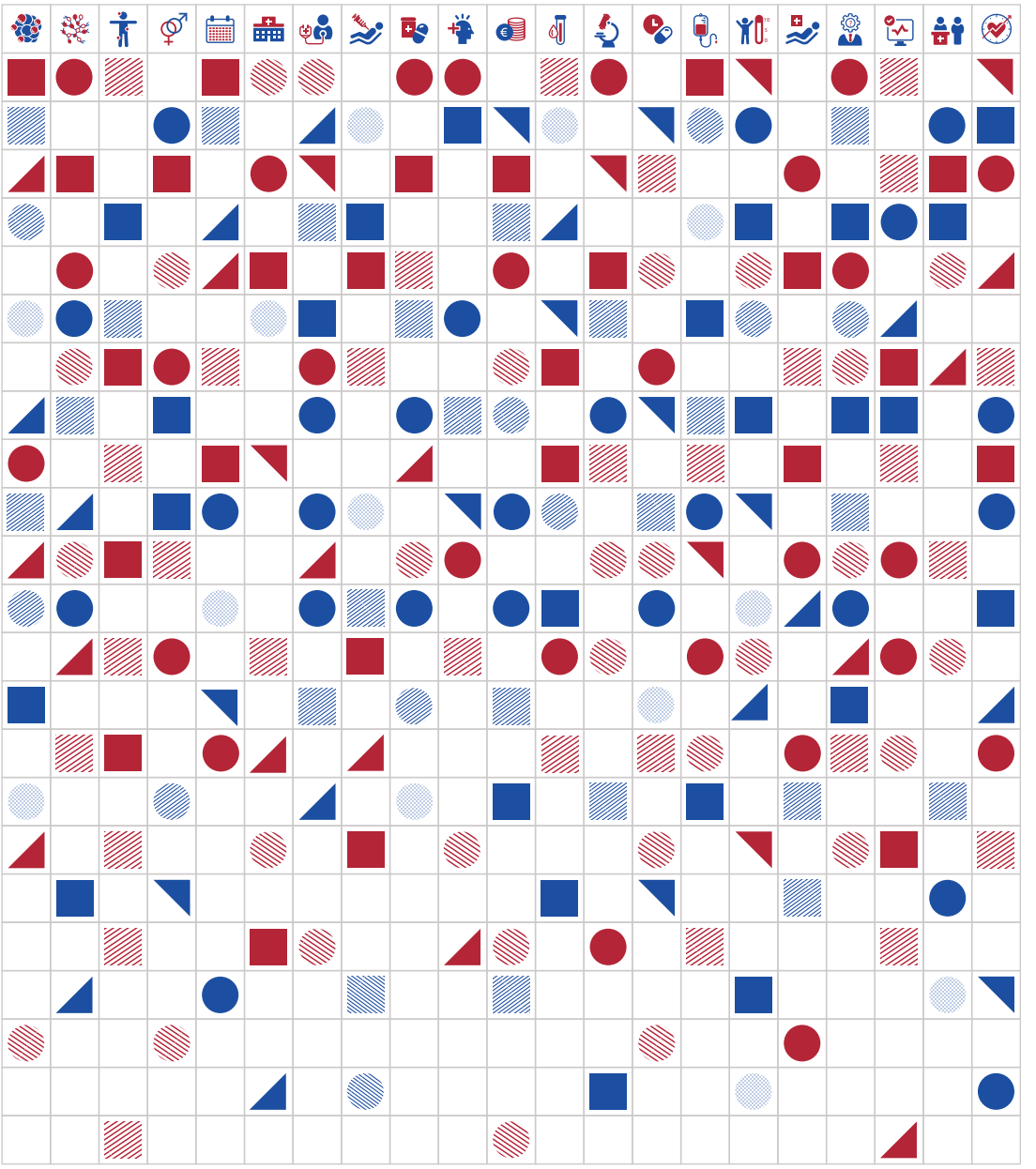

MAKING CANCER DATA MEANINGFUL

Using real-world data to evaluate cancer care and outcomes, and support evidence-based decision-making

MARIANNE LUYENDIJK

MAKING CANCER DATA MEANINGFUL

Using real-world data to evaluate cancer care and outcomes, and support evidence-based decision-making

MARIANNE LUYENDIJK

Author
Marianne Luyendijk
Design
Iris Jönsthövel, www.irisjonsthovel.com
Published and Printed by
Optima Grafische Communicatie, www.ogc.nl
ISBN: 978-94-6510-962-6
The research for this thesis was made possible by the Netherlands Comprehensive Cancer Organisation (Integraal Kankercentrum Nederland, IKNL), in close collaboration with the Erasmus School of Health Policy and Management (ESHPM) at Erasmus University Rotterdam, and with additional academic input from Erasmus Medical Centre (Rotterdam).

Making Cancer Data Meaningful

Using real-world data to evaluate cancer care and outcomes, and support evidence-based decision-making

Het genereren van zinvolle data

Het gebruik van real-world data om kankerzorg en -uitkomsten te evalueren en besluitvorming te ondersteunen

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr.ir. A.J. Schuit

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 21 november 2025 om 10.30 uur

door

Marianne Luyendijk
geboren te Al-Khobar, Saoedi-Arabië.

PROMOTIECOMMISSIE

Promotoren:	prof.dr. C.A. Uyl-de Groot prof.dr. S. Siesling prof.dr. A. Jager
Overige leden:	prof.dr. V.P. Retèl prof.dr. S.C. Linn prof.dr. M.J. IJzerman
Copromotoren:	dr. H.M. Blommestein

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

TOWARDS A 'CURE' FOR CANCER

In 1969 *The Washington Post* published a whole-page advertisement, titled*, "Mr. Nixon: You can Cure Cancer" (1)*. At that time, the belief ran, that curing cancer was just a matter of willpower, proper planning and sufficient funding (2). Unfortunately, half a century later, with 18 million new diagnoses and 10 million deaths worldwide each year, a cure for all cancer types still seems far beyond our grasp (3).

The belief that a cure for cancer was a realistic ambition descended from the breakthrough discoveries in the 1960s and 1970s. Aggressive chemotherapy combinations were then shown to successfully achieve durable response and even cure in lymphoblastic leukemia, advanced Hodgkin's lymphoma and metastatic testicular carcinoma — cancers that had previously responded poorly to treatment (2, 4, 5, 6, 7, 8). At that time, scientists assumed that there was *one* mechanism and *one* cure for cancer. If one cancer could be cured with combination chemotherapy, then other cancers should follow (9). Consequently, the number of studies investigating chemotherapy proliferated during the years to come. Various combinations with the available therapies were tested, doses were escalated, and treatment durations were extended—all with the objective to lengthen patient's lives. Regrettably, by 1985, only minor improvements in survival—typically of a couple of weeks to months—were observed at the population level (2, 10).

Surgeons also took radical approaches until the 1980s. They assumed that the more tissue they cut away, the higher the probability of cure. In women with breast cancer, for instance, not only was the breast removed, but also the pectoral muscles, axillary nodes, the wall of the chest and sometimes even parts of bones (e.g., ribs, sternum, clavicle) (2, 11). Surgeons started to question these radical approaches when a better understanding of the origin of tumors and their patterns of spread emerged (12, 13). Indeed, in the early 1980s, the results of a trial investigating radical mastectomy versus an approach in which only malignant lumps were removed — followed by radiation to the breast, showed no difference in cancer recurrence and survival (14, 15). This approach changed the standard of care and the basic principles of this type of surgery (lumpectomy) are still applied today. But it also contributed to a theory that had been proposed centuries ago: solid cancers could be either localized and curable by surgery (with or without systemic therapy) or metastatic and incurable by surgery (2, 16, 17).

Following the chemotherapy studies during the 1980s, it became clear that a better understanding of the biologics of cancer was needed in order to improve treatment (10). One important step forward in this respect was the discovery of the role of hormones

in the genesis and growth of breast and prostate cancer cells (2, 18, 19). This finding not only led to the introduction of effective anti-hormonal therapies, but it also contributed to the realization that cancer is a highly heterogeneous disease. Later, the hypothesis emerged that certain targets (e.g., proteins) in cells might play a role in cancer cell growth and that a possible treatment approach could be to inactivate these targets. Indeed, this resulted in the development of effective 'targeted' medicines (2, 10). One of the first such agents, the monoclonal antibody trastuzumab, reduced the size and number of metastases (visible on a CT scan) in metastatic HER2-receptor overexpressing breast cancer. In 1999, trastuzumab was approved for use in clinical practice after clinical studies showed improvements in survival of about 5 months (from 20.3 months to 25.1 months) in patients with metastatic breast cancer (20).

Many other targeted therapies for various cancer types followed and a new era of personalized oncology began where providing the right treatment to the right patient, reducing toxicity and improving quality of life became important considerations (21). While there have been profound successes with targeted medicines in the treatment of metastatic cancers (e.g., imatinib for gastrointestinal stromal tumors and chronic myeloid leukemia, trastuzumab for breast cancer, vemurafenib for melanoma), the response to these agents is often not durable (22). As with chemotherapy, cancer cells will become resistant to these agents sooner or later (2, 23). Newer strategies aimed at increasing response duration included the use of multiple targeted therapies and the application of multiple lines of treatment. Interestingly, the discovery of cancer targets has also enabled treatment to de-escalate. That is, if a diagnostic test shows that there is no overexpression of HER-2 receptors in a patient's cancer, it is of no use to give this patient a HER-2–targeting agent. A trend towards "doing less" continues in both local and systemic treatment. For instance, with the introduction of gene-expression profiling tests (e.g., mammaprint/oncotype DX), it is possible to predict which patients with early-stage breast cancer have a very good prognosis and can safely forgo chemotherapy (24, 25).

Paradigm shifts in cancer medicine development were rewritten in the 2010s with approval of immunotherapies and cell- and gene- based therapies for use in clinical practice. Both therapies provide novelty, by using new modes of action: immunotherapies re-activate the immune response, allowing the body to attack cancer cells, while gene therapies inactivate genes that are causing problems (26, 27). Additionally, both show remarkable benefits in specific subgroups of patients. For instance, immune checkpoint inhibitors have revolutionized the treatment of patients with metastatic melanoma, improving the 5-year survival rate from historical levels of less than 10% to over 50% in clinical studies (28, 29). Gene therapies offer potential cures for childhood leukemia,

with patients previously considered incurable now living for over 10 years (30). While these therapies are obviously valuable additions to the treatment arsenal, they are unfortunately not effective in all patients.

In the history of curing cancer, prevention and early detection also gained considerable attention. Important discoveries in this regard include the identification of cancer causing substances (predominantly from smoking) and the development of screening tests that are capable of detecting (pre-)malignant growth that can be cured with surgery (e.g., the Pap smear test to detect abnormal cervical cells before they have a chance to become cancerous) (31, 32). Unfortunately, prevention and early detection also have not delivered a universal 'cure' for cancer (33, 34). As history shows, 'curing' cancer is a cumulative process that results from decades of research (2). A process that will likely continue in the coming decades.

WE MUST HAVE DATA

In the early days of cancer medicine discovery, scientists already recognized that objective data were needed to demonstrate the effects of their treatments in humans. Physicians, so they recognized, could not simply perform experiments without biases arising regarding the types of patients they treated and their characteristics (2, 35). A famous statistician, named Bradford Hill, proposed a study design to overcome these problems: the randomized controlled trial (RCT). By randomly assigning a large group of patients to an intervention and a control group, both observable and not-observable patient characteristics could be balanced across groups. Besides randomization, Hill's methodology involved a clear definition of patient eligibility, treatment exposure, and outcomes measured (i.e., method and time points). All these features made the RCT the most rigorous method for determining a cause-and-effect relationship (36).

Almost a century later, the RCT is still the most important source of evidence in medicine (37). Results from RCTs are used by various stakeholders, including scientists, pharmaceutical companies, regulators, policy makers and physicians (38, 39, 40, 41). Companies and other organizations conduct RCTs to demonstrate that new technologies (including medicines) are safe and effective. Federal agencies involved in regulating the market authorization (MA) of medicines utilize these results to draw conclusions about the safety, efficacy, and quality of new medicines. Policymakers use them as input for their assessments of (cost)effectiveness and to make decisions regarding reimbursement. Physicians use results to make evidence-based treatment decisions.

Despite their prominent role, RCTs do not provide all the answers relevant to the three different stakeholders in oncology (42). When RCTs are performed, which is usually the case for medicines, evidence gaps remain. Firstly, because the generalizability of results from RCTs to daily practice is often hampered by the (strict) eligibility criteria and highly controlled environment. In daily practice, patients are usually less fit, and their treatment adjustments are more flexible (43). Secondly, in the era of personalized oncology, where medicines target relatively small groups of patients, RCT design is becoming increasingly complex (22). Finally, driven by societal pressure and the competitive pharmaceutical industry, there are incentives to bring products to the market quickly. This appears to play a role in the way studies are designed (e.g., patient inclusion, choice of comparator arm, and outcomes), performed and results are communicated (35, 44, 45, 46). Besides the uncertainties surrounding the effects of medicines in oncology, there is often even more uncertainty about the actual effects of other health technologies (e.g., medical devices). Evidence of the performance of medical devices is usually not based on RCTs because the licensing procedures are less regulated than that of medicines (47). Moreover, the characteristics of technologies (e.g., the performance of a device may vary depending on the user) make it more challenging to conduct RCTs and interpret their results (48)

The different stakeholders in oncology need data to fill evidence gaps. As such, there is an increased interest in using data from everyday clinical practice, well known as 'real -world data' (RWD), to fill research gaps (49, 50, 51)]. Like data from RCTs, too, the use of RWD to evaluate outcomes in oncology is not a new phenomenon. By 1986, Bailar & Smith published one of the first scientific papers based on population-based cancer registry data. This study showed that mortality increased by almost 9% from 1962 to 1985, even though researchers had hypothesized a decrease in in the number of cancer-related deaths as a consequence of the introduction of curative treatments (e.g., cisplatin for testis carcinoma) (2, 52). The rising mortality rates were a consequence of an enormous increase in the number of smokers and a resulting increase in the incidence of smoking-related cancers (mainly lung cancer) (53).

The paper by Bailar & Smith (1986) is one of the first examples of how RWD can generate valuable insights to improve cancer care and outcomes. Ever since this publication, the amount of RWD has proliferated. Patient data, potentially relevant to answer research questions, is everywhere. For example, healthcare providers report information regarding diagnoses and treatments in electronic health records and they submit insurance claims to receive payment for services delivered. Patients generate data by filling out quality of life questionnaires, recording side effects in apps and by using wearables. Organizations aiming to provide insights into the burden of cancer (e.g., incidence, sur-

vival and prevalence) and the quality of cancer care, collect data from several sources, preferably including cancer registries (54, 55). These data can generate insights that cannot be obtained from RCTs, for instance, to identify disparities in the use of new medicines in practice (e.g., difference in uptake between age groups), to understand the long-term safety and other outcomes related to medicines and to evaluate the costs of care (54, 56).

While potentially relevant, RWD studies also come with challenges. Potential flaws in using RWD for comparative effectiveness research include the risk of selection bias and confounding, which both lead to incomparability of treatment groups under study (57). Selection bias occurs when treatment and control groups are selected in such a way that they are systematically different from each other. This happens, for instance, when fitter patients receive other types of treatment than patients who are less fit or when patients drop out of a study due to side effects related to a specific treatment. Confounding occurs when an apparent relationship between a treatment (or another exposure) and an outcome is distorted by a third (often unmeasurable) factor. This for instance happens when the choice of treatment for patients is related to their prognosis (e.g., due to older age, comorbidity or lifestyle such as smoking). Other often voiced concerns for any type of RWD study include data completeness, quality and validity (55, 58).

HOW MUCH CAN WE SPEND?

Partly due to advances in treatment, the prognosis of cancer has improved over the decades. While this is good news, there is also a backside. That is, the costs of cancer care in Europe doubled from about 50 billion in 1995 to over 100 billion in 2018, with 32 billion spent on cancer medicines and 71 billion on other aspects of healthcare (59, 60). In most European countries, healthcare is collectively financed, either through tax or insurance-based systems. The increasing costs of care overall, and cancer treatment specifically, put pressure on healthcare budgets, raising concerns about the financial sustainability of health care systems and the accessibility to care and new treatments. Policymakers, who need to balance spending on cancer control alongside other policy priorities, increasingly have to make choices on what care to reimburse and what not (61).

The rising costs of care are driven by the ageing population and the growth in the number of patients in need of care. Yet, they are also driven by the introduction of new health technologies, including medicines and medical devices, as well as the need for a multidisciplinary team to provide care. The number of treatment options is increasing,

and these options are often used as adjuncts rather than as replacements for existing treatments (61). For instance, prior to 1990, about 40 systemic therapies were available to treat metastatic solid cancers, while over 175 options were available in 2020 (61, 62, 63). Additionally, the costs of treatments are increasing as well. Nowadays, €100.000 per patient for a treatment course is no exception in oncology (64). Medicines are higher-priced, partly because they are more expensive to develop and produce and they target smaller populations, making it more difficult to earn back investments. But also, because expensive diagnostic tests are needed to identify patients who most likely benefit from targeted medicines. The rising treatment and diagnostic costs have led to the question about whether health technologies (including medicines) truly add value to patients and society (65).

One approach to controlling costs without comprising outcomes, is to identify and abandon high-cost aspects of care that add no or very little benefit. In many European countries, national authorities already systematically evaluate the added value of new health technologies (including medicines) to make decisions regarding their reimbursement (61). Evaluations of effectiveness of the new technology compared to the standard of care, and sometimes also cost-effectiveness are tools used by these authorities (66). While different components of cancer care contribute to the total costs, evaluations of the (cost)effectiveness are in theory relevant for any type of health technology. In practice, however, reimbursement decisions are typically limited to medicines, owing to the nature and regulations of the pharmaceutical market (67).

DECISION-MAKING IN ONCOLOGY: WHO, WHY, WHAT?

In an era of rapid development of medicines and other technologies for cancer, value considerations have become more and more important for different stakeholders (65). Value in terms of efficacy and safety is relevant for regulatory agencies (and Notified Bodies) that decide whether medicines and technologies can be allowed on the market. The value in terms of added effect and costs is relevant to policymakers who need to decide on reimbursement. Finally, the value of a medicine or technology in improving a patient's life is relevant to clinicians who must determine which treatment is best for which patient. Unfortunately, these important decisions often have to be made based on uncertain evidence (68). Below a detailed description is given of each of the stakeholders (regulators, policymakers, clinicians), their roles in the decision-making process, why choices need to be made, and what evidence is available to inform the decisions.

Marketing authorization: Is it safe, efficacious and of sufficient quality?

The question of whether health technology is allowed to be sold on the market is nowadays a concern of regulatory agencies (for medicines) and Notified Bodies (for medical devices). History has shown that market forces alone do not sufficiently ensure that technologies are safe and effective. In 1956, a medicine named thalidomide was marketed as a sedative and used, among other things, to reduce nausea in pregnant women. At that time, no specific proof of medicine safety was required. While several studies on thalidomide were performed, these did not include testing in pregnant animals (or humans), nor was any randomized study performed. The lack of rigorous evidence resulted in a disaster. By the end of the 1950s, physicians observed a remarkable increase in the number of babies born with missing or shortened limbs. These birth defects were soon thereafter linked to the use of thalidomide. In 1961, thalidomide was withdrawn from the market by the pharmaceutical company (69).

The disaster of thalidomide resulted in changes to the regulatory system for medicines worldwide. Not only were explicit phases of testing defined (phase I – III clinical studies), but randomized studies, in large groups of patients, to document the efficacy of medicines became the new standard. Most of these requirements, although slightly adjusted, are still in place today (69). To obtain marketing authorization of a new medicine or indication, pharmaceutical companies must submit a comprehensive dossier that includes all data generated during the different phases of study. Regulatory agencies in Europe, the European Medicines Authority (EMA), then assess all this evidence and decide on approval (70). Notably, the procedure for medical devices, including high-risk devices, such as implants, is regulated differently. Unlike medicines, which require RCTs to generate insights into safety and efficacy, medical devices are only required to perform 'clinical investigation' to verify safety, performance, and the benefit to risk ratio (not effectiveness) in Europe. These investigations do not necessarily have to be RCTs, nor do they need to adhere to specific requirements concerning the study design (47). While regulations have been adapted recently, after several scandals with widely used hip and breast implants, evidentiary standards of devices are still lower than that of medicines (47, 71).

While the three-phase process of medicine testing and the thorough evaluation by regulatory agencies have been implemented for obvious reasons, there is also a downside to it: it takes a lot of time (72). Patients with life-threatening forms of cancer, without effective treatment options available, are not willing to wait for access to promising medicines. To facilitate earlier access, the EMA has implemented different pathways to expedite the process of studies and approval for certain medicines (73, 74). These pathways require less rigorous evidence. For instance, in order to reduce follow-up time,

studies using surrogate outcomes—defined as outcomes that are reasonably likely to forecast clinical benefit (in oncology usually progression free survival)—are accepted (74). Such outcomes are for traditional approvals not accepted. Additionally, single arm studies are sometimes sufficient to bring products to the market (75). Increasingly, these are the routes by which oncology medicines are approved. At the time a decision of approval must be made, the evidence of efficacy and safety is consequently increasingly uncertain (68, 76).

Reimbursement: (How much) should we pay?

The process of marketing authorization is independent of the decision to reimburse a new medicine or device. The latter are made at the national level, while the former are made at the level of the European Union. Typically, the aim of reimbursement authorities is to optimize health outcomes within a budget-constraint environment. To support reimbursement decisions, each country has its own specific evidence criteria. Some require only evidence of effectiveness, whereas others are also interested in cost-effectiveness (additional costs/additional effects) and budget impact (expected number of patients × price) (38, 66).

The key source of data to inform reimbursement decisions is usually obtained from the registration studies that are used to gain marketing authorization (68). However, the requirements of reimbursement agencies are not considered in the design of these studies. This brings additional uncertainties, on top of those already faced by regulatory agencies. That is, healthcare authorities are interested in the effectiveness of the new product compared to the standard of care in that country. This is not always the same as the comparator arm in the clinical study. Moreover, reimbursement agencies are interested in the extent to which new treatments improve survival and quality of life. With the changed criteria of regulatory agencies (e.g., EMA), robust information on survival is often not available yet by the time reimbursement decisions need to be made. Finally, what is truly relevant for health authorities is what the value of new medicines and devices is in daily clinical practice (68, 77, 78). Since the population in daily practice is broader and more heterogeneous than patient populations in RCTs, and because patients in practice are not treated under controlled circumstances, this remains unknown at the time a decision must be made.

Clinical: Which treatment for which patient?

The question of which treatment among the available, approved and reimbursed alternatives best fits the individual patient, including considerations related to the appropriate timing, dose and duration, is the responsibility of the treating physician in consultation with the patient. Clinical guidelines inform the decision, but oncology is a

quickly evolving medical field. As such, the latest advances are often not yet included in the guidelines (79). Moreover, RCTs of medicines typically evaluate one (often a new) treatment, compared to another. The relative value of all competing therapeutic options remains uncertain (80).

To fill this gap, several organizations provide advice regarding the value of new anti-cancer medicines. In Europe, the ESMO (European Society for Medical Oncology) has developed a standardized tool, the Magnitude of Clinical Benefit Scale (MCBS), to grade the clinical benefit of new medicines. The scale ranges from 1 to 5 for non-curative medicines and from A to C for curative medicines, with 1 and A representing the highest scores. Evidence of overall survival, toxicity, quality of life and surrogate outcomes (e.g., progression free survival) is considered in the scale (41, 65). Typically, the scores are applied to all new medicines approved by the EMA and updated when new evidence becomes available. The ESMO-MCBS can be used by clinicians to gain a better understanding of the actual effectiveness of a medicine, but also to weigh the relative value of treatment alternatives in the absence of direct comparisons (65). In the Netherlands, the committee for evaluation of oncology medicines (CieBOM) additionally performs assessments and publishes advice regarding the value of new medicines to treat cancer using the PASKWIL criteria (Palliatief, Adjuvant, Specifieke bijwerkingen, KWaliteit van leven, Impact van behandeling, Level of evidence - PASKWIL) (81, 82).

Important to note is that all assessments are based on the clinical studies performed to obtain marketing authorization (68). Predictably, clinicians face similar uncertainties as policymakers and regulators: What is the added value of a medicine compared to the standard of care in their country and patient population in terms of outcomes that are relevant for patients? Other unanswered questions include those of what medicines can safely be used in a broader patient population (e.g., the elderly), what the optimal treatment duration is, what sequence of different treatments delivers the best outcomes, and what long-term issues can be expected (80).

DECISION-MAKING IN ONCOLOGY: WHAT IS NEXT?

Each of the above-mentioned stakeholders must make decisions with imperfect evidence available. The proliferation of RWD in routine practice provides opportunities to use these data, both prior to approval and post-approval. Nowadays, RWD is used by regulatory agencies to evaluate long term safety and to provide information about the natural history and epidemiology of a disease or the standard of care (83). For reimbursement decisions, RWD is also used to generate contextual information regarding

the disease and its treatment, as well as to populate cost-effectiveness models. Moreover, in some countries, including the Netherlands, post-reimbursement registries have been developed to generate better insights into the (cost)effectiveness of expensive medicines in daily practice (84). Finally, the number of publications in oncology based on RWD has increased over the years (27, 56, 85). Results from such studies may be used by clinicians to support their decisions.

So far, RWD studies in any type of decision have been used on a rather ad hoc basis, but all stakeholders see a broader role of RWD in the decision-making process. For example, RWD may be used to generate a historical comparator arm when it is not possible to conduct a randomized study, to support regulatory approval of new indications of medicines, to provide insights in long-term outcomes, and offer information on effectiveness and costs in daily practice post-reimbursement, or in particular subgroups underrepresented in RCTs (e.g., elderly) (49, 86). However, currently there is still a lack of trust in RWD studies due to doubts about the quality of data and validity of applied methods. In order to further unlock the potential of RWD, guidance for conducting such studies, both in general and specifically within oncology, have been developed by different stakeholders (49, 50, 51, 87). Core principles of these frameworks include, ensuring quality of RWD, addressing the key risks of bias by using appropriate analytical techniques, and generating transparency in each phase of a RWD study (planning, conduct, reporting).

