localization for ductal carcinoma in situ of the breast: comparable resection margins. Agahozo MC, Berghuis AMS., van den Broek E, Koppert LB, Obdeijn I & van Deurzen CHM

CONFERENCE CONTRIBUTIONS

ABSTRACTS & PRESENTATIONS

Identifying the optimal use of CTCs in the early staging phase of breast cancer. Berghuis AMS, Koffijberg H & IJzerman MJ (2015). A348. DOI: 10.1016/j.jval.2015.09.624

What do you believe? An online tool for cost-effectiveness analysis of circulating tumor cell detection based on user beliefs. Berghuis AMS, Koffijberg H & IJzerman MJ (2016). Poster session presented at 38th Annual North American Meeting of the Society for Medical Decision Making, SMDM 2016, Vancouver, Canada.

The cost of expensive breast cancer drugs. Poster session presented at ESMO 2017 Congress, Madrid, Spain. Berghuis AMS, Koffijberg H, Terstappen LWMM, Sleijfer S & IJzerman MJ (2017).

Estimating The Drug Treatment Cost of Breast Cancer. Value in health, 20(9), A430-A431. DOI: 10.1016/j.jval.2017.08.189 Berghuis AMS, Koffijberg H, Terstappen LWMM, Sleijfer S & IJzerman MJ (2017).

The Impact Of Cluster Selection Methods In Two-Stage Bootstrapping To Assess Uncertainty. Kip MMA, Berghuis AMS, Nijsten MW, IJzerman MJ, & Koffijberg H. (2018). In Health Economic Outcomes In Cluster Randomized Controlled Trials. Value in health, 21(Suppl. 1), PRM111. DOI: 10.1016/j.jval.2018.04.1544

Rapid assessment of the cost-effectiveness of liquid biopsies – an online tool to explore and update the evidence. Berghuis AMS, IJzerman MJ & Koffijberg H. Presented at the lowlands health economic study group (IolaHESG) conference 2019.

OTHER CONTRIBUTIONS

Course leader CuriousU Summer School: Evaluating Future Health Technologies (2017, 2018, 2019).

Vice President ISPOR University of Twente Student Chapter (2017-2018).

Member ISPOR Student Network Chapter Development Committee (2017-2018)

President ISPOR University of Twente Student Chapter (2018-2019).

Co-Chair ISPOR Student Network Education Committee (2018-2019)