In this thesis, we explored the potential of different RWD sources to generate insights into cancer care and outcomes to support decision-making. We focus on retrospectively collected data; thus, data collected without the goal of answering a specific research question. The thesis begins with a description of data sources used in this thesis including legal aspects (**Part 1, Chapter 2**). We assess how RWD has been used in literature to generate evidence of comparative effectiveness in oncology (**Part 1, Chapter 3**). In **Part 2**, we present examples of how RWD can contribute to filling research gaps in oncology. Specifically, we used RWD in a modelling study to assess the cost-effectiveness of a medical device (**Chapter 4**), in a study to evaluate utilization and access to new medicines (**Chapter 5**) and to evaluate long-term outcomes in large groups of patients (**Chapters 6** and **7**). In **Part 3** (**Chapter 8**), we applied state-of-the-art guidance developed by regulatory and health technology assessment organizations to identify relevant RWD and assess its suitability for a defined research question. Finaly, we reflect on our experiences with gathering, analyzing, and interpreting the data in the discussion. Following the core principles of RWD frameworks, we discuss the challenges and opportunities for improvement regarding the broader use of RWD in oncology decision-making.

PART 1

REAL-WORLD DATA: SETTING THE SCENE

CHAPTER 2

REAL-WORLD DATA SOURCES

Real-World Data (RWD) refers to different types of routinely collected data (88). In the field of oncology, important types of RWD include electronic health records (EHR), administrative claims data and registries (55). Each of these data sources has its own strengths and limitations. In this thesis, we primarily used data from the Netherlands Cancer Registry (NCR) and administrative claims data and examined the possibilities of linking the two data sources. According to RWD guidelines, an essential step in any type of RWD study is to carefully consider the data source (49). Understanding the original purpose of RWD, possible subsequent changes, data content, coverage accessibility quality (completeness, accuracy) is import for determining whether the RWD source is suitable for a specific research question (49, 55). We start this thesis by introducing the main RWD sources used in our research and briefly discuss legal consideration that may influence data usability.

The Netherlands Cancer Registry

Purpose and origin

The NCR is hosted by the Netherlands Comprehensive Cancer organization (IKNL) and is a nation-wide registry which has the objective to provide insights into cancer epidemiology, care and outcomes of patients with cancer. It includes retrospectively collected data of newly diagnosed patients with cancer in the Netherlands (89). Newly diagnosed patients are notified by the national archive of pathology (PALGA [in Dutch: Pathologisch Anatomisch Landelijk Geautomatiseerd Archief])) and the National Registry of Hospital Discharge and Diagnoses hosted by the Dutch Hospital Data (DHD). Once identified, data managers record the relevant data, directly from electronic health records (EHRs). Annually, NCR data are merged by the Municipal Personal Record Database in order to retrieve information on patients last available vital status (alive or death) (90).

Quality and completeness

The NCR data includes >95% of all patients with cancer in the Netherlands (89). It includes information on tumor characteristics, diagnostic procedures, diagnosis, initial treatment and outcomes (overall survival). The patient, tumor and treatment information in the NCR is quite comprehensive when compared to essential elements of cancer datasets reported in the literature (91). It includes diagnostic information (morphology, topography, grade, behavior, stage, date of diagnosis) and patients' demographic information (sex, date of birth and postal code) and for selected patients' information on performance status and co-morbidities. It codes according to national and international guidelines and coding systems (92, 93, 94). It is also quite comprehensive with regard to radiation and surgical procedures performed, as well as to systemic treatment. It reports details about procedures and the type of systemic treatment (i.e., substance) as well as

the start date of prescription (and for chemotherapy also stop date) (95, 96). Typically, information on the key variables collected contain few missing values.

Despite its completeness there are a couple of aspects that the NCR does not include but which are relevant for various research question in oncology. Mainly, cancer recurrences (e.g., metastatic cancer years after diagnosis of primary disease which are often not pathologically confirmed, and since pathology is the main notification source of the NCR this is not identified) and information on treatment that is given beyond first line (usually medicines for metastatic disease). Additionally, it does not include outcomes other than overall survival, such as information on quality of life and progression free survival. Additional data can be collected on request, but this requires additional funding from an external organization (95).

Relevance

The NCR data are relevant to answer various research questions. The primary purpose of the NCR is to describe the nature and the burden of cancer in the Netherlands via stage at diagnosis, incidence, prevalence, mortality and survival statistics published on www.cijfersoverkanker.nl. These numbers can, for instance, can be used to identify trends, to compare statistics and trends between countries (or over time) in order to identify possible disparities and drive improvement. Additionally, the NCR data are used to evaluate population screening (e.g., breast cancer screening), to describe patterns of care and to support various epidemiological or clinical studies (91, 95). The NCR can be enriched by additional registration or linkage to other sources.

Accessibility and legal aspects

Patient-level data from the NCR is available on request. Stakeholders interested in these data must therefore submit a formal request including a protocol, research objectives and variables of interest to answer the research question. Permission has to be obtained through the Privacy Committee (looking at privacy of patients, care givers and hospitals) and the tumor specific Scientific Committees (looking at e.g. relevance of the data items requested and quality of the research question). Depending on the type and amount of data requested, the data can either be shared via a protected transfer program or an account is created within the protected cloud of IKNL. The dataset is anonymized for external users (e.g., variables containing a date are transformed). For scientific research, NCR data are typically provided at no charge, unless additional data, linkages or preparation is needed (97). Authorized internal users (employees of IKNL) have access to the complete dataset (including dates) under specific circumstances (98).

The NCR uses personal data as prescribed in the General Data Protection Regulation (GDPR) and other applicable laws and regulations in the Netherlands that aim to protect personal data (98, 99). Typically, laws relating to data protection state that patient data can only be used for scientific research if a patient gave consent to do so (100). Under specific circumstances, it is allowed to use patient data without specific consent. For instance, when it is not reasonably possible to gain consent (e.g., patients died, the sample size is too large). In such circumstances, additional requirements hold for the use of patient data for research (e.g., the organization needs to implement an transparent opt-out system) (100). The NCR operates under this exceptional rule.

Administrative Claims Data

Purpose and origin

Administrative claims data are data are recorded for reimbursement purposes and include information about diagnoses and health care services provided (e.g., health care utilization, including diagnostic procedures, inpatient care and prescribed medications) during patient visits to health care facilities (101). In the Netherlands these data are collected and stored by various organizations, including hospitals, insurance companies, data aggregators. The latter are organizations often established to enhance the quality of care by combining data from multiple sources. In this thesis we used the data from a data aggregator: the Dutch Hospital Data, and claims data from two individual hospitals in the Netherlands.

DHD is an organization that collects claims data from all hospitals in the Netherlands. Its objective is to improve the quality of healthcare in the Netherlands by collecting, analyzing, and sharing hospital claims data. Researchers can request access to this data, but availability may vary based on the specific research question and the type of data requested (102). Hospitals collect their own claims data primarily for billing and reimbursement in order to receive payment for the services delivered.

Quality and completeness

Typically, claims data is structured using codes, which facilitates the translation of data from care systems into an analytical dataset. The completeness of this data varies based on the organizations that collect or store it. Regarding the data used in this thesis, DHD offers comprehensive coverage, it included claims data from all hospitals in the Netherlands (103). Claims data from individual hospitals only includes the services provided at those specific facilities, although it may, under certain circumstances, be integrated with other patient data available within the hospital (104).

Because claims data are generated for reimbursement purposes, it can be expected that the accuracy of data concerning specific costly treatments (e.g., surgery, expensive oncological medicines) is high (55). However, there may be information that is relevant for cancer research but not for reimbursement, that is not accurately registered (e.g., exact date of treatment). Additionally, patient, tumor and outcome information in administrative claims datasets are typically limited (55).

Accessibility and legal aspects

In the Netherlands, patient claims data can generally only be used for the purposes for which it was originally collected (i.e., reimbursement), unless patients gave explicit consent to use their data for additional purposes (100, 104). This principle is in line with data protection regulations, including the General Data Protection Regulation (GDPR) (99). Hospitals in the Netherlands and the DHD are allowed to use administrative claims data to improve the quality of care (without explicit consent), but the data can only be used for research under specific conditions. That is, the data must be completely anonymized which means that all personally identifiable information has been removed from the data. This includes removing names, addresses, identification numbers, dates and any other data that could potentially link back to an individual. Additionally, any data that could enable re-identification should also be excluded. For example, patient IDs in the hospital should be replaced with random numbers, and the key used for this replacement must be discarded afterward (100, 104).

CHAPTER 3

REVIEW OF THE QUANTITY & QUALITY OF REAL-WORLD DATA STUDIES

Published as:
Assessment of Studies Evaluating Incremental Costs, Effectiveness, or Cost-Effectiveness of Systemic Therapies in Breast Cancer Based on Claims Data: A Systematic Review.

Luyendijk M, Vernooij RWM, Blommestein HM, Siesling S, Uyl-de Groot CA 2020.
Value in Health. 23(11):1497–1508

ABSTRACT

Objectives: Large secondary databases, such as those containing insurance claims data, are increasingly used to compare the effects and costs of treatments in routine clinical practice. Despite their appeal, however, caution must be exercised when using these data. In this study, we aimed to identify and assess the methodological quality of studies that used claims data to compare the effectiveness, costs, or cost-effectiveness of systemic therapies for breast cancer.

Methods: We searched Embase, the Cochrane Library, Medline, Web of Science, and Google Scholar for English-language publications and assessed methodological quality using the Good Research for Comparative Effectiveness principles. This study was registered with PROSPERO (number CRD42018103992).

Results: We identified 1251 articles, of which 106 met the inclusion criteria. Most studies were conducted in the US (74%) and Taiwan (9%) and were based on claims datasets (35%) or claims data linked to cancer registries (58%). Furthermore, most included large samples (mean 17,130 patients) and elderly patients, and they covered various outcomes (e.g., survival, adverse events, resource use, and costs). Key methodological shortcomings were the lack of information on relevant confounders, the risk of immortal time bias, and the lack of information on the validity of outcomes. Only a few studies performed sensitivity analyses.

Conclusions: Many comparative studies of cost, effectiveness, and cost-effectiveness have been published in recent decades based on claims data, and the number of publications has increased over time. Despite the availability of guidelines to improve quality, methodological issues persist and are often inappropriately addressed or reported.

INTRODUCTION

Data from randomized controlled trials (RCTs) are typically used to show the comparative effectiveness of treatment options and are often used in cost-effectiveness analyses to support drug reimbursement decisions (105). However, although RCTs provide a gold standard methodology for assessing efficacy and safety, it is increasingly recognized that they fail to reflect either the true effectiveness or, if measured, the true costs of a treatment in daily practice because they use highly selected cohorts in controlled conditions (43, 54, 106). Moreover, the primary endpoints of many phase III RCTs are only intermediate outcomes, such as relapse-free or progression-free survival, whereas final outcomes are more relevant to patients, such as survival and quality of life (46, 107, 108). To improve our understanding of treatment effects and costs in routine clinical practice, healthcare decision makers are therefore increasingly prioritizing real-world data (RWD) collected from sources other than traditional RCTs (88, 101, 109).

The digital era has resulted in a proliferation of RWD (56). Large electronic databases, rich in longitudinal patient information for large cohorts, are routinely being generated as a byproduct of clinical care and financial transactions (110). Typical examples are electronic health records and claims and billing databases,(111) and although the information they contain is not collected or stored for research purposes, it can be used to assess the costs, effectiveness, and cost-effectiveness of drugs used in routine clinical practice (110, 112). Billing and claims data, in particular, are an appealing source to researchers because they are relatively inexpensive (given the large sample sizes), easily accessible, and structured with codes (112, 113, 114). Consequently, such data sets have been used not only to examine care patterns (e.g., guideline adherence or regional variation) and costs but also to evaluate treatment effectiveness (115, 116, 117, 118).

Despite the clear benefits of claims data, several concerns must be addressed when using them for comparative research. For instance, the data may lack important clinical information (e.g., diagnosis and outcomes), we may be uncertain of the accuracy of the codes, and we must acknowledge that the data may only reflect a select patient population (e.g., the insured) (110). The absence of treatment randomization is another important concern inherent to all observational data (119, 120). These issues threaten the validity of study results and the usefulness of those results for decision makers. Nevertheless, many issues can be addressed through proper research design, appropriate analysis, and standardized reporting (101). To guide comparative research based on non- RCT data (including claims data) and to assist decision makers in judging the validity of such studies, several good research practice guidelines have been developed. These include the Good Research for Comparative Effectiveness (GRACE) initiative, the ISPOR series on Good

Practice for Comparative Effectiveness Research, and a technical support document by the National Institute for Health and Care Excellence (110, 121, 122, 123, 124). The extent to which these recommendations are implemented by researchers is currently unclear. In the present study, we aimed to describe the quantity and to assess the quality of published (cost)effectiveness research of systemic therapies for breast cancer based on claims data. We focused on breast cancer because it is one of the most prevalent cancers worldwide, placing significant health and financial burdens on society,(125, 126, 127) and because the number of innovative therapies for this disease is increasing (128, 129).

METHODS

Search strategy

A systematic review was conducted to obtain an overview of all (cost-)effectiveness studies of systemic therapies for the treatment of breast cancer based on claims data. We primarily searched the Embase, Cochrane Library, Medline, Web of Science, and Google Scholar databases; however, this was supplemented by screening the reference lists of studies deemed relevant based on full-text reviews of the introductions and discussions of selected articles. A detailed list of the keywords used for each database can be found in File 1 in Supplemental Materials, available online. The searches focused on titles and abstracts, had no time restrictions, and were restricted to articles published in English. The main database search was conducted on July 4, 2018. The study protocol was registered with the International Prospective Register of Systematic Reviews before conducting the review (number CRD42018103992).

Study Selection

The full inclusion and exclusion criteria are detailed in File 2 in Supplemental Materials, available online. Broadly, we included studies based on the PICO framework:(130) (1) patients were diagnosed with invasive/metastatic breast cancer; (2) the analysis compared interventions (or compared with none) based on patient-level claims data (including those linked to other sources); (3) interventions included systemic anticancer treatments, with or without radiotherapy or surgery, and with no restriction on comparators; and (4) outcomes were either costs (or resource use), effects (clinical outcomes, adverse events, including cancer recurrence and the development of other disease in later life, treatment switching, or patient-reported outcomes), or both (i.e., cost-effectiveness). Titles and abstracts were screened for inclusion by 2 independent reviewers (ML and RV) before they reviewed the full texts of articles to identify those eligible for data extraction. At both stages of article selection, disagreements were resolved by discussion, deferring to a third reviewer (SS or HMB) to make the final decision when no agreement could be reached.

Data Extraction

Data were extracted from each study using a form designed by ML and pilot tested by ML and RV. The following data were collected: author, year of publication, country, time horizon, study design, database, sample size, patient selection method, patient characteristics, treatment type, comparison type, outcome type, outcome measures, and statistical methods.

Quality Assessment

All included studies were assessed for methodological quality by 2 independent researchers (ML and RV), using the validated GRACE checklist. This checklist was specifically developed to evaluate the quality of comparative effectiveness research based on observational data (124). It comprises 6 items concerning data quality (e.g., availability of information in the data set) and 5 items concerning the methodology (e.g., study design and analysis). Additional criteria were also used to improve consistency among the researchers and across the assessment of the included publications. These criteria were based on previous literature,(131, 132) pragmatic literature searches, and the researchers' judgments. The full checklist, including descriptions of the main and additional criteria, can be found in Table 1 and File 3 in Supplemental Materials, available online. We planned to resolve disagreements in the quality assessment through discussion until consensus was reached.

Table 1. GRACE checklist and additional criteria.

	Description	Answer options GRACE (124)	Additional criteria	Source (additional criteria)
D1	**Were treatment and/or important details of treatment exposure adequately recorded for the study purpose in?**	**YES:** Reasonably necessary information to determine treatment or intervention was adequately recorded for study purposes (e.g., for drugs, sufficient detail on dose, days supplied, route, or other data important).	**YES:** If there is some information on dose/duration (e.g., no. of treatment lines).	Reference (131)
		NO: Data source clearly deficient or not enough information in article.	**NO:** No information on dose/duration.	

Table 1. GRACE checklist and additional criteria. (continued)

	Description	Answer options GRACE (124)	Additional criteria	Source (additional criteria)
D2	**Were the primary outcomes adequately recorded for the study purpose (e.g., available in sufficient detail through data source(s))?**	**YES:** Information to ascertain outcomes were adequately recorded in the data source (e.g., if clinical outcomes were ascertained using ICD-9 diagnosis code(s) in an administrative database, the level of sensitivity and specificity captured by the code(s) was sufficient for assessing the outcome of interest.)	**YES**: 1) The outcome(s) based on algorithm/codes and the sensitivity, specificity, or PVV is reported in the article or in in the article referred to. 2) Mortality outcomes based on cancer registry and/or death registry. 3) Other outcomes obtained through medical chart review.	
		NO: data source clearly deficient (e.g., the code(s) captured a range of conditions that was too broad or narrow, and supplementary information such as that from medical charts was not available), or not enough information in article.	**NO:** One or more of the outcomes used for the analyses is based on an algorithm or codes AND the algorithm/code is not validated and/or sensitivity, specificity, or PVV are not reported in the paper or reference.	
D3	**Was the primary clinical outcome(s) measured objectively rather than subject to clinical judgment (e.g., opinion about whether the patient's condition has improved)?**	**YES:** clinical outcomes were measured objectively (e.g., hospitalization, mortality). **N/A:** primary outcome not clinical (e.g., PROs).	n/a.	
		NO: e.g., clinical opinion about whether patient's condition improved, or not enough information in article.	n/a.	

Table 1. GRACE checklist and additional criteria. (continued)

	Description	Answer options GRACE (124)	Additional criteria	Source (additional criteria)
D4	**Were primary outcomes validated, adjudicated, or otherwise known to be valid in a similar population?**	**YES:** Outcomes were validated, adjudicated, or based on medical chart abstractions with clear definitions, e.g., a validated instrument was used to assess patient reported outcomes (e.g., SF-12 Health Survey); a clinical diagnosis via ICD-9 code was used, with formal medical record adjudication by committee to confirm diagnosis or other procedures to achieve reasonable sensitivity and specificity; billing data were used to assess health resource utilization, etc.	**Yes**: 1) The outcome(s) based on algorithm/codes and the sensitivity, specificity, or PVV is reported in the article or in in the article referred to. 2) Mortality outcomes based on cancer registry and/or death registry. 3) Other outcomes obtained through medical chart.	
		NO: No, or not enough information in article.	**NO:** One or more of the outcomes used for the analyses is based on an algorithm or codes AND the algorithm/code is not validated and/or sensitivity, specificity, or PVV are not reported in the paper or reference.	
D5	**Was the primary outcome(s) measured or identified in an equivalent manner between the treatment/intervention group and the comparison group(s)?**	**YES.**	n/a.	
		NO, or not enough information in article.	n/a.	
D6	**Were important covariates that may be known confounders or effect modifiers available and recorded?**	**YES:** most if not all important known confounders and effect modifiers available and recorded (e.g., measures of medication dose and duration).	**YES:** A list of important confounders/covariates was determined with a pragmatic literature search per outcome and disease stage (see S3). This item was judged to be sufficient (YES) if the confounders/covariates in the list were available in the dataset of the study.	S3
		NO: At least one important known confounder or effect modifier not available and recorded (as noted by authors or as determined by user's clinical knowledge), or not enough information in article.	**NO:** If one or more of the specified confounders/covariates were missing.	S3

Table 1. GRACE checklist and additional criteria. (continued)

	Description	Answer options GRACE (124)	Additional criteria	Source (additional criteria)
M1	**Was the study (or analysis) population restricted to new initiators of treatment or those starting a new course of treatment?**	**YES**—Only new initiators of the treatment of interest were included in the cohort, or for surgical procedures and devices, only patients who never had the treatment before the start of study follow-up were included.	**YES:** If efforts were done to exclude patients who could have had the treatment before. Or if it is very unlikely that patients had the treatment before (e.g., first treatment after first primary diagnosis of breast cancer).	
		NO: or not enough information in article.	**NO:** If it is possible that patients had the treatment before and no efforts were done to check if patients had the treatment before diagnosis (e.g., enrolment in insurance at least a certain period prior to diagnosis) and/or to exclude patients who could have had the treatment before (e.g., patients with other cancer prior to breast cancer or recurrent disease).	
M2	**If one or more comparison groups were used, were they concurrent comparators? If not, did the authors justify the use of historical comparisons group(s)?**	**YES:** Data were collected during the same time period as the treatment group ("concurrent") or historical comparators were used with reasonable justification, e.g., when it is impossible for researchers to identify current users of older treatments or when a concurrent comparison group is not valid—(i.e., uptake of new product is so rapid that concurrent comparators differ greatly on factors related to the outcome).	**YES**: The timeframe of patients selection was ≤ 3 years (e.g., patients diagnosed with breast cancer between 2002 and 2004) or the timeframe was >3 years and it was evident that treatments were provided during the same time period.	
		NO: Historical comparators used without being scientifically justifiable, or not enough information in article.	**NO:** The timeframe of patient selection was > 3 years and it was unclear when treatments (i.e., intervention and comparator(s)) were provided.	

Table 1. GRACE checklist and additional criteria. (continued)

	Description	Answer options GRACE (124)	Additional criteria	Source (additional criteria)
M3	**Were important covariates, confounding and effect modifying variables considered in the design and/or analysis?**	**YES:** Most if not all important covariates that would be likely to change the effect estimate substantially were accounted for, e.g., measures of medication dose and duration.	**YES:** A list of important confounders/covariates was determined with a pragmatic literature search per outcome and disease stage (see S3). This item was judged to be sufficient (YES) if the confounders/covariates in the list were considered in the analysis.	S3
		NO: Some important covariates were available for analysis but not analyzed appropriately, or at least one important covariate was not measured, or not enough information in article.	**NO:** If one or more of the specified confounders/covariates were not considered in the analysis.	S3
M4	**Is the classification of exposed and unexposed person-time free of ITB?**	**Yes.**	**YES**: If the study is at low risk of ITB: 1) Time-dependent covariates analysis was used. 2) Landmark analyses or restriction in selection of patient populations (e.g., patients had to be alive at a certain period). 3) If it was highly unlikely or impossible for an outcome to occur prior to the start of treatment (e.g., recurrence/mortality prior to the start of adjuvant chemotherapy in early stage breast cancer). 4) Comparison of two or more treatments without a comparison to no treatment.	Reference (132)
		NO: Or not enough information in this article.	**NO:** At high risk of ITB.	
M5	**Were any meaningful analyses conducted to test key assumptions on which primary results are based?**	**Yes:** And primary results did not change substantially. **Yes:** But primary results changed substantially.	n/a.	
		None reported, or not enough information in article.	n/a.	

Abbreviations: GRACE, Good Research for Comparative Effectiveness; ICD: International Classification of Diseases; ITB: Immortal Time Bias; S3: File 3 in Supplemental Materials; SF-12: Short Form; PROs: Patient Reported Outcome Measures; PVV: positive predictive value

Data Analysis

The characteristics of the selected studies are presented as numbers and percentages for categorical variables and as means and standard deviations (minimum to maximum) for continuous variables. For the quality assessment, we estimated the proportion of studies that fulfilled each criterion and compared the GRACE scores before and after 2010. This cutoff was chosen because good research guidelines had been published at the end of 2009 (110, 122, 133, 134).

RESULTS

Descriptive Statistics

We identified 1251 unique studies, and we excluded 1047 based on title and abstract review and 98 based on full-text review. Thus 106 studies were included for data extraction and quality assessment (Figure 1; Files 4 and 5 in Supplemental Materials, available online).

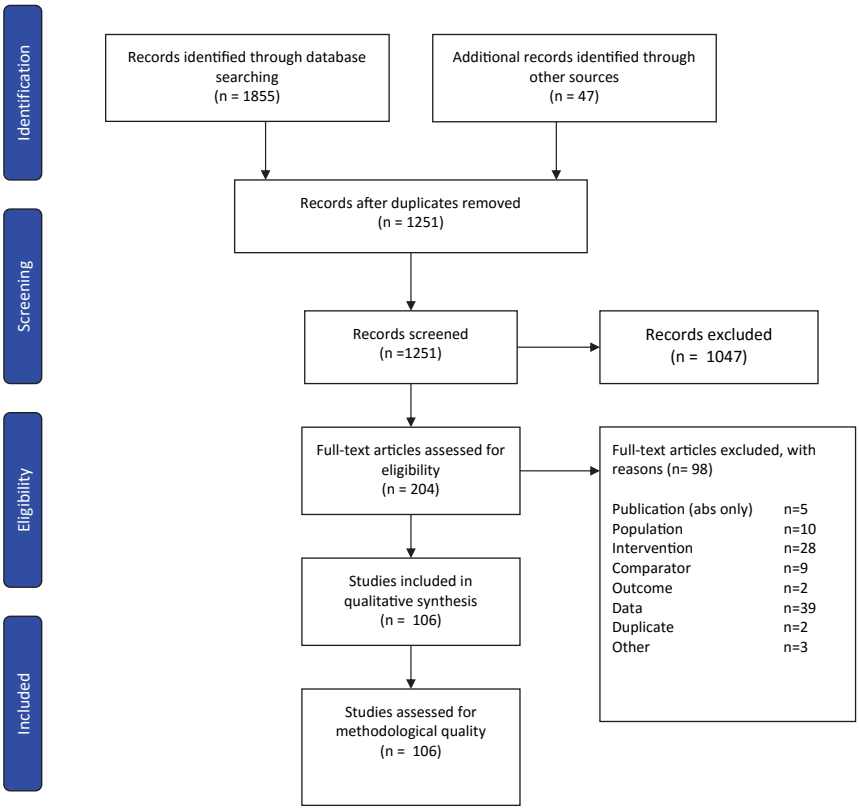

Figure 1. PRISMA flowchart for the inclusion and exclusion of publications.

Study Details

A full list of the included studies and their details is presented in File 6 in Supplemental Materials, available online. In the following text, we describe specific features of the included studies in more detail.

Publication year, country, database, design

The earliest study in this review was published in 2002, and the number of studies showed a trend to increase over time (Figure 2a). Most studies originated from the United States (78 studies; 74%), Taiwan (10 studies; 9%), and Canada (7 studies; 7%), with the remainder from other countries (11 studies; 10%) (Figure 2b). About one-third (37 studies; 35%) used claims data not linked to a database with patient and/or clinical characteristics, but the remainder linked claims data with either cancer registry data (61 studies; 58%) or other sources (e.g., RCTs; 8 studies; 7%) (Table 2). The Surveillance Epidemiology and End Results–Medicare linkage database was used most frequently (41 studies; 39%), followed by the Truven MarketScan database (claims data only; 12 studies; 11%) and the Taiwan National Health Insurance Research Database (claims data only; 10 studies; 9%). Finally, 83 (78%) of the studies compared only the effectiveness of treatments, 12 studies (11%) compared only the costs of different therapies, 9 (9%) estimated both the effects and incremental costs of therapies, and 2 (2%) performed full cost-effectiveness analyses, reporting the costs, effects, and incremental cost-effectiveness ratios (see Table 2).

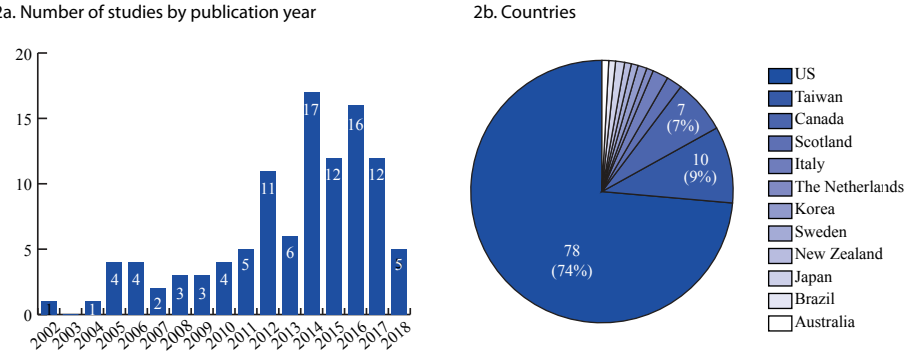

Figure 2. Publications per year and countries of the studies (N = 106)

Patient populations, sample size, cohort selection

Study populations comprised patients with locoregional disease (43 studies; 41%), metastatic disease (27 studies; 25%), any stage, (21 studies; 20%), or unclassified stage (15 studies; 14%). They also focused on patients aged ≥65 years (43 studies; 41%), adults aged about ≥18 years (29; 27%), or patients with no age limit specified (22 studies; 21%).

If specified, patients were rarely younger than 18 or 65 years. Sample sizes ranged from 175 to 190,620 (mean: 17 130; median: 6433). Samples for the linked data sets were mostly selected based on cancer registry data (52 studies; 49%) or a combination of an algorithm and cancer registry data (6; 6%), whereas claims-only studies were selected based on algorithms (33; 31%) or single claim codes (7; 7%). Unfortunately, most of the algorithms/codes were not validated (see Table 2).

Table 2. Key characteristics of the 106 included studies

Number of studies	**N = 106**		**%**
Types of databases			
Claims data	37		35%
Claims data linked to cancer registry data	61		58%
Claims data linked to other sources	8		7%
Study design			
Comparative effectiveness	83		78%
Costs comparison	12		11%
Comparison of effects and costs separately	9		9%
Costs-effectiveness	2		2%
Patient population			
Disease			
Locoregional	43		41%
Metastatic	27		25%
Both	21		20%
Not specified/other[1]	15		14%
Age group			
About ≥18 years[2]	29		27%
About ≥65 years[2]	43		41%
All patients[3]	22		21%
Other[4]	12		11%
Data	**Mean (SD)**	**Median(min/max)**	**N**
Sample size	17,130 *(26,438)*	6,433 *(175/190,620)*	106
Study period in years[5]	7.8 *(4.7)*	7.0 (0/life time[6])	106
Median follow-up in months[7]	40 *(22)*	n/a	39
Treatment and comparator			
Comparison of[8]:			
a) *≥1 therapy vs no therapy*	41		39%
b) *≥2 therapies*	36		34%
c) *Different timing*[9]	13		12%
d) *Adherence vs non-adherence*[10]	12		11%
e) *Other*	4		4%
Treatment type			

Table 2. Key characteristics of the 106 included studies (continued)

Number of studies	**N = 106**	**%**
Chemotherapy	46	43%
Hormone therapy	27	26%
Targeted therapy	13	12%
Chemotherapy & Hormone therapy	3	3%
Chemotherapy & Targeted therapy	12	11%
Not specified	3	3%
All	2	2%
Outcomes	**N = 173** [11]	
Overall survival	37	21%
(Breast) cancer specific survival	16	9%
Recurrence or other cancer	16	9%
Treatment switching	9	5%
Adverse events [12]	48	28%
Health care costs (total costs/cancer related costs / both)	23 (n = 16/n = 5/n = 2)	13%
Health care visits	18	11%
Other [13]	7	4%
Cohort selection		
Cancer registry	52	49%
Cancer registry + algorithm	6	6%
Algorithm – adapted from literature	11	10%
Algorithm – validated	2	2%
Algorithm – not validated	20	19%
Single code with BC diagnosis	7	7%
Other	8	8%
Statistical methods for selection bias	**N = 140** [14]	
Stratification	19	14%
Regression analyses (multivariate)	82	58%
Linear regression	3	Total >82; multiple methods were used in some studies
Log/logit regression	14	
Poisson regression	2	
Negative binominal	2	
Tobit regression	1	
Generalized linear models [15]	11	
Cox regression	46	
Competing risk regression	1	
Cox regression with time-dependent confounders/covariates	16	
Propensity score methods	27	19%

Table 2. Key characteristics of the 106 included studies (continued)

Number of studies	N = 106	%
Propensity score matching	12	Total >27; multiple methods were used in some studies
Propensity score weighting	6	
Propensity score stratification	4	
Propensity score covariate adjustment	11	
Doubly robust methods [16]	2	
Not specified	2	
Instrumental variables	0	0%
None	11	8%
Other [17]	1	1%

Notes: 1) other is for instance stage 1 & 2 or stage 2–4; 2) the category about ≥18 years includes studies that defined the age as approximately 18 without an upper age limit, the category ≥65 defined the age as approximately 65; 3) in these studies age was not specifically specified (e.g., patients with invasive BC), we assumed no age limits were applied; 4) other includes >50/<65/>80, postmenopausal etc.; 5) study period includes period for selection of patients + follow-up; 6) lifetime was assumed to be 34 years as in this study pts of ≥65 years were included (max 99 years) & 0 years of follow-up was for a cross-sectional study; 7) median follow-up was reported by 38 studies, other studies reported mean follow-up/person years or did not report follow-up duration; 8) some studies made multiple comparisons (e.g., comparison of treatment vs not treatment & comparison of timing of the treated), we included only the main comparison; 9) different timing includes for instance delay in treatment initiation, treatment restarting after discontinuation; 10) most studies that compared adherence to no-adherence also compared persistence to non-persistence; 11) many studies evaluated multiple outcomes, the number of outcomes therefore sum up to more than 106; 12) short term adverse events (AEs) such as hospitalizations for neutropenia, fever, thrombocytopenia and long term AEs (e.g., development of cardiovascular events/diabetes/depressive disorders); 13) other outcomes e.g., out of pocket payments, incremental cost-effectiveness ratio, non-cancer survival; 14) some studies used more statistical methods thus does not add up to 106; 15) including generalized linear models & extended estimation equations; 16) combine the IPTW and regression model; 17) randomization
Abbreviations: BC: breast cancer; max: maximum; min: minimum; SD: standard deviation

Treatment and Comparator

Chemotherapy (CT) was evaluated in 46 studies (43%). These either assessed all agents, a specific subclass, or a pre-specified regimen (e.g., doxorubicin, cyclophosphamide, and taxane). Hormonal therapy (HT) was assessed in 27 studies (26%), typically evaluating the risk of developing health problems in the future (e.g., cardiovascular problems and diabetes) or the effect of treatment adherence on survival and related outcomes. Another 13 studies (12%) evaluated targeted therapy (TT), 3 (3%) compared CT with HT, 12 (11%) compared CT with TT, and some did not specify the treatment (e.g., "any" systemic therapy). The comparator varied among studies. Over one-third (41 studies; 39%) included no therapy as a comparator (e.g., HT vs no HT), and about one-third (36 studies; 34%) compared 2 or more therapies or regimes (e.g., some compared HT and CT whereas others compared multiple different CT regimens). Some studies compared the effects of different times to initiation of therapy (13 studies; 12%) and the effects of adherence and non-adherence (12 studies; 11%).

Outcomes

Multiple outcomes were evaluated in most studies (see Table 2), mainly related to survival (i.e., overall or breast cancer specific survival), adverse events, and recurrence. Other frequently studied outcomes included healthcare visits and costs. Of the studies that assessed costs, 5 measured cancer-related costs, 16 measured total costs (i.e., costs related to cancer care and costs of unrelated conditions), and 2 measured both types.

Statistical methods to control for selection bias

Most studies used multiple methods to control for selection bias, including regression analyses and propensity score matching, but none used an instrumental variable method. Only 11 studies did not attempt to control for confounding, but most of these did not intend to draw inferences about the relative effectiveness/ costs or were pilot studies for which sample sizes were too small to control for confounders (see Table 2).

Quality Assessment GRACE Checklist

The results of the quality assessment are summarized in Table 3 and File 7 in Supplemental Materials, available online. Although agreement was reached on all items, a couple of points are noteworthy. First, we had some difficulties reaching consensus on the issue of immortal time bias (ITB), so we asked a third reviewer to check 25 studies that we had initially disagreed on; however, this did not change our original conclusions. Second, we also had extensive discussions about the issue of relevant confounders, although again, we ultimately agreed on all items.

Only 17 studies (16%) reported information on the dose or duration of the study treatment (item D1). Fewer than half of the studies reported the accuracy of outcomes based on the International Classification of Disease/claim code(s) or algorithms. Given that many studies used adverse events or cancer recurrence as outcomes, GRACE items D2 and D4, which evaluate the recording and validity of primary outcomes, were not fulfilled by 63 (59%) and 62 (58%) studies, respectively. Nevertheless, the primary outcomes of most studies were based on either International Classification of Disease/claim codes or death registry data, resulting in 104 studies (98%) using objective outcomes (item D3). Given that the same claim codes/mortality data were used for the intervention and comparison group, 100% of the studies fulfilled the criteria for item D5. Many data sets lacked information on relevant confounders for the outcome, with 47 studies (44%) each meeting the criteria for items D6 and M3. For example, information on major risk factors (e.g., lifestyle indicators for cardiovascular diseases), disease severity (e.g., stage), performance status, and HER2 status were frequently unavailable in the data sets (see File 3 in Supplemental Materials (available online) for a full list of confounders deemed relevant per outcome).

In most studies, patients were selected based on specified criteria, such as the type of breast cancer and the period of diagnosis. Although most described a long timeframe for selection (e.g., diagnosed between 1998 and 2008), they often failed to describe when the treatments under comparison were given. As a result, item M2 was deemed to be sufficient in only 26 studies (25%) because we could not be certain that data were collected during the same period in 80 studies (75%). Authors sometimes required the inclusion of information about patients and treatments before the study follow-up period, such as requiring patients to be enrolled in an insurance program for at least 12 months before diagnosis so that the authors could select new treatment initiators for analysis. Only 45 studies (42%) therefore met the criteria for item M1.

Item M4 addressed the issue of ITB. Immortal time refers to a period during cohort follow-up in which the event of interest could not have occurred, with bias arising when this is not correctly dealt with in the study design or analysis (135, 136). The risk of ITB was high in 24 studies (23%) in our review; indeed, these made no attempts to deal with the issue in their design or analysis, despite the high likelihood that the event occurred before the exposure definition was fulfilled. Most of these studies evaluated treatment adherence and persistence or compared exposed and non-exposed patients over long periods within the definition of exposure (e.g., ever vs never users).

Table 3. Quality assessment using the GRACE checklist

Item	Question	Total	Pre-guidelines (2002–2009)	Post-guidelines (2010–2018)
		N = 106	N = 18	N =88
D1	Were treatment (details) recorded adequately?	17 (16%)	3 (17%)	14 (16%)
D2	Were outcomes recorded adequately?	43 (41%)	7 (39%)	36 (41%)
D3	Was the primary clinical outcome(s) measured objectively?	104 (98%)	18 (100%)	86 (98%)
D4	Was the primary outcome validated?	44 (42%)	7 (39%)	37 (42%)
D5	Were primary outcomes measured/identified in an equivalent manner between the treatment and comparator groups?	106 (100%)	18 (100%)	88 (100%)
D6	Were important covariates recorded?	47 (44%)	11 (61%)	36 (41%)
M1	Was the study restricted to new initiators of treatment?	45 (42%)	8 (44%)	37 (42%)
M2	Were comparison groups concurrent comparators?	26 (25%)	3 (17%)	23 (26%)
M3	Were important confounders considered?	47(44%)	11 (61%)	36 (41%)
M4	Is the analysis free of "immortal time bias?"	82 (77%)	15 (83%)	67 (76%)
M5	Were any sensitivity analyses performed?	37 (35%)	5 (28%)	32 (36%)

Finally, only 37 studies (35%) reported a sensitivity analysis (item M5). Nevertheless, unexpected differences were observed from before to after the good practice guidelines were published (see Table 3). Notably, after their publication, fewer studies included relevant confounders/covariates (items D6 and M3) and more studies appropriately reported sensitivity analyses (item M5).

DISCUSSION

We conducted a systematic review of studies assessing the incremental costs, effectiveness, or cost-effectiveness of systemic therapies for breast cancer based on claims data and rated the methodological quality of 106 included studies using the GRACE checklist. The earliest study was published in 2002, most were published in the United States and Taiwan, and there was a clear increase in number over time. Nevertheless, many studies had methodological shortcomings, and it was notable that the quality of studies did not improve after the publication of good practice guidelines.

The observed trend in the number of studies is not surprising given the proliferation of RWD and the growing interest in its use to improve clinical and regulatory decision making for oncology (110, 137, 138). RWD is believed to have the potential to complement evidence from RCTs thanks to certain well-known benefits, of which several were evident in the studies included in this review. First, we should note that many studies were based on the Surveillance Epidemiology and End Results Medicare–linked database in the United States, which has opened more opportunities to conduct studies using RWD. Second, the sample sizes were generally large, ranging from 175 to 190,620 (even covering entire countries, such as Taiwan (139)), which contrasts dramatically with the numbers in clinical trials (i.e., only 3% to 5% of all patients with cancer) (138). Third, about 40% of the studies included older patients (typically ≥65 years) and some even included the very old (>80 years), groups that are known to be underrepresented in RCTs (140). Fourth, diverse outcomes were studied in the claims-based studies, including healthcare resource utilization, costs, survival, and late or rare adverse events. These outcomes can be difficult to study in RCTs because of the controlled treatment settings, small sample sizes, and limited follow-up durations (46, 141, 142). Finally, claims studies allow for evaluations of different therapies that are not typically compared in head-to-head clinical trials; for example, combinations of doxorubicin, cyclophosphamide, paclitaxel, and trastuzumab were compared with docetaxel, carboplatin, and trastuzumab (118, 143).

Despite the advantages of using RWD, our review also showed that important methodological requirements were not always met in published studies. One frequently observed

issue, which is particularly relevant for comparative studies, was the lack of information about confounding variables. Most studies did link claims and cancer registry data, thereby including tumor and patient information (e.g., cancer stage), and many used proxies of health status (e.g., comorbidity index and performance status) (144, 145). Important confounders for the outcome of interest, however, were frequently lacking. In several studies that examined the effect of systemic treatments on cardiovascular problems, information on risk factors such as obesity or smoking status were not available in either the claims or the registry data (146). This absence of prognostic data is an established issue of RWD, especially in retrospective databases where researchers do not control the data collection (147). In these settings, failure to include all relevant confounders and to exclude all irrelevant confounders can lead to biased estimates of treatment effects (122).

Good practice guidelines recommend that all potential confounding factors be identified and that researchers preferably make use of directed acyclic graphs (DAGs) to visualize the relationships among variables (122). Such approaches make it possible to recognize relevant and irrelevant confounders, and they can guide researchers when interpreting model results (122). In our review, very few studies provided clear rationales for the choice of potential confounding variables, and none used DAGs. Given the trend toward using large longitudinal secondary databases for comparative research, we believe that it will become ever more critical to discuss and defend the reasons for including or excluding confounders. This transparency makes it easier for decision makers to judge the reliability of the observed effect.

Another quality concern that was quite common in the claims-based studies of this review was the high risk of ITB. This phenomenon can arise when there is a time period during cohort follow-up in which the event of interest cannot occur. Not appropriately accounting for this can produce misleading results (i.e., an overestimate of treatment effects) (135, 136). Previous systematic reviews have assessed the prevalence of ITB in observational studies published in high-quality medical journals (132, 148). Consistent with our research, they also found it to be surprisingly common (132, 136, 148, 149). Because ITB can be difficult for readers to recognize, we believe that researchers who use claims data should explicitly specify how they dealt with ITB (e.g., Chien et al),(139) in line with the RECORD-PE checklist (Reporting of Studies Conducted using Observational Routinely Collected Health Data Statement for Pharmacoepidemiology) (150).

We also found that outcomes based on claim codes or algorithms were often not validated and that performance characteristics (i.e., predictive values, sensitivity, specificity) were almost never provided. Validating outcomes is important because codes in claims data are not always accurate, and using inaccurate codes can lead to misclassification and loss of internal validity if the occurrence is not at random (i.e., classification bias)

(122). Such misclassification can vary by disease state, patient population, data source, and code (122, 151). Algorithms for cancer recurrence, an outcome frequently used in many studies in our review, have been shown to vary widely depending on the code and type of cancer. For example, the positive predictive values for breast, lung, and prostate cancers have been reported to be 30% to 87%, 72% to 94%, and, 30%, respectively (152). This supports the argument that authors should report their algorithm's performance to facilitate the accurate interpretation of their results.

Other issues were identified with the GRACE checklist. Few studies performed sensitivity analyses to explore how much of the estimated effect depended on underlying assumptions. This is an important omission because many assumptions are made about the patient population, the exposure, and the outcomes in claims- based studies, and because sensitivity analyses can provide insights into the extent to which study findings are dependent on them. Few studies also reported details of the doses or durations of the treatments under study, possibly because that information was unavailable or unreliable in their claims datasets (153). However, such information is particularly relevant when considering toxicity outcomes. Finally, it was often unclear whether the compared treatments were prescribed at the same or at different times during the study periods, preventing the reader from assessing the impact of change in the standard of care over time.

When using a checklist to grade the quality of studies, assessors must perform valid and reliable interpretations. We ensured validity by using the validated GRACE checklist for evaluating observation studies(124) and ensured reliability by using 2 independent assessors and a checklist clarification. Nevertheless, we still believe that aspects other than those listed in the GRACE checklist are relevant when evaluating studies based on claims data. For instance, the checklist does not cover either the selection of participants or the measurement of treatment exposure, and both are often identified by algorithms or codes and can result in misclassification. We found that few studies relying on claims databases alone used validated codes or algorithms to select samples. A similar finding was reported by Schulman et al (2014), who subsequently developed a checklist for selecting study cohorts in oncology research (154). This checklist appears to complement the GRACE checklist, and it could be suitable to combine these in future research (155). Another highly relevant consideration for claims-based studies is how results are interpreted in the context of other literature on the topic of interest. This is because confounding can never be totally excluded in the absence of randomization, which necessitates proper interpretation to increase confidence in the reliability of the direction and magnitude of observed findings (110). This is not included in the GRACE checklist, but it was raised in the discussions of many studies in this review.

We performed a comprehensive bibliographic database search using various combinations of MeSH/Emtree terms, free text terms, and synonyms to identify the studies included in this review. Nevertheless, it was difficult to find all the studies that used claims data because authors often used the database name rather than a general term for the data set, such as "claims data" or "administrative data." Moreover, MeSH/Emtree terms for claims data were only recently added to the bibliographic databases we used. To reduce the impact of this limitation, we also manually screened the references of all included publications. It should be noted that the large number of studies from the United States could result from the bias of including the MeSH/Emtree terms "Medicare" and "Medicaid" (US insurance programs). We nevertheless believe that this will have had little effect on our findings and recommendations because only minor differences in study quality were identified between those conducted in the United States and in other countries (File 8 in Supplemental Materials, available online).

Finally, few studies specified whether they were exploratory or hypothesis testing in nature, and moreover, there was a conspicuous lack of information to help reviewers consistently distinguish between these 2 types. According to Berger et al (2017), authors should specify this design element before conducting studies of treatment effectiveness (156). This is because, whereas hypothesis testing studies seek to support decision making, exploratory studies seek to generate hypotheses for further research (156). Therefore the requirements and practice recommendations for studies testing hypotheses of treatment effectiveness must be stricter. Providing greater clarity about the study objective in the aims, and specifically about whether a study is exploratory or hypothesis testing in nature, could help readers and decision makers judge whether a given study has sufficient quality for their needs.

CONCLUSION

Many comparative (cost)-effectiveness studies based on claims data have been published in recent decades, and the number of publications has clearly increased over this time. Our review highlights that methodological issues are frequently not addressed or reported appropriately despite the availability of good practice and reporting guidelines. Adherence to these guidelines must improve before the promise of claims data to increase insights into the effectiveness of cancer treatments can be fulfilled.

The supplemental materials can be found online at: doi: 10.1016/j.jval.2020.05.008

PART 2

FROM REAL-WORLD DATA TO EVIDENCE

CHAPTER 4

COST-EFFECTIVENESS ANALYSIS OF MAMMAPRINT®

Published as:
Cost-Effectiveness Analysis of MammaPrint® to Guide the Use of Endocrine Therapy in Patients with Early-Stage Breast Cancer.

Luyendijk M, Jager A, Buijs SM, Siesling S, Uyl-de Groot CA, Blommestein HM.
PharmacoEconomics. 2023;41(8):981-997.

ABSTRACT

Background: Gene expression profiling tests can predict the risk of disease recurrence and select patients who are expected to benefit from therapy, while allowing other patients to forgo therapy. For breast cancers, these tests were initially designed to tailor chemotherapy decisions, but recent evidence suggests that they may also guide the use of endocrine therapy. This study evaluated the cost-effectiveness of a prognostic test, the MammaPrint®, to guide the use of adjuvant endocrine therapy, in patients eligible according to Dutch treatment guidelines.

Methods: We constructed a Markov decision model to calculate the life-time costs (in 2020 Euro's) and effects (survival and quality adjusted life years) of MammaPrint® testing versus usual care (endocrine therapy for all patients) in a simulated cohort of patients. The population of interests includes patients for whom MammaPrint® testing is currently not indicated, but for whom it may be possible to safely omit endocrine therapy. We applied both a health care perspective and a societal perspective and discounted costs (4%) and effects (1.5%). Model inputs were obtained from published research (including randomized controlled trials), nation-wide cancer registry data, cohort data and publicly available data sources. Scenario and sensitivity analyses were conducted to explore the impact of uncertainty around input parameters. Additionally, threshold analyses were performed to identify under which circumstances MammaPrint® testing would be cost-effective.

Results: Adjuvant endocrine therapy guided by the MammaPrint®, resulted in fewer side effects, more (quality adjusted) life years (0.10 and 0.07 incremental QALYS and LYs, respectively) and higher costs (€18,323 incremental costs) compared to the usual care strategy in which all patients receive endocrine therapy. While costs for hospital visits, medication costs and productivity costs were somewhat higher in the usual care strategy, these did not outweigh costs of testing in the MammaPrint® strategy. The incremental cost-effectiveness ratio was €185,644 per QALY gained from a healthcare perspective and €180,617 from a societal perspective. Sensitivity and scenario analyses showed that the conclusions remained the same under changed input parameters and assumptions. Our results show the MammaPrint® can become a cost-effective strategy when either the price of the test is reduced (>50%), or the proportion of patients for which treatment is altered (i.e. those with ultra-low risk) increases to >26%.

Conclusion: Standard MammaPrint® testing to guide the use of endocrine therapy in our simulated patient population appears not to be a cost-effective strategy compared to usual care. The cost-effectiveness of the test can be improved by reducing the price or preselecting a population more likely to benefit from the test.

INTRODUCTION

A paradigm shift is taking place in oncology towards de-escalation of treatment with the aim of improving and personalizing care (157). Treatment de-escalation includes reducing ineffective care or care that provides patients with no net benefit (i.e. treatments for which the benefits do not counterbalance the harmful effects). Evidence-based de-escalation strategies offer advantages to patients because they can safely forgo therapy without compromising outcome (24). Additionally, avoiding ineffective treatments may also reduce health care costs (158). Risk stratification by gene expression profiling (GEP) is an approach to personalize and de-escalate treatment. Patients who are expected to benefit from therapy are distinguished from patients that can forgo therapy.

Breast cancer is the most common cancer in women and the most frequent cause of cancer-related death among women worldwide. Currently, several GEP tests are available for breast cancer and some of these are recommended for use in clinical practice, including MammaPrint®, Oncotype DX®, EndoPredict®, Prosigna® and Breast Cancer Index (24, 159). Retrospective and prospective studies have shown that these tests can accurately identify patients who have a low risk of disease recurrence and who can safely forgo adjuvant chemotherapy (24). Despite the fact that GEP tests are quite expensive, they are considered to be cost effective to identify patients who can forgo adjuvant chemotherapy in many countries. The costs of testing can be offset by gains in health-related quality of life (HRQoL) combined with savings in costs for chemotherapy and related adverse events (159, 160).

Recent studies suggest that MammaPrint® is also suitable for another purpose, that is, to guide endocrine therapy (ET) decisions in patients with early-stage breast cancer. This concerns patients who already have an excellent survival rate without chemotherapy. Nevertheless, (inter)national guidelines currently recommend (including for these patients) 5 years of adjuvant endocrine therapy (ET) to reduce the risk of disease recurrence (161, 162). Although ET is typically inexpensive, it can cause several side effects. While only a small proportion of patients develop severe side effects such as endometrial cancer and thromboembolism, less severe adverse events, such as hot flashes, arthralgia, vaginal dryness, emotional lability and symptoms of depression, are seen frequently (163, 164, 165). The latter are typically not life-threatening, but they often impact patients' quality of life and social functioning. Additionally, side effects are also associated with increased costs, for instance due to more visits to health care professionals and reduced work productivity (166).

In the Netherlands, MammaPrint® is currently commercially available but not reimbursed from the basic benefit package (167, 168). Possibly, the new indication of MammaPrint®

does qualify for reimbursement. For this purpose, information about the cost effectiveness is valuable but to the best of our knowledge currently lacking. As such, we conducted a cost-effectiveness analysis of MammaPrint® to guide ET decisions in patients in the Netherlands.

METHODS

Study Design

We constructed a decision analytic model to estimate the incremental costs per (quality-adjusted) life year (QALY/LY) of MammaPrint® testing to assign ET in patients with ER+/HER2-, lymph node negative and either grade 1 with a tumor size between 2-3 centimeter, or grade 2 and tumor size of 1-2 centimeter. We assumed 100% test accuracy for the MammaPrint®. The model simulates the course of events of 1,000 patients aged 63 years (i.e., average age of patients with the above mentioned characteristics in the Netherlands (96)) for two strategies: 1) MammaPrint® testing to guide ET or 2) not testing and give ET according to current guidelines to all patients. The latter includes, 2.5 years tamoxifen followed by 2.5 years aromatase inhibitors (AI) or 5 years AI for postmenopausal patients and five years tamoxifen combined with ovarian suppression for premenopausal patients (169, 170). Our analysis followed the Dutch pharmacoeconomic guidelines. As such, we used a societal perspective and included both direct and indirect medical and non-medical consumption costs and productivity costs (171). We also reported results from the health care perspective and for both perspectives we reported results without indirect costs. Costs are valued at 2020 euros and costs and health outcomes are discounted at a rate of 4% of and 1.5%, respectively (171).

Model Description

A decision tree was combined with a semi-Markov model to simulate a cohort of patients with early breast cancer (eBC) with the above mentioned characteristics. The decision tree included the two alternative strategies: 1) test patients with the MammaPrint® and guide their adjuvant ET accordingly; 2) treat all patients with adjuvant ET (Figure 1), and was used to calculate the proportion of patients with ultra-low risk. In each of the of the branches of the tree, the population was distributed among the clinical risk of recurrence scores as defined by the MammaPrint®: ultra-low (13%, n=131), low or high (50% low + 37% high = 87%, n=869) risk (162). Note: we assumed the same proportion of ultra-low risk patients in the usual care group.

The semi-Markov model simulated the patients with an ultra-low risk over different health states, until death, using transition probabilities and a cycle length of 3 months. We included the MammaPrint® tests costs for all patients that needed to be tested

(N=1,000). Treatment costs and outcomes were not included for the patients with a low and high risk (N=869) because these are not altered with the test. The health states in the Markov model included: i) 'alive' which are patients diagnosed with eBC (and the characteristics mentioned above), ii) 'endometrial cancer', 'other malignancies', 'thromboembolic event' which are three health states representing major adverse events (AEs) related to ET and; iii) 'death'. We did not include a breast cancer recurrence health state in our model because we assumed that the probability of recurrence would be the same in ultra-low risk patients in the MammaPrint® and usual care strategy (161, 162). All ultra-low risk patients, regardless of whether they received ET or not, initially entered the 'alive' state. From this health state they could die of any cause or develop a major AE and potentially die. Patients could only be in one health state at a time and experience one major AE in their lifetime. Patients who were cured from a major AE entered a new tunnel health state and were assumed to have the same survival as patients in the alive state.

Costs and health-related quality of life values (Qol) (also known as utilities) were attributed to the health states of the model. The total costs and effects of the two strategies were calculated by summing up the health state costs and effects of all cycles (for n=131 ultra-low risk patients). Test costs of the entire cohort (n=1000) were added for patients in the MammaPrint® strategy. Average costs and effects were calculated for the ultra-low risk patients (thus dividing the total costs and effects by 131). The incremental cost-effectiveness ratio (ICER) was calculated as follows: (average costs MammaPrint® strategy ' average costs usual care strategy) / (average effects MammaPrint® strategy ' average effects usual care strategy). To determine whether the MammaPrint® strategy was cost-effective compared to usual care, we compared the ICERs to the different willingness-to-pay (WTP) thresholds valid in the Netherlands (i.e. 20,000, 50,000 or 80,000 per QALY gained) (172).

Input Parameters

Model inputs used to estimate costs and effects were identified using targeted literature searches and obtained from published research (including randomized controlled trials (RCTs)), nation-wide cancer registry data, cohort data and publicly available data sources. Studies were selected based on their relevance to our model and the level of evidence according to the evidence based medicines criteria (173). For costs and utility parameters we preferred Dutch studies but if these were not available we used studies from other developed countries. A distribution around the input parameter was defined. All input parameters together with their uncertainty distributions are presented in Table 1a-1f.

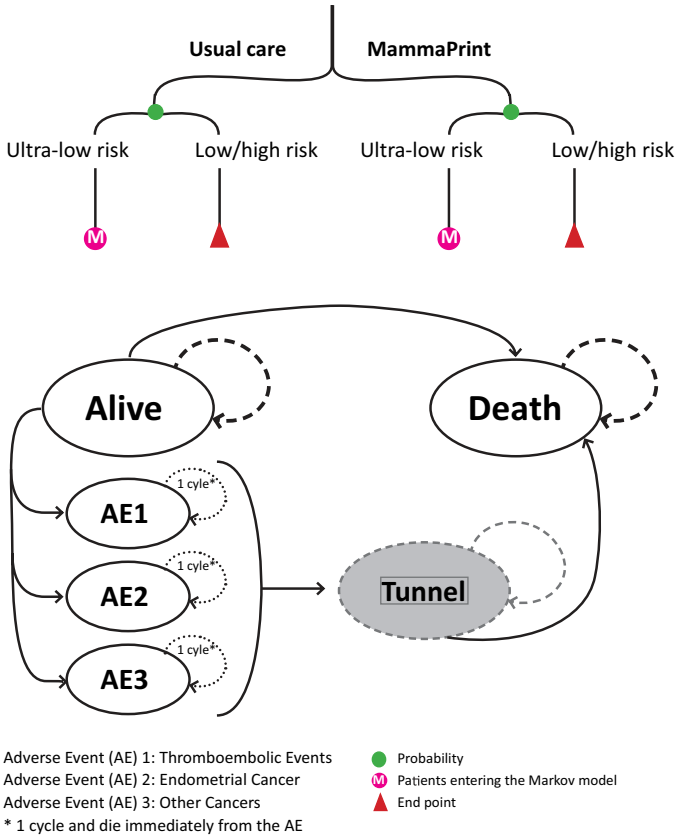

Figure 1. Schematic overview of the model

Probabilities

Mortality from any cause was assumed to be equal to that of the general female population aged 63 years old (174). We assumed no difference in all-cause mortality between the ultra-low risk group treated with or without ET. This assumption was supported by the excellent metastatic free and overall survival of ultra-low risk patients, both with and without treatment (161, 162). Dutch life tables were used to calculate death probabilities per cycle (174).

Probabilities of major adverse events (per cycle) were calculated for four groups: pre and postmenopausal women not treated with ET, postmenopausal women treated with an AI for 5 years, postmenopausal women treated with 2.5 years of tamoxifen followed by 2.5 years of AI, and premenopausal women treated with tamoxifen for 5 years combined with ovarian suppression. Probabilities were obtained from the literature (163, 175).

Supplemental file 1 (available online) provides more information about the RCTs and exact calculations.

Probabilities of death from major adverse events were calculated based on different publications. For the thromboembolic events, we weighted the proportion of patients with deep vein thrombosis (DVT) and pulmonary embolism (PE) (164, 176). For endometrium cancer and other cancers, we assumed that patients would only die from metastatic disease. The occurrence of metastatic disease was estimated for 21% of patients who would develop endometrium cancer and for 27% of patients who would develop other malignancies (Supplement file 2, available online). Probabilities of death were applied to all patients who entered the major AE health state for the first time. By using this approach, we assumed that patients would either die from the AE within 3 months of diagnosis or would survive the AE and die from any other cause. We also assumed that patients who experienced a major adverse event would cease treatment and thus not experience a reduction in quality of life and costs due to the further minor and major AEs beyond three months.

Minor adverse events related to ET were clustered in four overarching groups based on the publication of Kadakia et al. (2015) (177) which was an appropriate categorization according to a clinical oncologist: 1) vasomotor symptom which include hot flashes and night sweating; 2) vulvovaginal symptoms which includes symptoms such as vaginal dryness and bleeding; 3) mood symptoms which mainly includes depressive symptoms; 4) cognitive symptoms including tiredness and forgetfulness etc; 5) musculoskeletal symptoms for which we assumed that mainly fractures would result in costs and reductions in quality of life (177). Probabilities of these minor AEs for patients treated with and without ET were recorded for the same groups as the major adverse events, using the same data sources. We calculated the probability of having any minor AE by taking the highest proportion of the AE groups. We assumed that minor AEs would occur directly from the start of treatment and would remain for the full five years of treatment (except for fractures) (Supplemental file 3, available online).

It is well known that many breast cancer patients discontinue ET early and therefore assuming full adherence would overestimate costs. Since the proportions of adherence were not reported in detail in the RCTs, the discontinuation rates were obtained from a Dutch publication (178). This percentage was used to reduce ET medication costs.

Table 1a. Information about the characteristics of the patient population simulated in this study

Patient population					
Parameter	**Value**	**SE**	**Alpha**	**Beta**	**Source**
Mean age (years)	63.3	n.v.	n.a.	n.a.	Netherlands Cancer Registry (national data) (96)
Premenopausal (%)	18	0.003	5,257	23,351	Netherlands Cancer Registry (proportions based on national data) (96)
Postmenopausal (%)	82		23,351	5,257	
Adjuvant endocrine treatment	74	0.012	380	1,071	(178)
- 5 years of AI (%)	26		1,071	380	
- 2.5 years of tamoxifen followed by 2.5 years AI (%)					
Employed (%)	53	0.001	300,289	270,602	Statistics Netherlands (174)

Abbreviations: AI: aromatase inhibitor; n.v.: not varied; SE: standard error

Table 1b. Transition probabilities used to populate the decision tree and Markov model

Decision tree – probabilities					
	Value	**SE**	**Alpha**	**Beta**	**Source**
Ultra-low risk of recurrence based the MammaPrint® (%)	13	0.012	98	652	(162)
Low or high risk of recurrence based the MammaPrint® (%)	87		652	98	
Markov model - probabilities					
Probability of death					
Women with ultralow risk of recurrence based on the MammaPrint®	Based on life tables[a] (Supplemental file 10, available online)				(175, 179)
Probability to adhere to ET					
	Value	**SE**			**Source**
Year 1 (%)	87	n.v.	n.v.	n.v.	(178)
Year 2 (%)	78				
Year 3 (%)	69				
Year 4 (%)	63				
Year 5 (%)	49				

[a]Annual probabilities are reported by Statistics Netherlands, we calculated probabilities for each cycle. These probabilities are not varied in the PSA. Abbreviations: ET: endocrine therapy; n.v.: not varied; SE: standard error; PSA: probabilistic sensitivity analyses.

Table 1c. Probabilities of major adverse events in pre- and postmenopausal patients treated with and without ET and probabilities of death from major adverse events

Probabilities of major adverse events – probabilities of thromboembolic events are applied for 5 years and of endometrium cancer and other cancers for 10 years				
	Per cycle (SE – Alpha/Beta)			**Sources**
Post-menopausal women	AI (175)	Tamoxifen followed by AI (175)	No ET - RR (163)	
Thromboembolic events (%)	0.0484 *(0.0003 – 2.4/4,850)*	0.1038 *(0.0005 – 5/4,809)*	0.031 *(0.001 – 7/2241/0.001 – 11/2229[a])*	(163, 175)
Endometrium cancer (%)	0.0057 *(0.0004 – 0.2/3,075)*	0.0189 *(0.0004 – 0.6/,3044)*	0.0091 *(0.001 – 8/2,240 / 0.00 – 5/2,235[a])*	
Other cancers (%)	0.1933 *(0.0008 -5.9/3,069)*	0.1732 *(0.0008 – 5.3/3,040)*	0.1419[b] *(0.002 – 28/2,220 / 0.003 – 38/2,205[a])*	
	Per cycle *(SE – Alpha/Beta)*			**Sources**
Pre-menopausal women	Tamoxifen + ovarian suppression - (164, 180)	*No ET* – (164)		
Thromboembolic events (%)	0.063 *(0.0005 – 1/2,325)*	0.039 *(0.0007 – 0.5/1,005)*		(164, 180)
Endometrium cancer (%)	0.015 *(0.0004 – 1.7/3,577)*	0.010 *(0.0004 – 1.7/3,573)*		
Other cancers (%)	0.122 *(0.0004 – 1.7/3,577)*	0.109 *(0.0004 – 1.7/3,573)*		
	Per cycle *(SE)*			**Sources**
Death from:				
Thromboembolic events (%)	7.7 *(0.006 – 91/1,111[b])*			(164, 176)
Endometrium cancer (%)	21 *(0.021 – 78/296[b])*			(181)
Other cancers (%)	27 *(0.027 – 72/197[b])*			Assumptions

[a] Probabilistic RR were calculated by dividing two probabilities and using the SE of both. Probabilistic RRs were applied to the probabilistic; [b] SE assumed to be 10% of mean values of exemestane. Abbreviations: AI: aromatase inhibitor; ET: endocrine therapy; RR; relative risk; SE: standard error.

Table 1d. Probabilities of minor adverse events in pre- and postmenopausal patients treated with and without ET

Probabilities of minor adverse events in different subgroups - applied for 5 years (during ET)				
	Per cycle – for 5 years (SE – Alpha/Beta)			**Sources**
Post-menopausal women	AI – (175)[a]	Tamoxifen followed by AI – (175)[a]	No ET - RR from - (163)[b]	
Vasomotor (%)	35.1 *(0.007 – 1,703/3,149)*	40.4 *(0.007 – 1,945/2,869)*	27.9 *(0.010 718/1530/0.010 – 900/1340[d])*	(163)
Vulvovaginal (%)	6.6 *(0.004 – 320/4,532)*	8.4 *(0.004 - 403/4411)*	6.4 *(0.008 343/1905/0.008 352/1888[d])*	
Mood (%)	13.5 *(0.005 – 654/,4198)*	10.5 *(0.004 – 504/4,310)*	13.4 *(0.006 - 235/2013/0.006 – 236/2004[d])*	
Cognitive (%)	10.5[c] *(0.006 – 236/2,004)*	10.5[c] *(0.006 – 236/2,004)*	10.5 *(0.009 – 465/1783/0.009 – 532/1717[d])*	
Fractures (%)	0.26 *(0.001 – 13/4,839)*	0.17 *(0.001 – 8/4,844)*	0.25 *(0.005 – 143/2105/0.005 – 149/2091[d])*	
Any minor AE[f] (%)	35.1 *(n.v.)*	40.4 *(n.v.)*	27.9 *(n.v.)*	
	Per cycle – for 5 years *(SE – Alpha/Beta)*			**Sources**
Pre-menopausal women	Tamoxifen + ovarian suppression – (180)	*No ET – RR from (163, 164, 182)*[e]		
Vasomotor (%)	93.5 *(0.005 – 2,175/151)*	63.9 *(0.012 – 988/838/0.011 – 1233/577)*		(163, 180, 182, 183)
Vulvovaginal (%)	49.2 *(0.010 – 1,144/1,182)*	36.2 *(0.011 – 565/1261/0.011 – 669/1141)*		
Mood (%)	51.4 *(0.010 – 1,195/1,131)*	47.0 *(0.006 - 235/2013 /0.006 - 236/2004)*		
Cognitive (%)	59.5 *(0.010 – 465/1,783)*	41.5 *(0.009- 465/1783/0.009- 523/1717)*		
Fractures (%)	0.19 *(0.001 – 2/2321)*	0.17 *(0.004- 235/3340/0.004 – 240/3339)*		
Any minor AE[f] (%)	93.5 *(n.v.)*	63.9 *(n.v.)*		

[a] Crude probabilities from TEAM trial (took the highest % reported if multiple similar AEs were reported); [b] Applied RR to probabilities of exemestane. Probabilistic RR were calculated by dividing two probabilities and using the SE of both; [c]obtained from MAP3. because not reported in TEAM; [d] Probabilistic RR were calculated by dividing two probabilities and using the SE of both; [e] MAP3. Was also used because not all events were reported for pre and post-menopausal women separately in the IBIS1; [f] Max of the 5 overarching groups using the SE but using the max of the probabilistic values of all categories. Abbreviations: ET: endocrine therapy; IBIS1: International Breast Cancer Intervention Study; MAP.3: Mammary Prevention.3; n.v.: not varied using the SE but using the max of the probabilistic values of all categories; RR; relative risk; SE: standard error; SOFT: Suppression of Ovarian Function Trial; TEAM: Tamoxifen Exemestane Adjuvant Multinational.

Table 1e. Utility values applied in the model

Utility values					
	Per cycle	**SE**	**Alpha**	**Beta**	**Source**
First years after primary BC diagnosis	0.696	0.007	2719	1188	(184)
Subsequent years (= utility of the general population of women aged 55-65 years)	0.89	0.089	10	1	(185)
AE 1: Thromboembolic events - fatal	-0.056	0.006[a]	93	1,581	(171) (applied for 1 cycle)
AE 1: Thromboembolic events – chronic	-0.004	0.0002[a]	99	46,979	(171, 186) (see Supplemental file 4)
AE 2: Endometrial cancer	-0.036	0.0036[a]	95	2,581	(187) (applied for 1 cycle)
AE 3: Other cancers	-0.036	0.0036[a]	95	2,581	Assumed to be the same as endometrium cancer. (applied for 1 cycle)
Dis-utilities of minor adverse events values					
Dis-utilities of adverse events due to minor AEs – first 3 months	-0.083	0.002	2064	22.816	TOTAM study see Supplemental file 5 (available online), EQ5D5L Dutch Tarif
Dis-utilities of adverse events due to minor AEs – first 6 months – 24 months	-0.074	0.002	1288	16,130	
Dis-utilities of adverse events due to minor AEs – 24 – 60 months	-0.067	0.002	1782	24,830	

Abbreviations: AEs: adverse events; AI: aromatase Inhibitor; BC: breast cancer; EQ5D-5L: EuroQol-5 Dimensions – 5 levels.
[a] SE assumed to be 10% of mean.

Health Effects – Quality of Life

For each health state, survival was weighted by utility values to estimate QALYs. Utility values for the first year in the 'alive' health state were based on the study of Lidgren et al. (2007) who report utility values based on the Euroqol 5-dimensions 3-levels (EQ5D-3L) of breast cancer patients during the first year after diagnoses (184). For the remaining years, utility values were based on the female population from the Netherlands aged 55-65 years obtained with EQ5D-3L and valued with the Dutch tariff (185). The utility value of patients in the death state was assumed to be zero.

Table 1f. Costs input parameters (costs in euro's)

Direct health care costs					
	Once	**SE**	**Alpha**	**Beta**	**Source**
MammaPrint® costs	€2675	fixed	n.a.	n.a.	(192)
Endocrine therapy	**Per cycle**	**SE**			
Average of AIs	€ 31.03	3.10[a]	100	0.31	(193)
Tamoxifen (generic)	€ 19.11	1.91[a]	100	0.19	
Ovarian suppression – first cycle 3 injections per cycle	€ 908.91	90.9[a]	100	9.09	
Ovarian suppression – subsequent cycles 1 injections per cycle	€ 302.97	30.3[a]	100	3.03	
Minor adverse event costs					
	Per cycle	**SE**	**Alpha**	**Beta**	**Source**
Vasomotor	€ 22.17	0.22[a]	100	0.22	(196) + assumptions
Vulvovaginal	€ 22.17	0.22[a]	100	0.22	
Mood	€ 22.17	0.22[a]	100	0.22	
Cognitive	€ 22.17	0.22[a]	100	0.22	
	Once	**SE**	**Alpha**	**Beta**	**Source**
Fractures	€ 5,486	€ 1,321[b]	16	343	(197) (once)
Major adverse event health care costs					
	Once	**SE**	**Alpha**	**Beta**	**Source**
Thromboembolic events	€ 4,459	€ 1,130[b]	16	279	(198)
Endometrium cancer	€ 15,292	€ 562	739	21	(199)
Other malignancies	€ 29,897	€ 7,474[b]	16	1,869	(200, 201, 202)
Indirect health care costs					
	Per cycle	**SE**	**Alpha**	**Beta**	**Source**
Indirect medical costs	Based on PAID 3.0	n.v.	n.a.	n.a.	(195)
End of life costs		n.v.	n.a.	n.a.	
Patient and family costs					
Travel costs	**Per cycle**	**SE**	**Alpha**	**Beta**	**Source**
Patients with minor AEs	€ 4.64	€ 0.46[a]	100	0.05	(171)
Informal care costs					
Patients who died from major AEs	€ 2,680	€ 670[b]	16	168	(203)
Costs made in other sectors					
	Per cycle	**SE**	**Alpha**	**Beta**	**Source**

Table 1f. Costs input parameters (costs in euro's) (continued)

Direct health care costs					
	Once	**SE**	**Alpha**	**Beta**	**Source**
Costs due to productivity losses for patients treated with ET with related AEs – first 3 months	€ 3,094	€ 656	22	139	TOTAM study (see Supplemental file 7, available online) (188, 189)
Costs due to productivity losses for patients treated with ET related AEs – next 3 months	€ 907	€ 348	7	133	
Costs due to productivity losses for patients treated with ET related AEs – subsequent months	€ 846	€ 342	6	138	
Costs due to productivity losses for patients without ET related AEs – first 3 months	€ 320	€ 320	1	320	
Costs due to productivity losses for patients without ET related AEs – 3 - 6 months	€ 346	€ 163	5	77	
Costs due to productivity losses for patients without ET related AEs – subsequent months	0	€ 0.5	1	250	
Cost due to productivity losses for patients who died from major adverse events	€ 3,501	€ 875	16	219	(costs of 1 friction costs period = 102 days) (171)
Non-medical consumption costs in life years gained					
Non-medical consumption costs for each age ≥63	Based on PAID 3.0	n.v.	n.a.	n.a.	(195)

[a] SE assumed to be 10% of mean; [b] SE assumed to be 25% of mean. Abbreviations: AEs: adverse events; Ais: aromatase inhibitors anastrozole, letrozole, exemestane; ET: endocrine therapy; n.v.: not varied; PAID: Practical Application to Include Disease Costs.

Dis-utilities for thromboembolic events (chronic and fatal) and endometrium cancer were obtained from the literature (Supplemental file 4, available online) (187). We assumed dis-utilities for other cancers similar to disutility for endometrium cancer. The dis-utilities for acute thromboembolic events, endometrium cancer and other cancers were applied for one cycle whereas the chronic thromboembolic dis-utilities were applied for a lifelong duration.

Dis-utilities were also applied for patients who experienced any minor AE. These dis-utilities were calculated based on patient level data of the Therapeutic Drug Monitoring guided tamoxifen dosing (TOTAM) study (Dutch Trial Registry; NL6918) (Supplemental file 5, available online) (188, 189). This study evaluated therapeutic drug monitoring guided dosing of adjuvant tamoxifen and collected data on adverse events, quality of life (EQ5D-5L) and productivity losses (institute for Medical technology Assessment (iMTA)

Productivity Cost Questionnaire (PCQ)) 3, 6 and 24 months post initiation of tamoxifen (190). The EQ5D-5 levels (5L) data were valued using the Dutch tariff (191).

Costs

The costs of the MammaPrint® (€2675) were based on the commercial tariff reported in previous publications (192). Costs of ET and ovarian suppression (gonadotropin-releasing hormone agonist agonists) were based on national tariffs (193). Costs of ET related minor AEs included those of visits to health care professionals (including procedures for fractures) (Supplemental file 6, available online), travel costs and costs of productivity losses. Travel costs were calculated by multiplying the average distance to the health care facility with the average travel costs per kilometer (Dutch costing manual) (171). Productivity losses related to experiencing minor AEs were calculated based on patient level data of patients in the TOTAM study (Supplemental file 7 (available online) for more information) (188, 189). Costs of productivity losses were applied for 15 cycles because the average age of women in our model was 63.3 years old and the retirement age in the Netherlands is 67 years in 2024 (194).

Health care costs of ET related major AEs were based on published literature (Supplemental file 8, available online). These costs were applied to all patients entering the major AE health state. Productivity costs (of one friction period) of major AEs and costs of informal care (of 59 hours per month) were applied only to patients dying from the events (171).

End of life costs, unrelated health care costs and non-medical consumption costs in life years gained were obtained from the Practical Application to Include Disease Costs (PAID 3.0) (195).

Accounting for Uncertainty

We conducted univariate sensitivity analyses to describe the impact of uncertainty in the input parameters in our model. Input parameters were varied with +/-30% and we reported the most influential parameters (i.e., deviations from the base case ICER of > €10,000). In addition to the univariate sensitivity analyses we also performed a probabilistic sensitivity analysis using Monte Carlo simulations (1,000 iterations). Beta distributions were assigned to probabilities and utility parameters, and gamma distributions to cost parameters.

Scenario and Threshold Analyses

In scenarios we explored the impact of different assumptions for the incidence of 'other cancers', the share of ultra-low risk patients identified, the amount of productivity costs,

the utility values of minor adverse events and the share of patients who would receive the MammaPrint®. In addition, threshold analyses were performed to evaluate at what price and incidence rate (i.e., proportion of ultra-low risk patients identified with the MammaPrint®) testing would be cost-effective (Supplemental file 9, available online).

Model Validation

Our cost-effectiveness model was validated using the Assessment of the validation status of Health-Economic decision models tool (AdViSHE) (204). The conceptual model of our study followed the same structure as that of many previously conducted health economic models to evaluate interventions for early stage breast cancer. The model was constructed according to the 'Principles of good practice for decision analytic modeling in health-care evaluation: Report of the ISPOR task force on good research practices—Modeling studies' in close collaboration between health economic and clinical experts. Extreme value testing was performed for several parameters and validity checks were built into the model (i.e. constant number of patients in each cycle of the model). The model results were extensively discussed in the team. Finally, several scenario analyses were performed to assess the robustness of the results and back-of-the-envelope calculations were performed to assess whether the results were as expected (see Supplemental file 11, available online).

The analyses were performed in Microsoft excel and the list of important assumptions can be found in Supplemental file 12, available online.

RESULTS

The results of the base case analysis are presented in Table 2-4.

Effectiveness

The MammaPrint® strategy yielded more QALYs and LYs and fewer AEs than the usual care strategy (Table 2). The model estimated an average of 16.73 QALYs and 19.20 LYs per patient in the MammaPrint® strategy and 16.63 QALYs and 19.14 LYs in the usual care strategy (discounted results). Moreover, in the usual care strategy, more patients experienced a major AE. Specifically, in the MammaPrint® strategy 0.82 patients were estimated to experience a thromboembolic event (within 5 years from the start of treatment), 0.45 patients would develop endometrial cancer (within 10 years) and 6.64 patients another cancer (within 10 years). In the usual care strategy thromboembolic events, endometrial cancer and other cancers were estimated to occur in 1.59, 0.49 and 8.49 patients, respectively.

Table 2. Model results: Average discounted life years, quality adjusted life years per patient and total number of patients with major adverse events

	QALYs first 5 yrs	QALYs all yrs	LY first 5 yrs	LY all years	Patients with thromboembolic events	Patients with endometrial cancer	Patients with other cancers
Intervention					First 5 yrs	First 10 yrs	First 10 yrs
MammaPrint®	4.05	16.73	4.96	19.20	0.82	0.45	6.64
Usual care	4.00	16.63	4.95	19.14	1.59	0.49	8.49
Increments	0.05	0.10	0.01	0.07	-0.77	-0.04	-1.86

Abbreviations: LY: life years; QALYs: quality adjusted life years; yrs: years.

Costs

The average costs (over the 131 patients with ultra-low risk) from the health care perspective were €22,366 per patient for the MammaPrint® and €4,029 for the usual care strategy, respectively. The most important costs driver were the costs of the MammaPrint®, which were €20,472 per patient in the MammaPrint strategy and zero in the usual care strategy. The average costs from the health care perspective including costs in life years gained were €185,770 and €166,874 for the MammaPrint® and usual care, respectively. In addition to the costs of the MammaPrint®, important costs drivers were the indirect medical costs (total MammaPrint®: €138,590, usual care: €138,038), followed by the end-of-life costs (total MammaPrint®: €24,814, usual care: €24,807).

From the societal perspective (excluding costs in life years gained), the average costs were €165,062 vs €147,298 per patient in the MammaPrint® and usual care strategy, respectively. When including the costs in life years gained the costs were €328,466 (MammaPrint®) and €310,143 (usual care). Major costs drivers were the costs of the MammaPrint®, drug acquisition costs and the costs of productivity losses (MammaPrint®: €2,645 usual care: €3,641). When we included the costs in life years gained, these also contributed significantly (Table 3).

Incremental Cost-Effectiveness Ratio

Depending on the perspective and inclusion of costs in LY gained; ICERs ranged from €175,107 to €185,644 per QALY gained (Table 4). Incremental differences in QALYs were mainly driven by differences in survival time due to the occurrence of major AEs and differences in costs by the price of the MammaPrint®.

Table 3. Model results: average discounted costs per patient and increments (costs in euro's)

		Direct medical costs				Indirect medical costs		Costs in other sectors		Patient & family costs		
Perspective	**Intervention**	**Test costs**	**Drug acquisition costs[a]**	**Minor adverse events (hc)**	**Major adverse events (hc)**	**End of life costs**	**Indirect Medical costs**	**Productivity costs**	**Indirect consumption costs**	**Informal care costs**	**Travel costs**	**Total costs**
Health care – only direct hcc	MammaPrint®	20,472	n.a.	569	1,326	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	22,366
	Usual care	n.a.	1,683	640	1,706	n.a.	n.a.	N.a.	n.a.	n.a.	n.a.	4,029
	Increments	20,472	-1,683	-72	-380	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	18,337
Health care	MammaPrint®	20,472	n.a.	569	1326	24,814	138,590	n.a.	n.a.	n.a.	n.a.	185,770
	Usual care	n.a.	1,683	640	1706	24,807	138,038	n.a.	n.a.	n.a.	n.a.	166,874
	Increments	20,472	-1,683	-72	-380	7	552	n.a.	n.a.	n.a.	n.a.	18,896
Societal – excl costs in life years gained	MammaPrint®	20,472	n.a.	569	1326	n.a.	n.a.	2,645	139,989	33	29	165,062
	Usual care	n.a.	1,683	640	1706	n.a.	n.a.	3,641	139,545	43	40	147,298
	Increments	20,472	-1,683	-72	-380	n.a.	n.a.	-997	444	-10	-10	17,764
Societal	MammaPrint®	20,472	n.a.	569	1,326	24,814	138,590	2645	139,989	33	29	328,466
	Usual care	n.a.	1,683	640	1,706	24,807	138,038	3641	139,545	43	40	310,143
	Increments	20,472	-1,683	-72	-380	7	552	-997	444	-10	-10	18,323

[a] includes costs of bone density measures ones every 5 year for patients treated with ET. Abbreviations: excl: excluding; hc: health care; hcc: health care costs n.a.: not applicable

Table 4. Discounted incremental cost-effectiveness ratios (costs in euro's)

	Perspective			
	Health care – only direct hcc	**Health care**	**Societal – excl costs in LY gained**	**Societal**
Incremental costs (EUR) per LY gained	276,928	285,370	268,281	276,723
Incremental costs (EUR) per QALY gained	180,152	185,644	175,107	180,617

Abbreviations: Hcc: health care costs; EUR: euros; LY: life years: QALY: quality adjusted life years

Univariate Sensitivity Analysis

The 13 most influential parameters in our model are shown in Figure 2. Changing the probability of getting another cancer (post-menopausal women) with +30%, resulted in the highest ICER of €316,905 per QALY gained. Other influential parameters were the probability of patients with ultra-low risk as compared to the population tested with MammaPrint®, the probabilities of having any adverse events (in different subgroups) and the probability of other cancers for other subgroups.

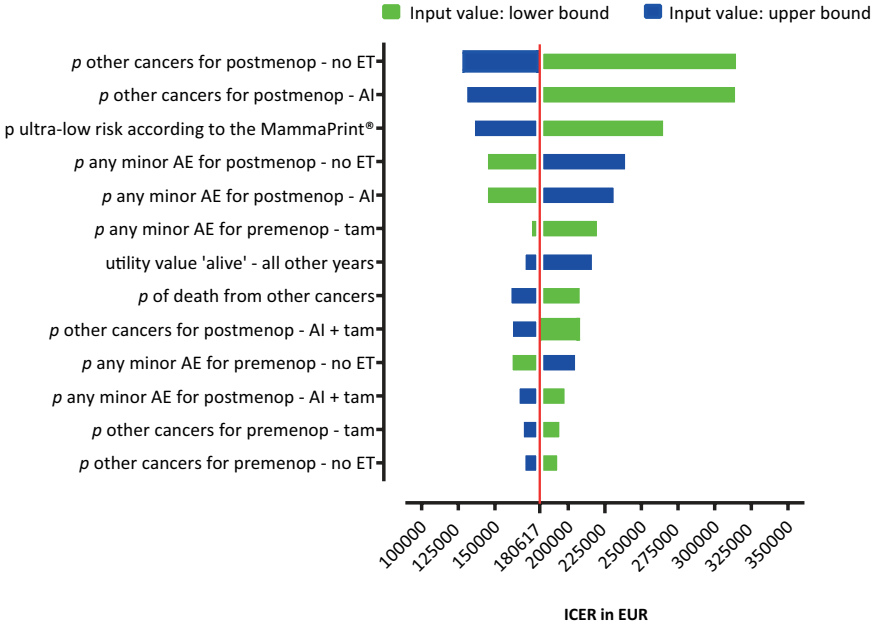

Figure 2. Tornado diagram (societal perspective, discounted)

Abbreviations: AE: adverse events; AI: aromatase inhibitor; ET: endocrine therapy; EUR: euro's; ICER: incremental cost-effectiveness ratio; postmenop: post-menopausal women; premenop: pre-menopausal women; p: probability; tam: tamoxifen.

Probabilistic Sensitivity Analysis

The results of the probabilistic sensitivity analyses are shown in the cost-effectiveness plane (Figure 3) for the societal perspective (discounted). None of the iteration fell below the WTP threshold of €80,000 per QALY, which is the highest WTP threshold used in the Netherlands.

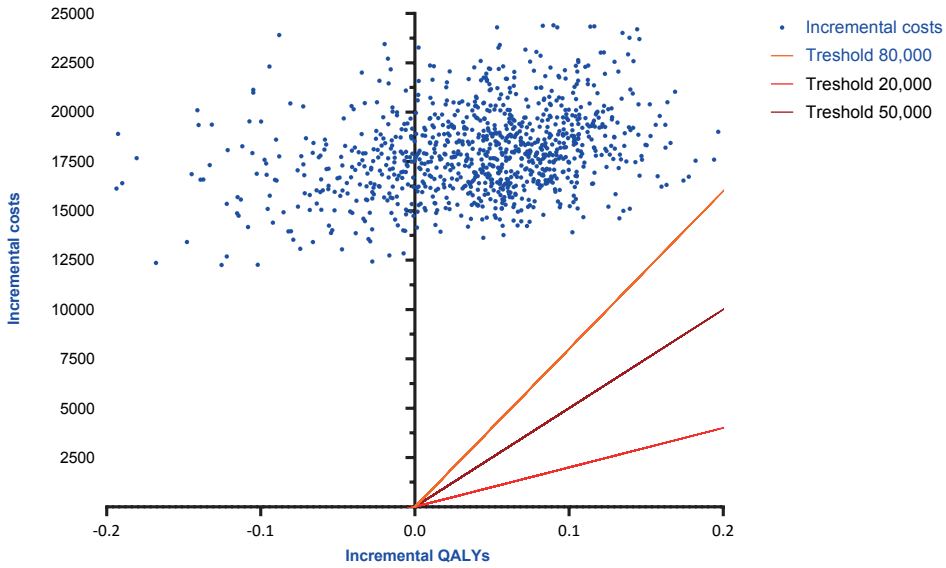

Figure 3. Incremental cost-effectiveness plane (QALYs)

Scenario and Threshold Analyses

The results of the scenario analyses are reported in Table 5. Mainly scenario 1, in which we assumed that patients treated with and without ET had the same chance of having another cancer in the coming 10 years instead of having higher chance in the ET group (base case), had a major impact on the results. Changing this assumption resulted in significantly less QALYs and LYs gained as compared to the base case (1 LY & 5 QALYs vs 10 LY and 13 QALYs). Because the costs and incremental costs barely changed, the ICER significantly increased.

The threshold analyses revealed that price of the MammaPrint® should be reduced with 50%, 65% or 80% are needed for the MammaPrint® strategy to be cost-effective depending on the WTP threshold. Another way for the MammaPrint® to become cost-effective, would be if the proportion of patients for which treatment could be altered (i.e., those with ultra-low risk) increased to at least 26%, 37% or 64% instead of the 13% in the base case (20,000, 50,000, 80,000, respectively).

Table 5. Scenario and threshold analyses (costs in euro's)

	Average life years gained			Average QALYs gained			Average costs (societal perspective)			ICER	
Scenario	**MammaPrint®**	**Usual care**	**Increment**	**MammaPrint®**	**Usual care**	**Increment**	**MammaPrint®**	**Usual care**	**Increment**	**Per LY**	**Per QALY**
Base case	19.20	19.14	0.07	16.73	16.63	0.10	328,466	310,143	18,323	276,723	180,617
Same probability of 'other cancers' for patients in MammaPrint® and usual care group	19.41	19.40	0.01	16.91	16.86	0.05	330,273	312,485	17,788	2,043,152	346,434
Smaller difference in productivity costs between MammaPrint® and usual care group	19.20	19.14	0.07	16.73	16.63	0.10	332,026	313,399	18,627	281,306	183.609
Reduced disutility values of minor adverse events	19.20	19.14	0.07	16.79	16.71	0.08	328,466	310,143	18,323	276,723	227.989
Costs of the MammaPrint® are attributed to 60% of patients	19.20	19.14	0.07	16.73	16.63	0.10	320,277	310,143	10,135	153,054	99,899
Costs of the MammaPrint only applied to those who experience AEs	19.20	19.14	0.07	16.61	16.47	0.14	331,388	313,985	17,403	262,818	126,074
	WTP threshold	Max price and reductions required to be cost-effective									
Varying the price											
Threshold analysis 1	80,000	Max price 1,341 euro, price reductions of 50%									

Table 5. Scenario and threshold analyses (costs in euro's) (continued)

	Average life years gained			Average QALYs gained			Average costs (societal perspective)			ICER	
Scenario	**MammaPrint®**	**Usual care**	**Increment**	**MammaPrint®**	**Usual care**	**Increment**	**MammaPrint®**	**Usual care**	**Increment**	**Per LY**	**Per QALY**
Threshold analysis 2	50,000	Max price 944 euro, price reductions of 65%									
Threshold analysis 3	20,000	Max price 546 euro, price reductions of 80%									
Varying the proportion of patients with ultralow risk identified with the MammaPrint® - proportion (n)											
Base case		13% (7.65)									
Threshold analysis 1	80,000	26% (3.83)									
Threshold analysis 2	50,000	37% (2.70)									
Threshold analysis 3	20,000	64% (1.56)									

Abbreviations: LY: life years; max: Maximum; persp: perspective; QALYs: quality adjusted life years; WTP: willingness to pay.

DISCUSSION

Our study showed that MammaPrint® guided adjuvant ET yields more QALYs and LYs but substantially higher costs in Dutch early breast cancer patients who are eligible for ET only. The costs were mostly driven by the relatively high number of patients needed to be tested to identify one patient who can safely forgo ET, and the price of the MammmaPrint®. Depending on the perspective, the ICERs ranged from € 174,450 to € 185,644. With the willingness to pay thresholds used in the Netherlands, the MammaPrint® strategy will not be a cost-effective strategy for guiding adjuvant ET for the population in our study. The proportion of patients for which treatment would be altered needs to be at least 26%, or the price of the MammaPrint® needs to be reduced by >50% to make the MammaPrint® a cost-effective strategy for the use in our study.

To our knowledge, this study is the first to evaluate the cost-effectiveness of the MammaPrint® to guide the use of adjuvant ET only. Based on the cost-effectiveness results, we would not recommend to reimburse the MammaPrint® from the basic benefit package to guide ET treatment decisions. The value for money of the test can possibly be improved by pre-selecting patients eligible for the test. In our comparison, many patients need to be tested with the MammaPrint®, to identify one patient that can be classified as being at ultra-low risk of recurrence and can safely forgo ET (i.e., 13%) (161, 162). This means that the total costs of testing, i.e., population size times the price of the test, are relatively high compared to the share of patients who benefit. Clearly, if less patients needed to be tested, the ICER will improve. Previous studies have suggested that the frequency of ultra-low risk breast cancers is higher in screen-detected cancers and in patients with more favorable characteristics (161). Additionally, one of our scenario analyses showed that the ICER improves if the test is only used in patients who develop minor symptoms. Further research identifying which patients benefit most from the test may be relevant to optimize the use of the MammaPrint®.

Despite the results of our study, we believe that de-escalating ET is relevant for patients with early breast cancer. Not only did the patients in the MammaPrint® strategy experience fewer side effects, but they also had a slightly better quality of life and survival rate. Moreover, health care costs and productivity losses were lower. Since adjuvant ET is frequently used, de-escalation strategies have the potential to reduce the treatment burden in many patients and also save costs to society (205). Unfortunately, de-escalation studies in this area are scarce, possibly because the consequences of ET related side effects on patients and society are understudied and underestimated and treatment costs are relatively low (206). In fact, we did not find a single European study which evaluated the effect of ET and adverse events on productivity losses and few studies which as-

sessed the impact on quality of life and health-care costs (196, 207). This observation is especially notable given the substantial amount of literature studying ET. Ultimately, to truly improve patient outcomes and the economics of health care, researchers should include broader outcomes (e.g., quality of life and costs) in their studies.

Quantifying the impact of relatively mild adverse events, such as those related to ET, is complex. As in most cost-effectiveness analyses, we expressed the benefits of omitting ET in QALYs, based on utility measures derived from a generic quality of life instrument (i.e., the EQ5D) (208). Generic instruments cover universal health aspects (e.g., self-care, mobility, pain), which makes them relevant for patients with all types of diseases. Disadvantage is that these instruments are usually not very responsive to specific health problems, such as the menopausal symptoms in patients treated with ET. As such, we may not have fully captured the beneficial effects of omitting ET. Possibly, the value based health care (VBHC) framework, would have been more sensitive to capture effects related to reductions in minor AEs because this framework is focusses on outcomes that are most relevant to patients (209). Nevertheless, even in this framework, finding the appropriate measure to value the outcomes would be a challenge.

There are a couple of limitations of our study. First, we assumed 100% accuracy of the MammaPrint® to stratify patients into the different risk groups and it is unlikely that this assumption holds. However, even with this optimistic assumption, the ICER is already far above the WTP threshold used in the Netherlands. A lower test accuracy would reduce the effects and increase the costs of the MammaPrint® strategy hence increase the ICER. Second, given the currently available evidence, we could not assess the ICER in different age groups. Further research in potential subgroups is recommended. For example, the share of patients with ultra-low versus low/high risk has a major impact on the ICER and could therefore lead to different conclusions regarding cost-effectiveness. Third, transition probabilities of minor AEs related to ET were based on different RCTs which used various definitions and described different levels of detail. For instance, one trial reported vasomotor symptoms without further specifying these, whereas another reported hot flashes, sweating and fatigue separately, even by grade (163, 165). Fourth, major AEs probabilities were based on studies that were not powered to report differences in these outcomes. The reported differences may thus have occurred due to chance. Finally, we assumed that all patients in the usual care strategy were treated with ET (in line with guidelines) and that all pre-menopausal women were treated with ovarian suppression. These assumptions are likely not entirely true. In fact, in daily practice only about 70% of the patients with the characteristics of those in our study actually initiate ET (196). Moreover, probably not all young women are willing to be treated with ovarian suppression. This suggests that ICERs would be even less favorable. Despite these limitations, we

think that our study convincingly showed that the use of the MammaPrint as described in this study, is effective but not yet a cost-effective strategy.

CONCLUSION

This study suggests that MammaPrint® testing to guide adjuvant ET in patients only eligible for ET, is an effective but not a cost-effective strategy compared to usual care. De-escalating ET appears to offer gains in survival time and quality of life and results in lower direct medical costs and productivity losses. Likely, the cost-effectiveness of the test can be improved if it would be possible to pre-select patients who can best benefit from the test.

The supplemental materials can be found online at: doi: 10.1007/s40273-023-01277-4

CHAPTER 5

REGULATORY APPROVAL, REIMBURSEMENT, AND CLINICAL USE OF CDK 4/6 INHIBITORS

Published as:
Regulatory Approval, Reimbursement, and Clinical Use of Cyclin-Dependent Kinase 4/6 Inhibitors in Metastatic Breast Cancer in the Netherlands.

Luyendijk M, Blommestein HM, de Groot CA, Siesling S, Jager A.
JAMA Network Open. 2023;6(2):e2256170.

ABSTRACT

Importance: The number of new cancer medicines that are being approved by regulatory agents is increasing exponentially. Yet little is known about the pace at which these medicines reach eligible patients in daily clinical practice during different phases of the postapproval access pathway.

Objective: To describe the entire postapproval access pathway of cyclin-dependent kinase 4/6 (CDK4/6) inhibitors in the Netherlands, from regulatory approval to reimbursement and to investigate the adoption of these medicines in clinical practice among patients with metastatic breast cancer.

Design, Setting, and Participants: This cohort study reviewed approval and reimbursement decisions of the CDK4/6 inhibitors palbociclib, ribociclib, and abemaciclib and estimated the number of patients with metastatic breast cancer who were eligible for these medicines compared with the actual use in clinical practice. The study used nationwide claims data that were obtained from the Dutch Hospital Data. Claims and early access data for patients with hormone receptor–positive and ERBB2 (formerly HER2)–negative metastatic breast cancer who were treated with CDK4/6 inhibitors from November 1, 2016, to December 31, 2021, were included.

Main Outcomes and Measures: Description of the postapproval access pathway, monthly number of patients who were treated with CDK4/6 inhibitors in clinical practice, and estimated number of patients who were eligible for treatment. Aggregated claims data were used, and patient characteristics and outcomes data were not collected.

Results: Three CDK4/6 inhibitors have received European Union–wide regulatory approval for the treatment of HR-positive and ERBB2-negative metastatic breast cancer since November 2016. In the Netherlands, the number of patients who have been treated with these medicines increased to approximately 1847 (based on 1 624 665 claims over the entire study period) from approval to the end of 2021. Reimbursement for these medicines was granted between 9 and 11 months after approval. While awaiting reimbursement decisions, 492 patients received palbociclib, the first approved medicine of this class, via an expanded access program. By the end of the study period, 1616 patients (87%) were treated with palbociclib, whereas 157 patients (7%) received ribociclib, and 74 patients (4%) received abemaciclib. The CKD4/6 inhibitor was combined with an aromatase inhibitor in 708 patients (38%) and with fulvestrant in 1139 patients (62%). The pattern of use over time appeared to be somewhat lower compared with the estimated

number of eligible patients (1847 vs 1915 in December 2021), especially in the first 2.5 years after approval.

Conclusion and relevance: This study found that CDK4/6 inhibitors rapidly reached many eligible patients with metastatic breast cancer and were adopted gradually over time in the Netherlands. Adoption of innovative medicines may be further optimized, and better transparency of the availability of new medicines during different phases of the postapproval access pathway is needed.

INTRODUCTION

Over the past 2 decades, a range of new targeted medicines and immune therapies have been developed for different types of cancer (63). These pharmaceutical innovations have the potential to improve the outcomes of patients with cancer (210). However, this improvement can be realized only when all patients who may benefit have access to these new medicines at the right time (211). Within oncology, a gap has been observed between the rapid development of new medicines and their actual adoption in clinical practice (212, 213).

After innovative medicines are approved by regulatory authorities, there are several procedures that need to be completed before the drugs become available to patients. For instance, in most high-income countries, new expensive cancer medicines are subject to health technology assessment (HTA) procedures to guarantee that scarce resources are spent on treatments that offer value for money (214). Moreover, after reimbursement has been agreed on, treatment guidelines should be adapted, supply should be guaranteed, and clinicians need to prescribe the new medicines to eligible patients. These postapproval procedures take time and are often blamed for delayed or restricted patient access to innovative medicines (60).

Even among the wealthiest countries in the world, achieving timely and sustained access to cancer drugs has been proven difficult (213, 215, 216). Monitoring access is thus desirable to identify possible delays in use and disparities in uptake and to assist with improvement. Previous studies on the access of patients with cancer to new medicines have covered different phases of the access pathway. For instance, some studies evaluated the degree to which new medicines were available, reimbursed, or subsidized in countries of varying economic status (212, 217, 218, 219). Other studies evaluated the time from approval to first sales and described sales over time (213, 220). These studies revealed delays in use and variations in access and uptake, but to our knowledge, no study has yet assessed the entire access pathway, from regulatory approval to reimbursement to adoption among eligible patients.

In this context, the cyclin-dependent kinase 4/6 (CDK4/6) inhibitors palbociclib, ribociclib, and abemaciclib, which are indicated for the treatment of patients with metastatic breast cancer, are interesting to study. These medicines were approved starting in 2016, and thus sufficient follow-up data are available for studying adoption rates over time. Moreover, these medicines offer renewed options for patients with hormone receptor (HR)–positive and ERBB2 (formerly HER2)–negative metastatic breast cancer. With improvements of approximately 10 months in progression-free survival and up to 14

months in overall survival reported in randomized clinical trials (RCTs), these 3 CDK4/6 inhibitors are seen as important breakthroughs in the treatment of metastatic breast cancer (221, 222).

In this cohort study, we evaluated the entire postapproval access pathway of CDK4/6 inhibitors from regulatory approval to reimbursement in the Netherlands, a high-income country in Europe, and investigated the adoption of these medicines in clinical practice. To identify the extent to which these medicines have reached eligible patients, we compared use with estimated use among eligible patients with metastatic breast cancer.

METHOD

Study Setting

In the Netherlands, different postapproval procedures need to be completed before newly approved, solid oncological medicines reach patients in need (eAppendix 1 and eFigure 1 in Supplement 1, available online). First, a clinical committee formulates a recommendation regarding the use of the new medicine in clinical practice (clinical assessment). This advice can be interpreted as clinical guideline recommendations (223). Second, expensive medicines require an HTA. Depending on the results of these analyses, the Ministry of Health, Welfare and Sport in the Netherlands decides whether the medicines can be reimbursed (67, 224). Third, if reimbursement is granted, the pharmacy should guarantee a sufficient supply and the government needs to ensure that declaration codes are available so that hospitals can bill insurance companies (logistic and administrative procedures). Ultimately, adoption in clinical practice should take place. New medicines may be available via alternative routes during the postapproval processes. This study was deemed exempt from review by Dutch legislation, which does not require informed consent from patients or approval by a medical ethics committee for this type of study, which did not contain directly identifiable patient data. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Study Approach and Data Sources

This retrospective cohort study comprised 3 parts: (1) a review of postapproval decisions for the CDK4/6 inhibitors, (2) data analyses of the number of patients who were treated in clinical practice, and (3) an estimation of the number of eligible patients. Aggregated claims data were used, but patient characteristics and outcomes data were not collected.

Two data sources, claims data and early access data, were consulted to identify the number of patients with HR-positive and ERBB2-negative metastatic breast cancer who were treated in clinical practice. First, the manufacturer of palbociclib provided data on the number of new patients (incident cases) who were treated each month during the wait for reimbursement. During this period of approximately 9 months, the manufacturer initiated an Expanded Access Program (EAP) for patients with HR-positive and ERBB2-negative metastatic breast cancer. The number of patients who received treatment (prevalent cases) per month were estimated (eAppendix 2 in Supplement 1, available online). Second, we obtained monthly (May 2016-December 2021) claims data on the CDK4/6 inhibitors by indication from the nationwide Dutch Hospital Data, which covers almost all (approximately 98%) hospitals in the Netherlands. Indications in the data set were coded according to the European Medicines Agency (EMA) label, and the number of prescriptions were specified per dose. We included only the prescriptions for patients with HR-positive and ERBB2-negative metastatic breast cancer. These data were used to calculate the monthly number of prevalent cases treated with each of the 3 CDK4/6 inhibitors after reimbursement was granted. We divided the number of prescriptions per month with the expected number of pills per patient per 4 weeks (i.e., 21 for palbociclib, 63 for ribociclib, and 56 for abemaciclib).

Statistical Analysis

The monthly number of patients who were eligible based on the indications from the EMA label (Table 1) was estimated using a 2-step approach. First, we estimated the number of incident cases eligible for first- and second-line CDK4/6 inhibitor combined with endocrine therapy (ET) each month. We used incidence statistics of patients with breast cancer from the Netherlands Cancer Registry as a starting point. We then calculated the number of patients with HR-positive and ERBB2-negative metastatic breast cancer per treatment line by applying proportions that were obtained from the literature (96, 225, 226, 227, 228, 229) (ie, mainly a Dutch cohort study (227); Table 2) (eAppendix 3 and eFigure 2 in Supplement 1, available online). In the Netherlands, the recommendation is to prescribe CDK4/6 inhibitors in the second line of treatment (230). Thus, the only patients who were eligible for CDK4/6 inhibitors in the first line of treatment were those who developed metastases within 12 months after the end of adjuvant treatment with an aromatase inhibitor (230). In addition to the EMA label indications, we assumed that a proportion of patients may not be eligible to receive CDK4/6 inhibitors. These were patients who already underwent 2 lines of ET (mainly older adults); who died while receiving first-line ET; who had an aggressive course of the disease, and thus cannot undergo ET; and who did not want any anticancer treatment. We assumed that patients who did not meet the criteria for use of CDK4/6 inhibitors in the first or second line of treatment at the time of reimbursement would not receive them in a later line.

Second, we estimated the number of prevalent cases per month by summing up the incident cases of each new month and subtracting the number of patients who experienced progression. Information regarding the probability of progression was obtained from Kaplan-Meier curves from RCTs that evaluated the efficacy of palbociclib (231, 232). We assumed that the efficacy of the other CDK4/6 inhibitors (ribociclib and abemaciclib) was similar. Because the reported follow-up in the RCTs was too short, we extrapolated the survival curves beyond the observed survival time using the approach of Hoyle and Henley (233, 234) (eAppendix 3 and eFigures 3-6 in Supplement 1, available online). Finally, we calculated adoption rates by dividing the number of patients who were treated with the estimated number of patients who were eligible.

Table 1. EMA-Label Indications for Approved CDK4/6 Inhibitors in the Netherlands

Medicine	Therapeutic indication according to the EMA product information
Palbociclib	Palbociclib is indicated for the treatment of HR-positive and ERBB2-negative locally advanced or MBC in combination with an AI; OR in combination with fulvestrant in women who have received prior ET.
Ribociclib	Ribociclib in combination with letrozole is indicated as initial endocrine-based therapy for postmenopausal patients with HR+ and ERBB2-negative advanced or MBC
Ribociclib (extension of indication)	Ribociclib is indicated for patients with HR-positive and ERBB-negative locally advanced or MBC in combination with an AI or fulvestrant as initial endocrine-based therapy or for patients who received prior ET. The population was extended to premenopausal patients and patients who received prior ET in combination with fulvestrant.
Abemaciclib	Abemaciclib is indicated for the treatment of women with HR-positive and ERBB2-negative locally advanced or MBC in combination with an AI or fulvestrant as initial endocrine-based therapy, or in patients who received prior ET.

Abbreviations: AI: aromatase inhibitor; CDK4/6: cyclin-dependent kinases 4 and 6; HR+/HER2-: hormone receptor positive, human epidermal growth factor receptor 2 negative; EMA: European Medicines Association; ET: endocrine therapy.

Table 2. Estimated New Patient Eligibility for First- and Second-line CDK4/6 Inhibitors in the Netherlands

Description	N	Explanation	Sources
Diagnosis			
BC	15,003	Total No. of new pts with invasive BC in 2017	Integraal Kankercentrum Nederland (96)
MBC at primary diagnosis	750	5% of new patients had MBC	Van de Meer et al. 2021 (225)
MBC after primary BC	2,851	20% Of patients developed metastatic recurrence	Lobbezoo et al. 2015 (226)
HR-positive and ERBB-negative MBC	2,377	66% Of patients had HR-positive and ERBB2-negative MBC	Lobbezoo et al. 2016 (227)
Total HR-positive and ERBB-negative MBC	2,377	Total No. of new patients with HR-positive and ERBB2-neagtive MBC	NA

Table 2. Estimated New Patient Eligibility for First- and Second-line CDK4/6 Inhibitors in the Netherlands (continued)

Description	N	Explanation	Sources
Eligibility for first-line ET with CDK4/6 inhibitors[a]:			
Developed metastases shortly after the end of or during adjuvant treatment with an AI	217	70% Of patients with HR-positive and ERBB2-ngative MBC treated withf first line ET	Lobbezoo et al. 2016 (227)
		13% Of patients with metachronous disease developed metastases within 24 mo after the diagnosis of primary BC: these patients were assumed to have received an AI shortly before developing metastases and therefore were not eligible for ET combined with CDK4/6 as second-line of ET	Lobbezoo et al. 2016 (227); assumptions
Initially treated with chemotherapy	238	22% Of patients with HR-positive and ERBB2-negative MBC received chemotherapy as initial therapy	Lobbezoo et al. 2016 (227)
		56% Of patients were eligible for first-line ET with a CDK4/6 inhibitor[b]	Lobbezoo et al. 2016 (227), Fietz et al. 2017 (228); assumptions
		80% Of patients eligible for first-line ET after chemotherapy were assumed to receive ET combined with palbociclib as first-line ET.	Assumption
Total eligible per y	455	Total No. of pateints eligible for first line ET combined with CDK4/6 inhibitor per year	NA
Total eligible per mo	38	Total No. of pateints eligible for first line ET combined with CDK4/6 inhibitor per month	NA
Eligibility for second-line ET with CDK4/6 inhibitors[a]:			
Initially treated with ET	760	70% Of patients HR-positve and ERBB2-negative MBC treated with first-line ET	Lobbezoo et al. 2016 (227)
		87% Of patients were not eligible for first line CKD4/6 inhibitors[c]	Lobbezoo et al. 2016 (227)
		58% Of patients were eligible for another line of ET[d]	Rugo et al. 2019 (229)
		90% Of patients were willing or able to use CDK4/6 inhibitors	Assumption
Initially treated with ET followed by a line of chemotherapy	121	70% Of patients with HR-positive and ERBB2-negative MBC were treated with first-line ET 87% Of patients were not eligible for first line CDK4/6 inhibitors[c]	Lobbezoo et al. 2016 (227)
		34% Of patients received chemotherapy after a line of ET	Assumptions
		35% Of patients were assumed to be elibible for second-line ET	Fietz et al. 2017 (228); assumptions
		70% Of patients were eligible for second-line ET combined with CDK4/6 inhibitors	Assumption

Table 2. Estimated New Patient Eligibility for First- and Second-line CDK4/6 Inhibitors in the Netherlands (continued)

Description	N	Explanation	Sources
Initially treated with chemotherapy followed by ET	15	22% Of patients with HR-positive and ERBB2-negative MBC received chemotherapy as initial therapy 56% Of patients were eligible for first line ET with CDK4/6 inhibitors[b]	Lobbezoo et al. 2016 (227), Fietz et al. 2017 (228); assumptions
		20% Of patients with HR-positive and erBB2-negative MBC received chemotherapy as initial therapy	Assumption
		35% Of patients were assumed to be eligible for second-line ET	Rugo et al. 2019 (229)
		70% Of patients were eligible second-liine ET combined with CDK4/6 inhibitors	Assumption
Total	895	Total No. of patients eligible for first line ET combined with CDK4/6 inhibitor per year	
Total	75	Total No. of patients eligible for first line ET combined with CDK4/6 inhibitor per month	

[a]We distinguish lines of ET from lines of chemotherapy (e.g., patients who had 1 line of chemotherapy are still eligible for first line CDK4/6 inhibitor combined with ET because it is their first line of ET).
[b]This percentage was calculated based on the assumption that 26% would receive second-line therapy (227) and 18% would not receive any therapy (228).
[c]Calculated as 100% minus 13% (of patients with metachronous disease who were assumed to have received an AI shortly before developing metastases and therefore were not eligible for ET combined with CDK4/6 inhibitors as second line).
[d]This percentage was based on the PALOMA-2 (A study of Palbociclib + Letrozole vs Letrozole for 1st line treatment of Postmenopausal Women with ER+/HER2- Advaced Breast Cancer) trial of Palbociclib combined with letrozole in 58% of patients treated with placebo (ET only) received ET as subsequent treatment.
[e]Calculated as 100% minus 32% of patietns who received another line of chemotherapy minus 33% of patients who died.
[f]Calculated as 100% minus 80% of patients who were eligible for first-line ET after chemotherapy and assumed to receive ET combined with Palbociclib as first line.
Abbreviations: AI: Aromatase Inhibitors; BC: Breast cancer; CDK4/6: Cyclin-Dependent Kinases 4 and 6; ET: endocrine therapy; MBC: metastatic breast cancer; pts: patients

To represent the implications of uncertainty for the number of patients, we calculated upper and lower bounds of the base case using more conservative or optimistic numbers of incident cases. Data were analyzed in Stata, version 17 (StataCorp LLC); Microsoft Excel, version 2016 (Microsoft Corp); and R, version 4.2.2 (R Foundation for Statistical Computing).

RESULTS

Three CDK4/6 inhibitors have received European Union–wide regulatory approval for the treatment of HR-positive and ERBB2-negative metastatic breast cancer since November 2016. In the Netherlands, the number of patients who have been treated with these medicines increased to approximately 1847 (based on 1 624 665 claims over the entire study period) from approval to the end of 2021. Palbociclib received EMA approval in November 2016; ribociclib received initial EMA approval in August 2017, and its extended indication was approved in December 2018; and abemaciclib received EMA approval in September 2018. After regulatory approval, all 3 medicines were recommended for clinical use, but reimbursement decisions were postponed. After price negotiations, reimbursements were granted for palbociclib in August 2017, ribociclib in May 2018, and abemaciclib in September 2019. The clinical assessment phase and HTA phase lasted 3 and 9 months, respectively, for palbociclib; 1 month (first use) and 8 months for ribociclib; and 3 months and 11 months for abemaciclib (Figure 1).

Early Access Phase

While awaiting clinical practice recommendations and reimbursement, 492 patients with HR-positive and ERBB2-negativemetastatic breast cancer were treated with palbociclib via an alternative access route (i.e., EAP). The number of new patients (incident cases) who participated in this EAP increased the first 3 months (from 24 to 60 patients), remained stable in the 5 months thereafter (from 52 to 65 patients), and decreased in the last month (from 65 to 40 patients) of the early access phase.

Access Phase

In the access phase, the adoption of the first approved and reimbursed CDK4/6 inhibitor, palbociclib, in clinical practice increased over time from 0 patients being treated to approximately 1847 patients over a period of more than 50 months. Use of palbociclib increased steeply after reimbursement was acquired. By the end of the study period, 1616 patients (87%) were treated with palbociclib. Only a small number of patients were treated with the other medicines of this class, ribociclib (n = 157 [7%]) and abemaciclib (n = 74 [4%]), by the end of the study period. A majority of patients (1139 [62%]) received CDK4/6 inhibitors in combination with fulvestrant, and 708 patients (38%) received it with an aromatase inhibitor (Figure 2).

Entire Postapproval Phase

The estimated number of patients eligible (prevalent cases) who were treated with CDK4/6 inhibitors increased during the study period, in the base case analysis, from approximately 110 patients in the month that regulatory approval was granted (November

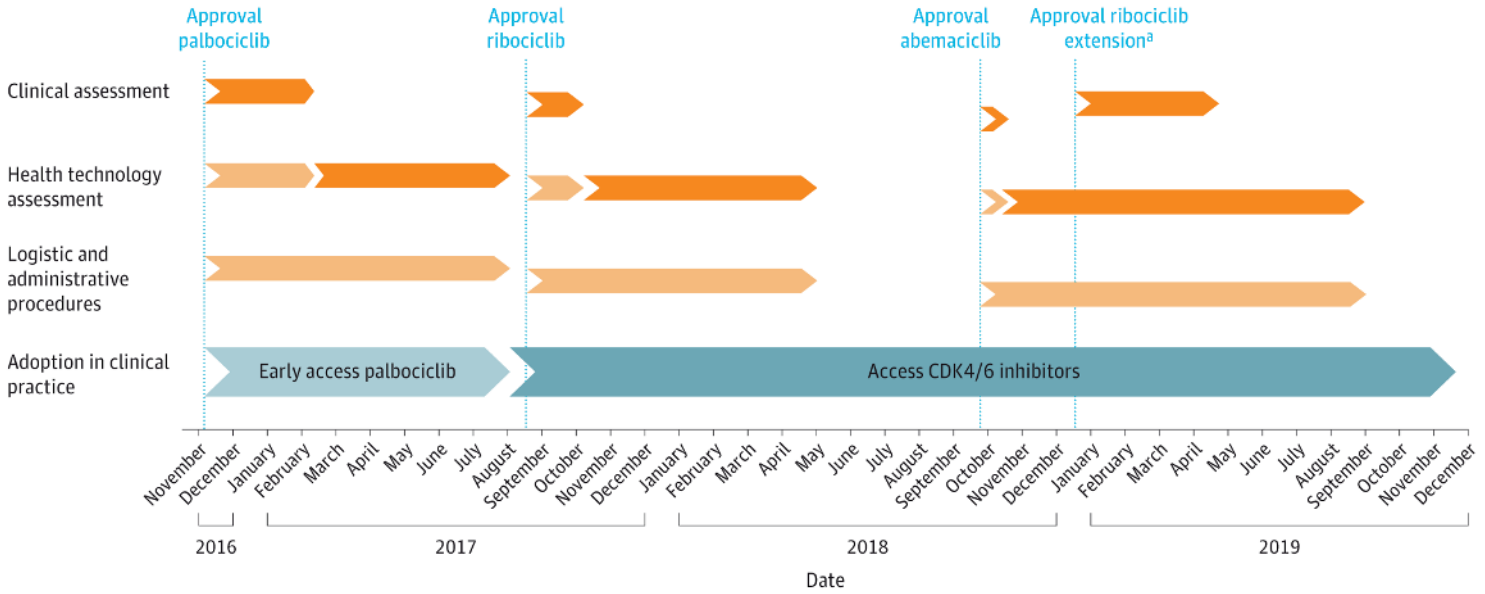

Figure 1. Post approval Phases Needed for Access to Cyclin-Dependent Kinase 4/6 (CDK4/6) Inhibitors

Postapproval phases were possible patient access delay (orange); possible delay overlapping with another phase (light orange); early access possibly not available to all patients (light blue); and access to reimbursed care (blue). Palbociclib combined with aromatase inhibitor in first-line treatment or with fulvestrant in second-line treatment. Ribociclib combined with letrozole in first-line treatment. Abemaciclib combined with aromatase inhibitor or with fulvestrant in first- or second-line treatment. Extended-indication ribociclib combined with fulvestrant in first- or second-line treatment. [a]There was no additional health technology assessment because the extension was already considered in the initial decision.

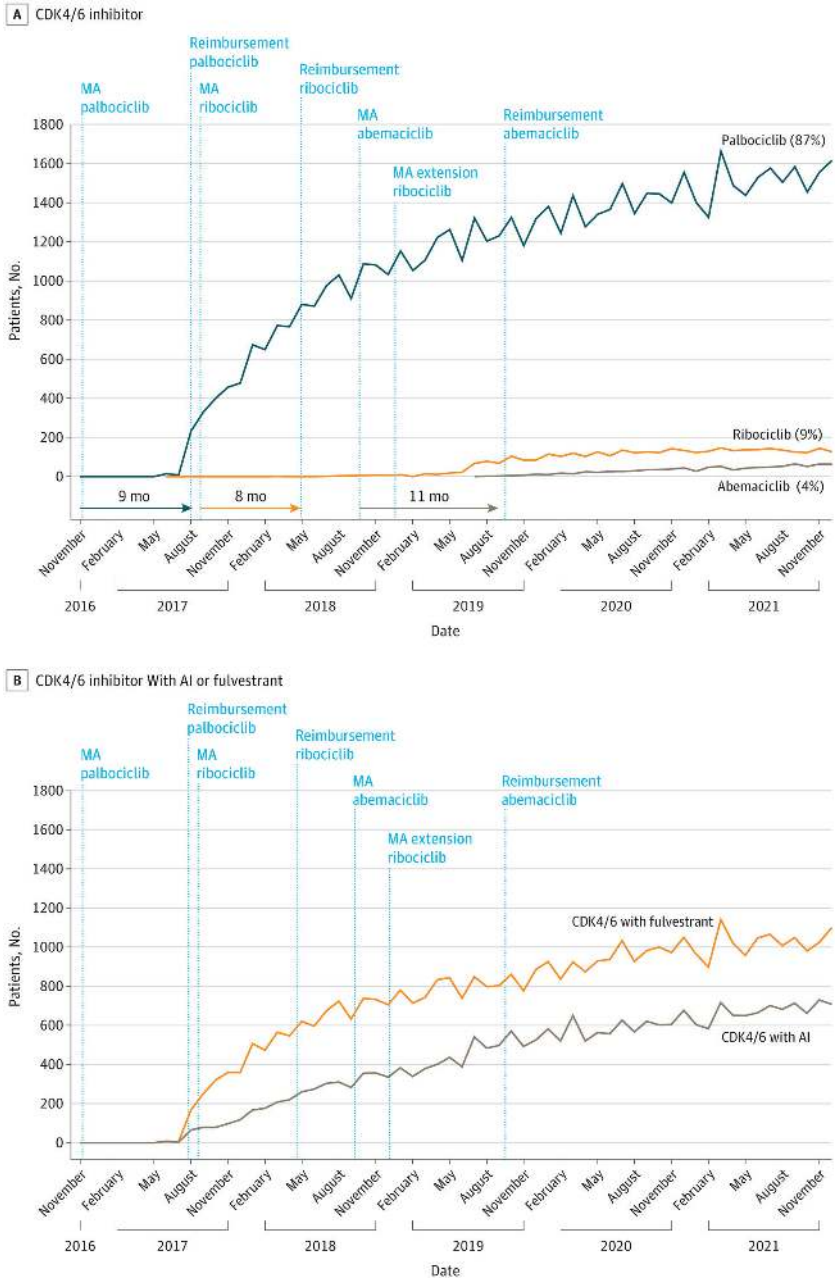

Figure 2. Obseraved Number of Patients Treated With 3 Cyclin-Dependent Kinase 4/6 (CDK4/6) Inhibitors and With a CDK4/6 Inhibitor Combined With Aromatase Inhibitor (AI) or Fulvestrant Over Time

Palbociclib combined with AI in first-line treatment or with fulvestrant in second-line treatment. Ribociclib combined with letrozole in first-line treatment. Abemaciclib combined with AI or with fulvestrant in first- or second-line treatment. Extended-indication ribociclib combined with fulvestrant in first- or second-line treatment; there was no additional health technology assessment because the extension was already considered in the initial decision. MA indicates marketing authorization.

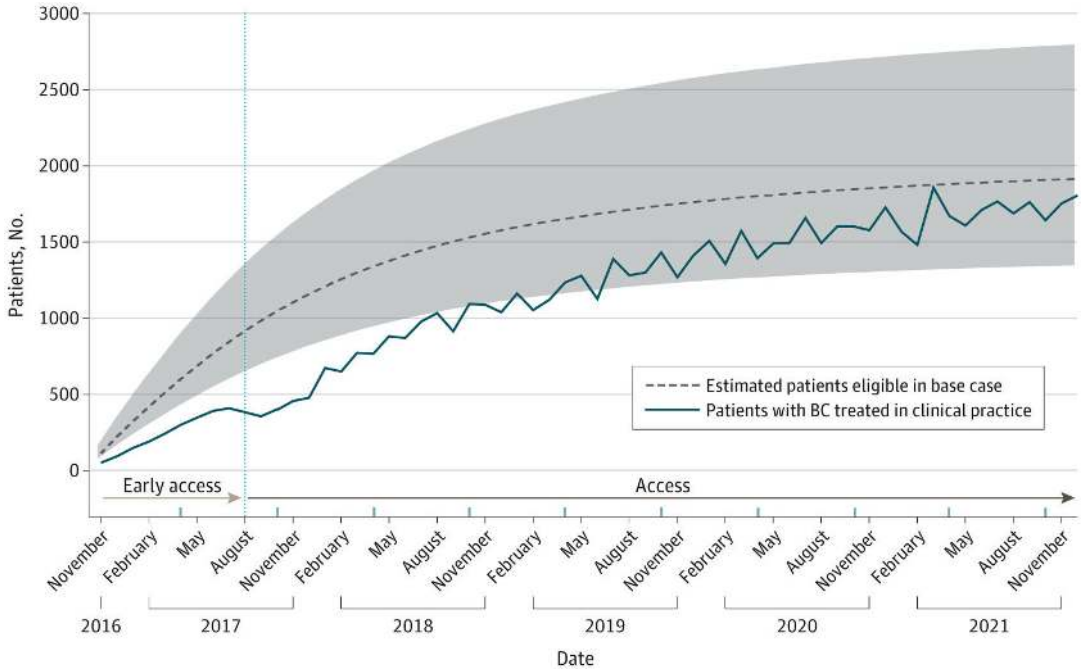

Figure 3. Observed Total Number of Patients Treated and Estimated Number of Prevalent Patients Eligible

The shaded area indicates the upper and lower bounds of uncertainty for the estimated patients eligible in the base case. Early access is access to cyclin-dependent kinase 4/6 inhibitors before approval. BC indicates breast cancer.

2016) to approximately 1915 patients in the final month of the study (December 2021) (Figure 3). The pattern of use over time appeared to be somewhat lower compared with the estimated number of eligible patients, especially in the first 2.5 years after approval. The estimated adoption (ie, number of patients who were treated divided by estimated number of patients who were eligible) of the CDK4/6 inhibitors increased from approximately 383 of 918 (42%) directly after reimbursement was granted (August 2017) to approximately 1847 of 1915 patients (96%) at the end of the study period. We observed a steep increase in the number of patients who received treatment when we considered only the numbers calculated based on the claims data (8 patients before August 2017 to 880 in May 2018) (Figure 2). In contrast, a notable decline in the use (from approximately 410 patients in July 2017 to 357 in September 2017) and adoption (from approximately 42%to 36%) of CDK4/6 inhibitors was observed when we took into account that patients were already treated as part of the EAP in the early access phase (Figure 3).

Sensitivity Analyses

In the base case, of the 198 incident cases, we estimated that 38 (19%) were eligible for first-line and 75 (38%) for second-line CDK4/6 inhibitors each month. Additionally, 85 incident cases (43%) were estimated to be ineligible (Table 2). We created 2 scenarios. First, we assumed that 119 patients (60%) would be ineligible for CDK4/6 inhibitors. Second, we assumed that 34 patients (17%) would be ineligible. The outcomes of these alternative assumptions are shown in Figure 3.

DISCUSSION

This cohort study showed that CDK4/6 inhibitors were gradually adopted in clinical practice in the Netherlands despite reimbursement decisions lagging 9 to 11 months behind the EMA approvals. The medicines reached many Dutch patients with breast cancer within a short period, but the number of patients who were treated in daily practice appeared to be somewhat lower than the estimated number of patients who were eligible in the first 2.5 years after regulatory approval.

While previous studies have reported delays of 9 months, on average, between the approval of cancer drugs and their first use in the Netherlands (2017-2020), the present study found that many patients received palbociclib immediately after regulatory approval was granted (213, 235). Almost 500 patients participated in an EAP, which allowed them to benefit from this medicine even though it was not being reimbursed yet. This observation emphasizes 2 important points. First, it suggests that HTA procedures do not necessarily delay patient access to new medicines. This finding builds on earlier

work that reported that different early access routes, either privately or publicly funded, offered options to patients in high-income countries (236, 237). Second, it suggests that frequently used quality indicators of medicine accessibility (e.g., time to reimbursement) should be interpreted with caution. These indicators are used not only to rank European countries according to their medicine accessibility but also to urge policy makers to improve accessibility (213, 235, 238). Based on the study findings, in-depth studies that evaluate the entire postapproval access pathway are needed to identify and elucidate the disparities in access to innovative medicines.

Despite the rapid use of palbociclib, there are some caveats worth mentioning with regard to its early availability via an EAP. First, there is little transparency about the availability and use of EAP programs in most countries, including the Netherlands (239). Because manufacturers are not allowed to advertise EAPs, not all clinicians may be aware of these programs. Lack of awareness can be a factor in inequalities in access between patients being treated by different physicians. Second, EAPs are not funded by the government in the Netherlands. Their implementation thus relies on the efforts and goodwill of pharmaceutical companies. While some companies may see benefits in EAPs (for instance, to speed up future sales), others may be unwilling or unable to allocate resources for implementing such programs (240, 241, 242, 243). Because HTA procedures take time to complete, this delay likely plays a role in inequalities in access between different groups of patients. Countries surrounding the Netherlands have used different approaches to funding early access. For instance, in Germany, all newly EMA-approved medicines are reimbursed by default at the price set by the company, which ensures immediate access for patients. Pricing negotiations take place after the medicines have been launched (243). In France, the government funds different EAPs to allow patients to use medicines before reimbursement and marketing authorization (236). It remains unclear which approaches best deliver timely, sustainable, and affordable access to innovative medicines for all patients in need.

We evaluated the use of 3 medicines of the same class that were approved for combination with either an aromatase inhibitor or fulvestrant. In this study, we observed that palbociclib, the first approved medicine of the 3 CDK4/6 inhibitors, remained the predominant medicine that was prescribed. This preference for palbociclib was also seen in other high-income countries, although there is no clear evidence that palbociclib is superior to ribociclib and abemaciclib (244, 245, 246, 247, 248). Possibly the pattern of use can be explained by physician preferences, the marketing strategies of the manufacturers, hospital policies, or differences in price rebates. We also observed that the CDK4/6 inhibitors were mainly combined with fulvestrant, which likely reflects the recommendation in the Netherlands to primarily prescribe CDK4/6 inhibitors in the

second line of treatment (230). The Dutch recommendation is contrary to that of other high-income countries (e.g., the US), where they advise to prescribe CDK4/6 inhibitors in the first line of treatment (249). First-line treatment with these drugs has been associated with a longer progression-free survival but also with more toxic effects and higher costs due to longer treatment durations (250). Undoubtedly, the recommendation to use the CDK4/6 inhibitors mainly in the second line of treatment is interesting from a budget-impact point of view (i.e., lower total costs), which may help to ensure sustainable health care financing and prevent access problems in the long run. It remains to be seen what strategy is the most optimal from the effectiveness and cost-effectiveness perspectives. This topic is currently being investigated in the Dutch SONIA (Endocrine Therapy Plus CDK4/6 in First or Second Line for Hormone Receptor Positive Advanced Breast Cancer) trial (250).

It is well known that the adoption of innovations in clinical practice takes time (251, 252). Therefore, it may be unsurprising that we observed a gap between the number of patients who were treated and the estimated number of patients who were eligible in the first 2.5 years after reimbursement was permitted. The disparities found in this study may partly reflect the uncertainty surrounding the estimated patient numbers, which relied on modeling rather than clinical practice data. However, we consider it plausible that the adoption of the 3 CDK4/6 inhibitors started more slowly than expected. Specifically, we observed a minor decline in the increasing use pattern from the moment that the EAP ceased and reimbursement was approved (August 2017). This decrease suggests that not all physicians responded directly to the changed reimbursement status of palbociclib. Raising awareness of this issue among treating physicians may further optimize the implementation of new medicines in the future.

Limitations

This study has some limitations. First, the number of patients who were eligible for the 3 medicines relied on decision analytic modeling, and the scenario analyses we performed showed that the results had substantial amount of uncertainty. Thus, the estimated number of eligible patients and adoption rates should be interpreted with caution. To fully understand access, detailed information is needed regarding patients with metastatic breast cancer, their treatments, and the decision-making process in daily practice. Second, we used aggregated claims data to evaluate the use of the CDK4/6 inhibitors, which means that we did not have data on patient characteristics or outcomes. As such, we could not assess whether patients who received the CDK/6 inhibitors immediately after approval and reimbursement differed from those who accessed the medicines later. Third, we studied the adoption of only 1 medicine class for the treatment of breast cancers in a single country. Future studies need to compare the adoption of medicines

across different countries with different HTA policies. Fourth, the COVID-19 crisis was ongoing during the study period (2020-2021). Nevertheless, we did not expect a large association between the crisis and study results because treatment recommendations for patients with metastatic disease during the pandemic were similar to those from the years before (253).

CONCLUSION

This cohort study found that the CDK4/6 inhibitors palbociclib, ribociclib, and abemaciclib were rapidly accessible to a group of patients with HR-positive and ERBB2-negativemetastatic breast cancer and were adopted gradually over time in the Netherlands, a high-income country. Although many patients received the medicines during the period when reimbursement was pending, there appeared to be barriers to adoption other than reimbursement. To improve accessibility of innovative medicines, we call for better transparency of their availability during the different phases of the postapproval access pathway and for more in-depth cross-country studies on the adoption of medicines, which may further optimize the process.

The supplemental materials can be found online at: doi: 10.1001/jamanetworkopen.2022.56170

CHAPTER 6

CHANGES IN SURVIVAL IN *DE NOVO* METASTATIC CANCER IN AN ERA OF NEW MEDICINES

Published as:
Changes in survival in *de novo* metastatic cancer in an era of new medicines.

Luyendijk M, Visser O, Blommestein HM, de Hingh IHJT, Hoebers FJP, Jager A, Sonke GS, de Vries EGE, Uyl-de Groot CA, Siesling S.
JNCI: Journal of the National Cancer Institute. 2023(0;0)

ABSTRACT

Background: Over the past decades, the therapeutic landscape has markedly changed for patients with metastatic solid cancer, yet few studies have evaluated its effect on population-based survival. The objective of this study was to evaluate the change in survival of patients with *de novo* metastatic solid cancers during the last 30 years.

Methods: For this retrospective study, data from almost 2 million patients diagnosed with a solid cancer between January 1, 1989, and December 31, 2018, were obtained from the Netherlands Cancer Registry, with follow-up until January 31, 2021. We classified patients as with or without *de novo* metastatic disease (M1 or M0, respectively) at diagnosis and determined the proportion with M1 disease over time. Changes in age-standardized net survival were calculated as the difference in the 1- and 5-year survival rates of patients diagnosed in 1989-1993 and 2014-2018.

Results: Different cancers showed divergent trends in the proportion of M1 disease and increases in net survival for M1 disease (approximately 0-50 percentage points at both 1 and 5 years). Patients with gastrointestinal stromal tumors saw the largest increases in 5-year survival, but we also observed substantial 5-year survival increases for patients with neuroendocrine tumors, melanoma, prostate cancer, and breast cancer.

Conclusion: Over 30 years, the survival of patients with *de novo* M1 disease modestly and unevenly increased among cancers. Metastatic cancer still remains a very lethal disease. Next to better treatment options, we call for better preventive measures and early detection to reduce the incidence of metastatic disease.

INTRODUCTION

Solid cancers are a major public health problem worldwide, with approximately 16 million new patients and 8.5 million deaths reported in 2020 (3). Approximately 4%-65% of cancers are diagnosed as metastatic disease, which is often associated with a poor prognosis (254, 255). Although the approval of over 80 novel systemic therapies since 1990 has expanded the treatment options for most metastatic tumors (63), improving survival for these patients remains a challenge (256).

Population-based cancer survival statistics, including 5-year survival rates, are important metrics for evaluating and prioritizing cancer control policy (257). Many prior studies have compared survival rates between countries or have assessed the differences in survival between periods irrespective of cancer stage (257, 258). However, nationwide studies focusing on changes in the survival of patients with distant metastases are scarce because many existing cancer registries have no or incomplete data on stage at diagnosis (259).

Systemic therapy constitutes the backbone of treatment for most metastatic cancers. Historically, this primarily included cytotoxic chemotherapy and hormonal therapy, but in recent decades, the therapeutic landscape has rapidly changed because of the approval of several targeted and immune therapies. Although randomized controlled trials (RCT) have demonstrated the safety and efficacy of each of these new medicines, the impact on survival in unselected population-based samples has been inadequately studied (260). This is an important knowledge gap because clinical trial results may not be representative of the general patient population (261). Moreover, RCTs do not usually include the cumulative benefit of sequential therapies. Analysis of the changes in population-based survival over several decades while considering the systemic therapies that have been introduced might offer valuable insights into the overall impact of new medicines for metastatic cancer.

In this study, we investigate the survival trends of patients with metastatic cancers at the time of diagnosis (*de novo* metastatic cancer [M1]) using data from the nationwide Netherlands Cancer Registry (NCR). We evaluate changes in the survival of patients presenting with a solid cancer between 1989-1993 and 2014-2018, and we discuss this in light of the systemic therapies introduced. Our aim is to determine whether M1 cancer survival has improved during a period in which many novel medicines have been approved.

METHODS

Study design

We performed a retrospective cohort study based on data from the NCR between 1989 and 2018. The primary outcomes of this study were the changes in the 1-year and 5-year net survival of patients with distant M1 disease. We divided the data into 2 cohorts (1989-1993 and 2014-2018) to ensure comparison of periods before and after the implementation of novel medicines. Additionally, we report on the trends in the proportion of patients with M1 disease (from 1989 to 2018).

We included adult patients aged 18 years and older, diagnosed with a solid primary cancer between 1989 and 2018. Information on patients' vital status was available until January 31, 2021. We excluded patients diagnosed on the date of death. Patients with multiple primary tumors were grouped by their first tumor only and were excluded if the first tumor was diagnosed before 1989. This study was approved by the NCR's Ethics Committee (written informed consent was not required).

Data source

The NCR reached national coverage in 1989 and has an estimated completeness of more than 95% of all malignancies in the Netherlands (89). Data managers record data on patient, tumor, and initial treatments directly from patient files. The International Classification of Diseases for Oncology is used for coding topography and morphology. Staging methods are considered when determining the stage at diagnosis, but no information is recorded on the modalities used or the outcomes. In most tumors, the stage is coded according to the Union for International Cancer Control tumor, node, metastasis (TNM) classification (4th-8th edition). Otherwise, tumors are classified by the extent of the disease (EoD) as local, regional, or distant. The NCR is linked each year on January 31 to the Municipal Personal Records Database to obtain information on vital status. Patients who emigrate are censored at the date of emigration.

Case definitions

Patients were classified as having M1 (*de novo* metastatic cancer) based on 1 of 3 criteria: 1) M1 disease according to the TNM classification (clinically and/or pathologically proven); or 2) distant disease according to the EoD classification system (clinical EoD and/or pathologically proven EoD=6 used in the NCR for only a minority of cancers); or 3) the presence of metastasis with an unknown primary tumor (C80.9) (Supplementary Table 1, available online). All other patients were classified as without *de novo* metastatic cancer (M0), including those with unknown stage. In the Netherlands, clinically

diagnosed metastases are typically based on radiological evaluation rather than clinical examination alone.

We defined 28 cancers based on the histology (carcinoma) and the primary site (most cancers) or the histology only (gastrointestinal stromal tumors [GIST], neuroendocrine tumors [NET], melanoma), including a group of "other" cancers and a group of cancers with an unknown primary site (C80.9) (Supplementary Table 2, available online). Approximately 75% of patients with metastatic melanoma had melanoma of unknown primary. These often present with lymph node metastases only, so they do not fit within our definition of M1 disease (distant metastases). Still, we included them in our analyses because they share similar treatment strategies (262).

Lung cancer was divided into small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) subgroups based on their clinical and treatment differences and into "lung other" (mostly cases without pathological confirmation). All patients with central nervous system tumors (i.e., meninges C70, brain C71, other parts of central nervous system C72, pituitary gland C75.1, and pineal gland C75.3) were excluded because these patients usually die from the local tumor burden rather than metastasis. Patients with Kaposi's sarcoma (C46) were excluded because most do not die from metastatic disease but from the underlying disease (mostly HIV/AIDS), and patients with mesothelioma (C45) were excluded because TNM data were unavailable for this cancer between 1989 and 1992.

Statistical analyses

The proportion of patients with M1 disease was calculated by dividing the annual number of patients with M1 disease per cancer by the total annual number of patients with and without distant metastases at diagnosis (M0 + M1). Trends over time of the proportion of M1 were plotted, and by visual inspection of the graphs we grouped the cancer types into a constant, increasing, decreasing, or a combination of an increasing and decreasing trend. We also reported changes over time with the number, and mean age, proportion of men, of patients diagnosed between 1989-1993 and 2014-2018. The 1- and 5-year net survival for patients with M1 disease were estimated for the 2 defined time periods, considering net survival to be the survival observed if cancer was the only possible cause of death (263). Net survival can be interpreted as cancer-specific survival. The Pohar-Perme estimator served to estimate the age-standardized net survival with 95% confidence intervals (264). Age standardization was conducted to account for changes in the distribution of age over time, which is important because the age distribution has a significant impact on the population's survival. The International Cancer Survival Standard was used for most cancer types and Dutch weightings for prostate

cancer (find the weights in Supplementary Table 2, available online). We used the same age distribution for all time periods. Age-standardized net survival analyses were not possible for certain cancers due to the limited number of diagnoses and events in 1 or more age strata. For these cancers, weighting groups were adjusted (e.g., testicular cancer) or the diagnostic cohorts were changed (e.g., vulva or vagina, hepatocellular cancer); additionally, analyses not standardized by age were performed for cancers with 20 or less patients in 1 or more age strata (see Supplementary Table 3, available online).

We used the cohort approach to estimate patient net survival. This approach can estimate the 5-year net survival of a cohort even in the absence of complete follow-up data (such as in our study for the cohort 2014-2018) by using the survival information of all patients until censoring (e.g., like the Kaplan-Meier and life table methods) (263). Survival changes over time were calculated as the arithmetic gain by subtracting the estimated survival rate of the 1989-1993 cohort from that of the 2014-2018 cohort. The statistical significance changes were assessed with the Z-test ($P < .01$) (265).

In supplemental analyses, we also evaluated the changes in median overall survival, and we reported trends in M1 survival over time by estimating the net survival of the intermediate cohorts.

RESULTS

Patient population

In total, almost 2 million adult patients were diagnosed with an eligible solid cancer between 1989 and 2018. We included 52 263 and 84 383 diagnosed with *de novo* metastases (M1) in the survival analyses for 1989-1993 and 2014-2018, respectively. Information on M stage at diagnosis was based on clinical or pathological classification in most patients (69% in 1989-1993 and 92% in 2014-2018). The number with M1 disease increased between 1989-1993 and 2014-2018 for all cancers except for stomach cancer, whereas the number of patients with an unknown primary cancer more than halved (Table 1; Supplementary Table 4, available online).

The proportion of patients with M1 disease at diagnosis and disease evolution differed between cancers. For instance, almost one-half of the patients with pancreatic cancer had distant metastases at diagnosis as opposed to approximately 1% of patients with head and neck squamous cell carcinoma. Furthermore, certain cancers (e.g., NSCLC or SCLC) showed a clear increasing trend over time, whereas others clearly decreased (e.g., prostate) (see Figure 1, A–D; Supplementary Table 4, available online).

Table 1. Per cancer type and time period, the number of newly diagnosed patients with *de novo* metastatic disease and the mean age

	1989-93	Mean age (95% CI)	Male (%)	2014-18	Mean age (95% CI)	Male (%)	Increase or decrease
Cancer[a]							
Prostate	5477	74 (73-74)	5477 (100)	8491	73 (73-73)	8491 (100)	↓
Testicular	210	32 (30-33)	210 (100)	293	35 (34-36)	293 (100)	↑
Thyroid	182	67 (65-69)	64 (35)	259	67 (66-69)	116 (45)	≠
Breast	2808	64 (63-64)	12 (0)	3826	63 (62-63)	42 (1)	↓
NET	304	64 (63-65)	158 (52)	840	64 (63-65)	466 (55)	≠
Vulva/vagina	27	70 (65-76)	0 (0)	85	71 (69-74)	0 (0)	↑
Melanoma	440	57 (55-58)	249 (57)	1041	63 (62-64)	639 (61)	↑
Ovarian	1060	66 (65-67)	0 (0)	1517	68 (68-69)	0 (0)	↑
Corpus uteri	248	69 (68-70)	0 (0)	672	70 (69-71)	0 (0)	
GIST	50	59 (56-63)	21 (42)	160	66 (64-68)	101 (63)	↑
Rectum	1956	67 (67-68)	1159 (59)	3644	66 (66-66)	2254 (62)	↓
Cervix uteri	144	59 (57-62)	0 (0)	373	58 (57-60)	0 (0)	↓
Sarcoma	390	57 (55-59)	214 (55)	545	60 (59-62)	329 (60)	↑
Colon	4753	68 (68-69)	2259 (48)	9332	68 (68-68)	4961 (53)	≠
Kidney	1489	65 (64-65)	918 (62)	2205	67 (67-68)	1467 (67)	↑
HNSCC	97	62 (60-64)	77 (79)	238	66 (64-67)	178 (75)	↑
Bladder	549	69 (68-70)	384 (70)	1768	70 (70-71)	1164 (66)	↑
Other solid cancers	811	65 (64-66)	364 (45)	1187	65 (64-65)	586 (49)	≠
Cancers with an unknown primary site	11659	69 (69-69)	6406 (55)	5280	73 (72-73)	2661 (50)	↑
SCLC	3575	66 (65-66)	2871 (80)	4873	67 (67-68)	2564 (53)	↑
Esophagus/Cardia	1646	65 (64-65)	1229 (75)	4387	67 (67-67)	3464 (79)	↑
Stomach	2720	69 (68-69)	1697 (62)	2148	69 (68-69)	1329 (62)	≠
NSCLC	7510	65 (65-65)	6055 (81)	19868	67 (67-67)	11,262 (57)	↑
Lung other	1045	70 (70-71)	874 (84)	2890	76 (75-76)	1706 (59)	↑
Bile Ducts	289	69 (67-70)	133 (46)	1347	67 (66-67)	665 (49)	↓
Hepatocellular	114	62 (59-65)	69 (61)	508	68 (67-69)	395 (78)	↑
Gallbladder	339	71 (70-72)	73 (22)	383	70 (69-71)	111 (29)	↓
Pancreas	2371	67 (66-67)	1288 (54)	6223	69 (69-69)	3275 (53)	↑
Total	52,263	67 (67-67)	20,002 (62)	84,383	68 (68-68)	48,519 (57)	↑

Abbreviations: CI: Confidence Interval; GIST: Gastrointestinal stromal tumors; HNSCC: head and neck squamous cell carcinoma; NETs: Neuroendocrine Tumors; NSCLC: non-small cell lung cancer; SCLC: small cell lung cancer sign: significance; SD: standard deviation. ; a See Supplementary Table 2. for a detailed description of the cancer types; ↑increase; ↓ decrease

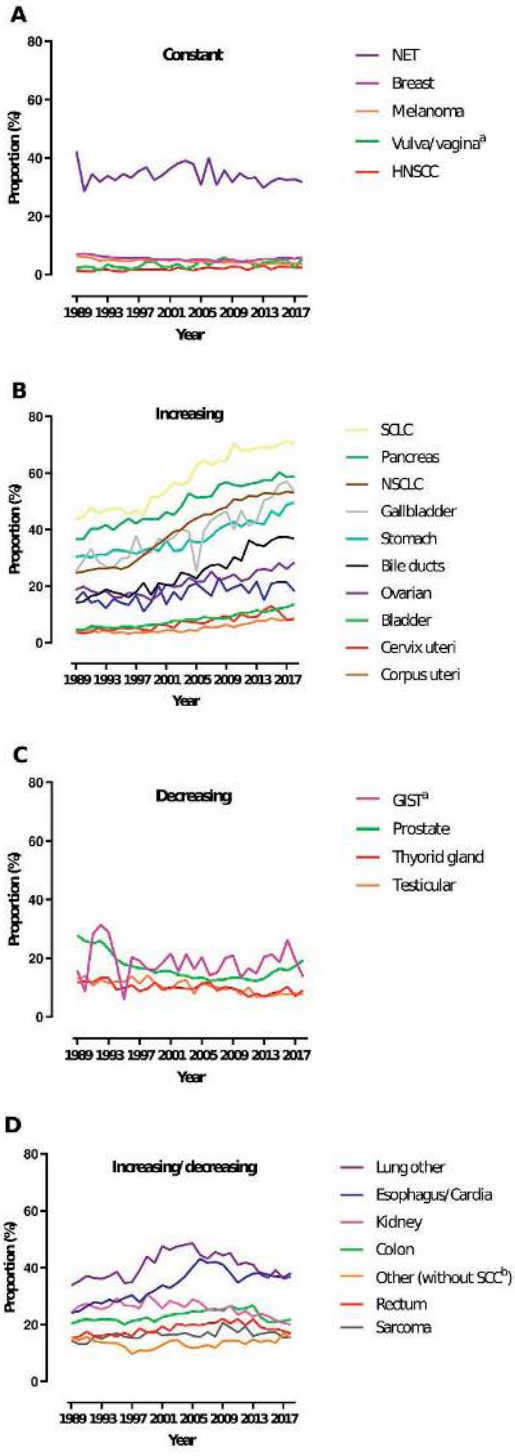

Figure 1. Trends in the proportion of patients with metastatic cancer at diagnosis M1/(M0 + M1)) per cancer type from 1989 to 2018. A) Constant trend. B) Increasing trend. C) Decreasing trend. D) Increasing and decreasing trend.

[a]Fluctuations because of small patient numbers in each year. [b]Patients with SCC excluded. Note: The increases in the proportion of M1 disease may be caused by various reasons (e.g., changes in the proportion of M1/(M0_M1) or changes primarily in the incident of M1 or M0). Therefore, the absolute patient numbers with M0 and M1 are also provided in Supplementary Table 4, available online. GIST=gastrointestinal stromal tumors; HNSCC=head and neck squamous cell carcinoma; M0=patients with locoregional cancer; M1=patients with de novo metastatic disease; NET=neuroendocrine tumor; NSCLC=non-small cell lung cancer; SCC=quamous cell carcinoma; SCLC=small cell lung cancer.

Changes in M1 survival

The 1- and 5-year age-standardized net survival rates increased over time for most cancers (Figure 2; Supplementary Figure 1; Supplementary Tables 5-7, available online). There remains a large gap in survival between cancers with the most and the least favorable survival. The percentage point increases in 1-year net survival ranged from 0 to 50, and that for 5-year survival ranged from 0 to 46. The largest gains in 1-year net survival were observed for GIST, NET, and cancers of the rectum, colon, and ovaries. The largest increases in 5-year net survival were seen in patients with GIST, NET, and cancers

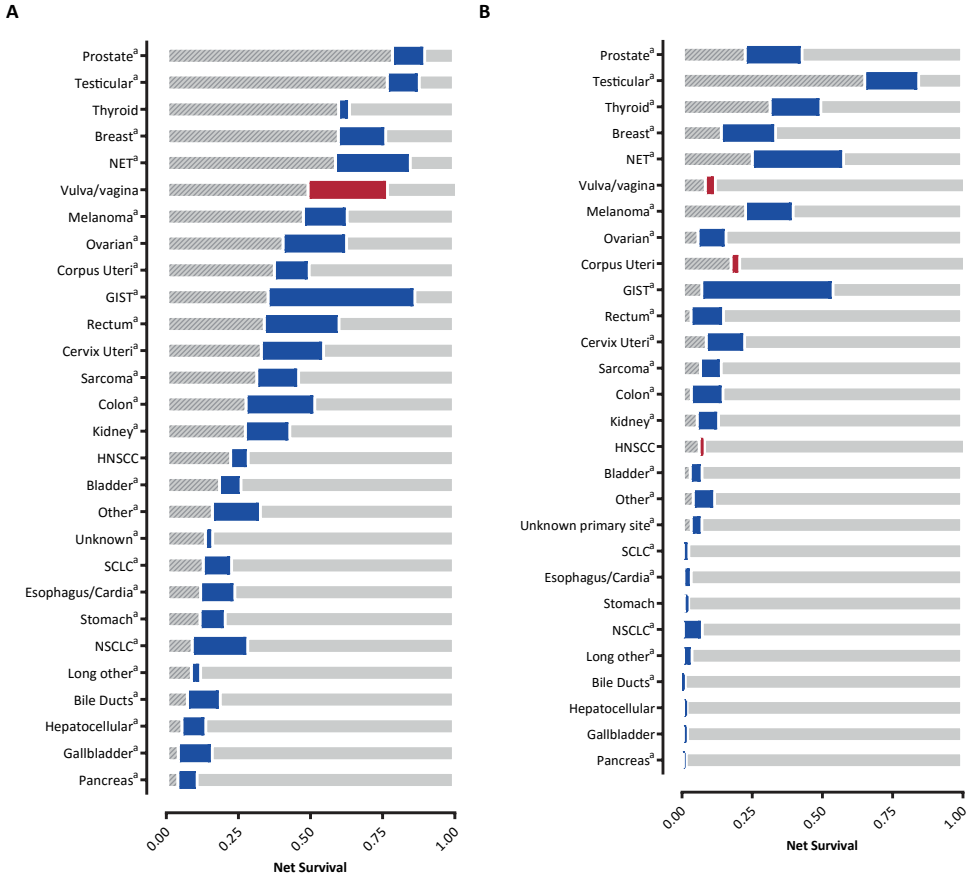

Figure 2. Changes in 1- and 5-year net survival of *de novo* metastatic cancer patients with primary cancer diagnosis over 2 time periods: 1989-1993 vs 2014-2018 (**blue** = increase, **red** = reduction). A) Changes in 1-year net survival. B) Changes in 5-year net survival. The **striped gray bar** on the left represents the net survival in the period 1989-1993; the **thicker blue bar** represents the survival change from 1989-1993; and **transparent gray bar** on the right represents the proportion of survival that is still to gain.

[a]Significant difference in net survival of patients in 1989-1993 vs those in 2014-2018 (a=.01). GIST=gastrointestinal stromal tumors; HNSCC=head and neck squamous cell carcinoma; NET=neuroendocrine tumor; NSCLC=non-small cell lung cancer; SCLC=small cell lung cancer.

of the prostate, breast, testicles, and thyroid (all subtypes). The increases were mainly seen from 1994-1998 onward for GIST, during the entire study period for NET, from 2004-2008 onward for prostate cancer, from 1999-2003 onward for breast cancer, in the earliest period for testicular cancer (i.e., 1989-1993 and 1994-1998), and from 1994-1998 onward for thyroid cancer (see Supplementary Figure 1, available online). Changes in median overall survival were also largest in GIST, NET, breast, and prostate cancer (see Supplementary Table 8, available online).

Overall, the net survival rate for M1 disease exceeded 50% at 1 year for 11 cancers and exceeded 20% at 5 years for 8 (2014-2018).

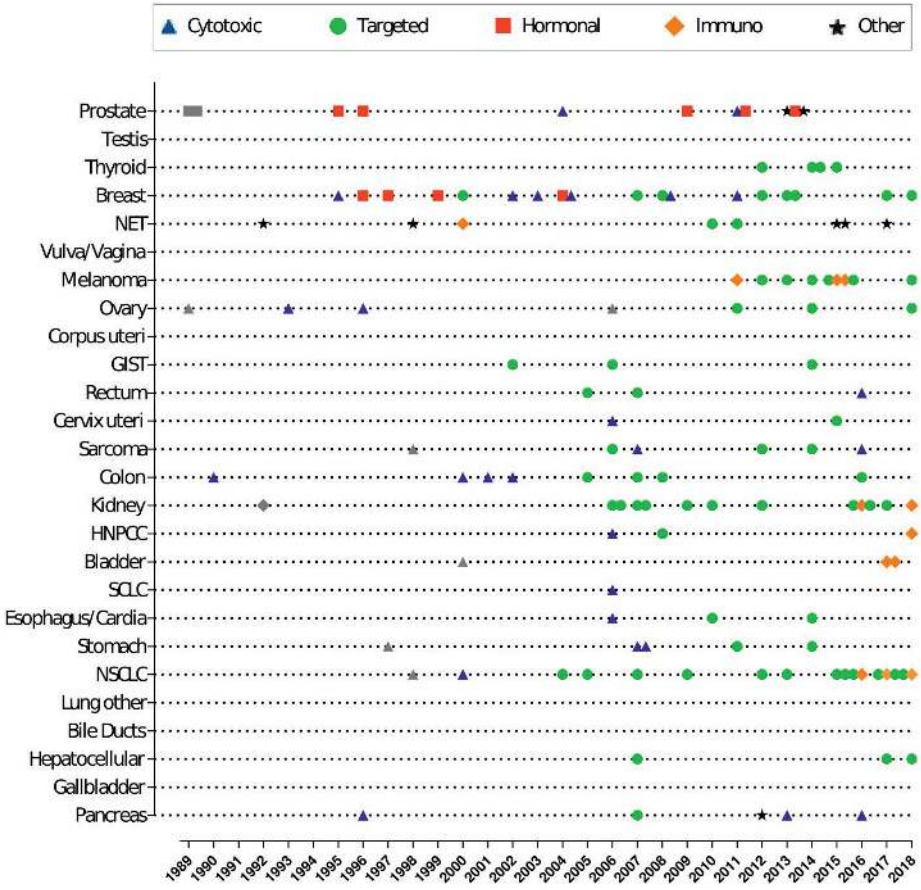

Figure 3. Timeline of newly approved medicines for patients with metastatic cancer per cancer type in the Netherlands. **Gray dots**, **squares**, and **triangles** show medicines for which the year of approval is uncertain (see Supplementary Table 10, available online). GIST=gastrointestinal stromal tumors; HNSCC=head and neck squamous cell carcinoma; NET=neuroendocrine tumor; NSCLC=non-small cell lung cancer; SCLC=small cell lung cancer.

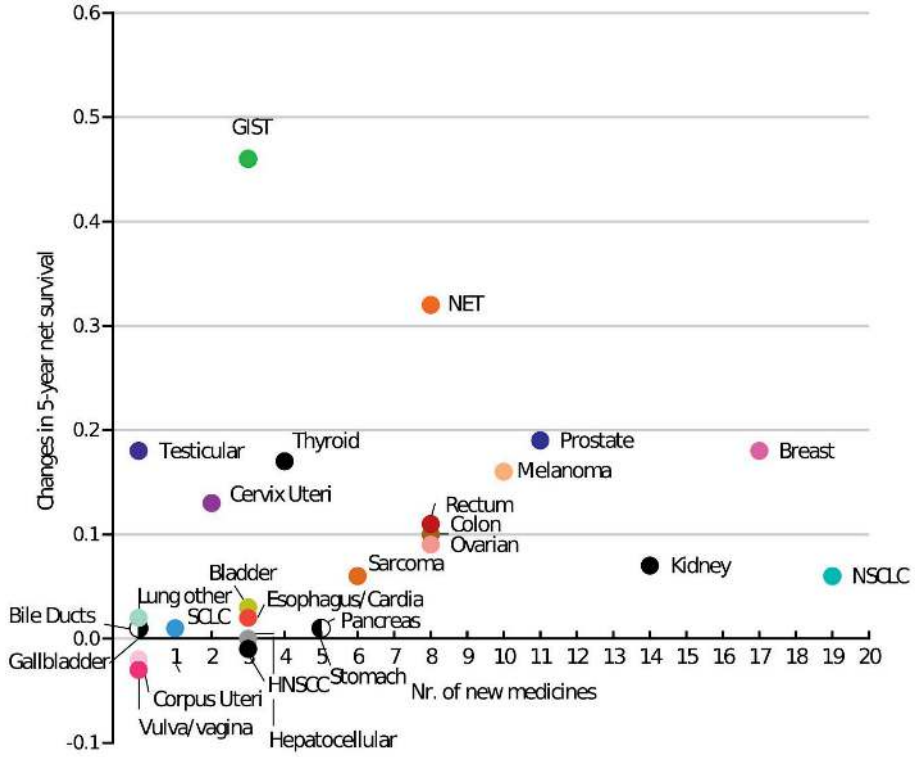

Figure 4. Association between the number of new medicines approved during the period 1989-2018 and the 5-year net survival changes from 1989-1993 and 2014-2018 of patients with different solid cancers. GIST=gastrointestinal stromal tumors; HNSCC=head and neck squamous cell carcinoma; NET=neuroendocrine tumor; NSCLC=non-small cell lung cancer; SCC=squamous cell carcinoma; SCLC=small cell lung cancer.

DISCUSSION

This study showed that the 1-year survival increased for most patients with M1 disease, but the magnitude of increase differed substantially between cancers. Additionally, we observed increases (>15%) in the 5-year survival for GIST, NET, melanoma, and cancer of the prostate, breast, thyroid, and testis. Approval was granted for more than 80 novel medicines and multiple new indications for existing medicines to treat metastatic disease (Figure 3; Supplementary Tables 9 and 10, available online). Although introducing new medicines probably contributed to some of the observed survival increases, the changes in survival did not appear to relate to the number of new medicines approved (Figure 4). Besides, we cannot assume causation because we do not know how many of the patients actually received the newly approved medicines. Moreover, a higher number of new treatment options does not always result in better survival at the population level.

Interpreting survival trends over several decades is complex because various factors can explain any observed changes. Together with better treatment, factors such as better staging due to improvements in imaging and changes in registration, coding practices and TNM classifications (both causing stage migration), and distribution of prognostic factors (e.g., subtypes) could account for the survival improvements. Nevertheless, it appears plausible that at least part of the survival increases observed in our study represents true treatment progress, such as the substantial 18-46 percentage point increases in 5-year survival for breast cancer, NET, and GIST that coincided with the approval of novel medicines.

The tyrosine kinase inhibitor imatinib, for instance, was approved in 2002 for the treatment of metastatic GIST (see Figure 3). In RCTs, patients treated with this agent showed median survival times of approximately 50-60 months, increasing from historic medians of just 9 months (266, 267). Approximately 84% of patients in our 2014-2018 cohort received a tyrosine kinase inhibitor as first-line therapy, suggesting that imatinib has been widely implemented in the Netherlands (Supplementary Table 11, available online). In addition, survival mainly increased from 1994-1998 to 1999-2003 (Supplementary Figure 1, available online). Hence, at least some of the observed survival increase likely reflects the adoption of effective treatment options for *de novo* metastatic GIST.

Concerning *de novo* metastatic breast cancer, the 5-year net survival in our study increased by 18 percentage points during a period when 17 new medicines were approved. Several of these, such as the targeted agents trastuzumab and pertuzumab, can improve survival in patients with HER2+ metastatic breast cancer. Although this subgroup accounts for only approximately 15% of patients, we consider it plausible that the successful implementation of these medicines accounts for some of the observed improvements in 5-year survival, especially given that most patients in the 2014-2018 cohort received one of these medicines as a first-line therapy (Supplementary Table 11, available online). Unfortunately, we could not separately evaluate the survival of this subgroup because HER2 status was not recorded in the NCR before 2005. Other medicines that probably contributed to the improved survival in our study include aromatase inhibitors and taxanes (268).

Octreotide represented an important breakthrough in the treatment of metastatic NET when implemented in the Netherlands in 1992 (269). Several other therapies were also granted regulatory approval since then. In our study, the number of patients diagnosed with NET markedly increased over time, likely reflecting improved pathology assessment and greater awareness of these uncommon cancers. Moreover, NETs are a heterogeneous group of tumors for which survival depends on multiple factors, including the histology and primary site, the distribution of which varied between the periods in this

study (270, 271). Therefore, the substantial survival increase of 32 percentage points is probably not only the effect of advances in treatment.

In addition to these cancers, our data uncovered substantial improvements of 16 percentage points in long-term survival from metastatic melanoma. These findings are relevant because the treatment landscape has dramatically changed with the advent of immune checkpoint inhibitors (ICIs) and multiple targeted agents over recent years. ICIs represent particularly important breakthroughs, improving the 5-year survival from metastatic melanoma by up to 52% in RCTs from historic levels less than 10% (210, 272). Studies investigating the impact of ICIs in daily practice have also revealed an improved prognosis (273). The NCR regrettably lacks data on therapies not administered first line, and thus we do not know how many patients received ICI. However, previous data in the Netherlands suggest the rapid implementation of new systemic options for melanoma (274). The survival increases observed in our study may therefore reflect this trend of ICI application.

We found long-term survival increases of 6 percentage points in patients with NSCLC. Stage migration since implementation of fluorodeoxyglucose positron emission tomography in 1997 likely accounts for some of this increase in the Netherlands. Since then, the percentage of patients with M1 NSCLC increased from 29% to 53% (Figure 1), but the limited survival increases are disappointing given that it received most novel medicines (Figure 4). The minor changes in population based long-term progress can be explained in several ways. First, some patients in the 2014-2018 cohort may not have received the newest and most effective agents because many of these had only recently been implemented (e.g., ICIs in 2016). Second, we cannot expect long-term benefits for all patients (e.g., 5-year survival of 16%-33% with ICIs in RCTs) (272, 275). Third, our follow-up duration may not have been long enough to capture improvements in 5-year survival because of the latest medicines. Future studies should confirm whether survival from metastatic NSCLC improves over time.

Despite the breakthroughs in systemic treatment for metastatic solid cancers, debate persists regarding the effectiveness and rising costs of new cancer medicines (276). Some may bring few health benefits, with a recent study estimating that approximately 22% of all newly approved medicines for solid cancer between 2003 and 2013 offered no overall survival benefits compared with standard care (277). Although other medicines had a positive impact, with a mean 3-month improvement in survival, the magnitude was often modest (277). This may also explain why the 1- and 5-year survival rates of some cancers have changed little in the last 30 years. Nevertheless, even minor benefits

in survival or other outcomes (e.g., quality of life) may represent progress in treating patients with metastatic cancer.

Analysis by stage requires appropriate data registration and few unknown or missing data. The NCR provided accurate stage information for the entire period, which facilitated the analysis of changes in M1 survival in a comprehensive list of cancers. By providing an overview of survival changes for all M1 solid cancers, our study supplements existing knowledge. Nevertheless, further research is needed to improve our understanding of the population impact of new medicines for metastatic cancer. As a prerequisite, databases should include data beyond the first-line treatment together with confounding variables (i.e., co-morbidity) and other relevant outcomes (i.e., quality of life, adverse events). A complete overview will also require further study in patients with metachronous metastatic disease. Unfortunately, it is not possible to easily obtain such data.

Our results show that the survival of patients with *de novo* metastatic cancer improved slowly over 30 years but that these gains were typically modest and unevenly distributed among cancers. Unfortunately, metastatic cancer remains a very lethal disease for almost all cancer types. Next to better treatment options to improve survival in patients with metastatic cancer, we call for better preventive measures and early detection to reduce the incidence of metastatic disease

Supplemental materials can be found online at: doi: 10.1093/jnci/djad020

CHAPTER 7

COMPARISON BETWEEN *DE NOVO* AND METACHRONOUS METASTATIC BREAST CANCER

Published as:
Comparison between *de novo* and metachronous metastatic breast cancer: the presence of a primary tumour is not the only difference —a Dutch population-based study from 2008 to 2018.

de Maar, Luyendijk M, Suelmann M, van der Kruijssen DEW, Elias SG, Siesling S, van der Wall E. Breast Cancer Research and Treatment. 2023;198:253–264.

ABSTRACT

Purpose The aim of this study was to compare characteristics and survival of patients with *de novo* and metachronous metastatic breast cancer.

Methods Data of patients with metastatic breast cancer were obtained from the Netherlands Cancer Registry. Patients were categorized as having *de novo* metastatic breast cancer (n = 8656) if they had distant metastases at initial presentation, or metachronous metastatic disease (n = 2374) in case they developed metastases within 5 or 10 years after initial breast cancer diagnosis. Clinicopathological characteristics and treatments of these two groups were compared, after which multiple imputation was performed to account for missing data. Overall survival was compared for patients treated with systemic therapy in the metastatic setting, using Kaplan Meier curves and multivariable Cox proportional hazards models. The hazard ratio for overall survival of *de novo* versus metachronous metastases was assessed accounting for time-varying effects.

Results Compared to metachronous patients, patients with *de novo* metastatic breast cancer were more likely to be ≥ 70 years, to have invasive lobular carcinoma, clinical T3 or T4 tumours, loco-regional lymph node metastases, HER2 positivity, bone only disease and to have received systemic therapy in the metastatic setting. They were less likely to have triple negative tumours and liver or brain metastases. Patients with *de novo* metastases survived longer (median 34.7 months) than patients with metachronous metastases (median 24.3 months) and the hazard ratio (0.75) varied over time.

Conclusions Differences in clinicopathological characteristics and survival between *de novo* and metachronous metastatic breast cancer highlight that these are distinct patients groups.

INTRODUCTION

In the Netherlands, around 14,500 patients annually are diagnosed with invasive breast cancer (205). Around 5% of these patients present with *de novo* distant metastases at the time of initial diagnosis (278). Moreover, in 15–20% distant metastases are diagnosed in the years following their initial breast cancer diagnosis (metachronous metastases) (279, 280). While systemic treatment, with a palliative intent, is the standard of care for both *de novo* and metachronous metastatic breast cancer (MBC) (170, 281), there are specific therapeutic considerations for each group. For instance, in *de novo* MBC the best approach regarding the primary tumour is still unclear. Many studies suggested an overall survival (OS) benefit of local treatment (282, 283, 284), but recent randomized studies have refuted this (285, 286). Unlike patients with *de novo* MBC, many patients with metachronous MBC have already received (neoadjuvant or adjuvant) systemic treatment in addition to loco-regional treatment following diagnosis of the primary tumour. Recurrence following these previous systemic therapies could reflect resistance to these drugs or mean that the maximum tolerated cumulative dose of these drugs was already reached. Moreover, patients can suffer from lasting side effects and therefore be less fit for further systemic treatment. These specific considerations illustrate the importance of understanding differences between patients with *de novo* and metachronous MBC.

So far, characteristics and OS of patients with *de novo* MBC have been analysed in several cohorts (287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302) and in some studies have been compared to patients with metachronous MBC (226, 280, 296, 303, 304, 305, 306, 307, 308, 309, 310). Typically, patients with *de novo* MBC have more favourable disease characteristics and longer OS compared to metachronous patients (226, 280, 296, 303, 304, 305, 306, 307, 308, 309). However, in younger patients (≤ 40 years) the opposite appears to be true (i.e., larger tumours, more widespread metastatic disease and more brain metastasis in *de novo* MBC) (296). Differences in gene expression profiles between *de novo* and metachronous MBC have been found, which lead to believe that these tumours possess biologically different behaviour (310). Moreover, whereas studies consistently reported improvements in OS over time for *de novo* MBC (225, 278, 311, 312), little evidence supports such a positive trend in metachronous MBC (312), which again emphasizes the difference between these groups. While literature on *de novo* MBC is often based on nationwide registry data, data on metachronous MBC is usually based on regional cohorts (312). Our study demonstrates nationwide data of patients with *de novo* MBC and a large cohort of patients with metachronous MBC diagnosed in 2008–2018 in the Netherlands, to compare clinicopathological characteristics, treatment and survival.

PATIENTS AND METHODS

Data source

The Netherlands Cancer Registry (NCR) is a nationwide cancer registry hosted by the Netherlands Comprehensive Cancer Organisation (IKNL) and includes all patients with newly diagnosed cancer, with an estimated coverage of 96% (89). Cancer diagnoses are notified through the nationwide Pathology Archive (PALGA) and the National Registry of Hospital Discharge Diagnoses. Trained data managers register data on diagnosis, clinicopathological characteristics and primary treatment directly from the patient files. Tumour location and morphology are coded according to the International Classification of Diseases for Oncology (ICD-O, third edition) and tumour stage is coded following the Tumour Node Metastasis (TNM) classification. Estrogen receptor (ER) and Progesterone receptor (PR) positivity of the primary tumour are set at ≥ 10% according to Dutch nationwide guidelines. Additional data on recurrences (including local, regional recurrences and distant metastases) were collected by the NCR retrospectively. Specifically, for patients with a primary diagnosis in 2003 and 2005 all recurrences up to 10 years after initial breast cancer diagnosis were identified. In addition, for half of the patients diagnosed in 2008, the first recurrence up to 5 years afterwards was identified and for patients diagnosed in the first quartile of 2012 (Q1 2012), all recurrences up to 5

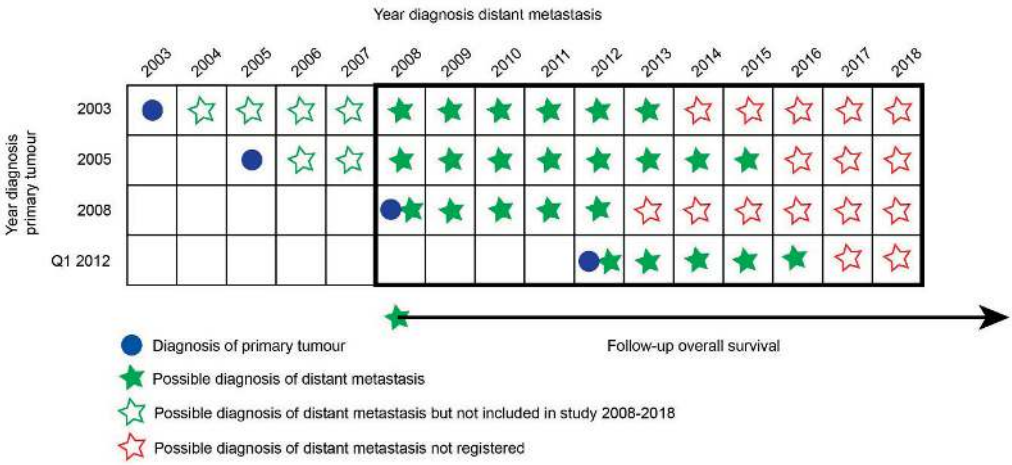

Figure 1. Selection of patients with metachronous metastatic breast cancer.
Patients diagnosed with metachronous distant metastases in 2008–2018 were identified among patients diagnosed with primary tumours in 2003, 2005, 2008 and the first quartile of 2012 (Q1 2012). For patients with a primary diagnosis in 2003 and 2005 all recurrences up to 10 years after initial breast cancer diagnosis were identified. For half of the patients with a primary diagnosis in 2008 the first recurrence up to 5 years after initial diagnosis was identified. For patients diagnosed in the first quartile of 2012, all recurrences up to 5 years post diagnosis were identified. Blue circles indicate primary tumour diagnosis, green filled stars indicate a possible diagnosis of distant metastasis included in our study. Empty stars indicate a possible diagnosis of distant metastasis that was not included in our study because it was outside the study period (green) or because it was not registered (red)

years after diagnosis were recorded (Figure 1). Recurrence locations were also registered following the ICD-O third edition. After diagnosis of metastases, the first systemic and/or local treatment is registered. Information on vital status was derived from the official municipal population database.

Patient selection

All patients diagnosed with *de novo* MBC between 2008 and 2018, as well as registered patients diagnosed with distant metachronous metastases between 2008–2018 were selected from the NCR. Patients with distant metastasis within 3 months of the primary diagnosis were excluded, because they are often considered as *de novo* MBC. For the survival analyses we excluded a small group of patients (n = 333, 3.0%) with tumour morphologies other than ductal carcinoma NOS or invasive lobular carcinoma (ILC). Moreover, we excluded patients from the survival analyses who had not received systemic therapy in the metastatic setting, because untreated patients likely have severe comorbidities and inherently have a different prognosis. Moreover, knowledge about patients eligible for systemic treatment can support medical oncologists' clinical decision-making.

Data definition

Patient characteristics (age at diagnosis of metastasis, sex, performance status), primary tumour characteristics (morphology, multifocality, tumour grade, clinical T stage, receptor and HER2 status), clinical N stage at primary diagnosis and location of metastases and treatment (local and systemic therapy after diagnosis of primary tumour and metastases) were analysed. Receptor and HER2 status of the metachronous metastases were not available. Period of metastasis diagnosis was categorized in 2008–2011, 2012–2015 and 2016–2018. Within metachronous MBC, metastasis free interval (MFI), defined as time between primary diagnosis and distant recurrence, was categorized in 3–12 months, 12–24 months, 24–60 months and MFI > 60 months. OS was analysed using time between diagnosis of distant metastasis and death or end of follow-up. If patients were alive at the end of follow-up (January 31st 2022), they were censored.

Statistical analysis

Descriptive statistics were used to depict clinicopathological and treatment characteristics and to describe missing data. Chi-squared tests or Fisher's exact tests were used to test difference in characteristics between *de novo* and metachronous MBC. To describe OS of the groups, Kaplan Meier curves were plotted and multivariable Cox proportional hazards (PH) analyses were performed including important confounders. Because some confounders included missing data, we used multiple imputation by chained equations (MICE) (313). Data were incomplete for any of the chosen variables in 19% of patients,

therefore we considered 19 simulated datasets to have a sufficiently reliable estimation of missing values (314, 315). The imputation model included MFI, which was used to categorize *de novo* versus metachronous MBC, and the following confounders: year of metastasis diagnosis; sex; age at metastasis diagnosis; tumour morphology; tumour multifocality; clinical T stage; clinical N stage; receptor status; metastasis location; tumour grade; performance status (assessed at primary diagnosis) and therapy variables, plus the outcome variables (vital status and the Figure 1 Selection of patients with metachronous metastatic breast cancer. Patients diagnosed with metachronous distant metastases in Nelson-Aalen estimator) as was recommended in literature (316, 317, 318, 319).

The multivariable Cox PH model comparing survival of patients with *de novo* and metachronous MBC included the same variables included in the imputation model, except for tumour grade and performance status, which were considered confounders but contained too many missings. For the variable 'age at metastasis diagnosis' we used a restricted cubic spline with four knots. We also excluded therapy variables for several reasons. First, therapy choices are partly determined by the variable of interest (*de novo* versus metachronous MBC influences therapy choices) and are therefore not a confounder but an intermediary variable. Second, specifically in *de novo* MBC a RCT reported that local therapy of the primary tumour did not improve OS (285, 286), so local therapy is not a confounder either. Likelihood ratio tests of Cox PH models were used to compare OS between the groups. The PH assumption is the most important assumption underlying the Cox model (320). The assumption was tested for each variable included in the Cox model using Schoenfeld residuals plots (321)). Confounding variables that did not meet the PH assumption were added as strata to the model. For the variable of interest (*de novo* versus metachronous MBC), we visualised the time-varying effect of the hazards by plotting the hazard ratio (HR) against time. For this purpose, we generated a time dependent Cox model with an interaction between the variable (*de novo* vs metachronous) and a restricted cubic spline of survival time with five knots (322). This time dependent model was based on one imputation dataset.

In addition to the described analyses we performed supplementary analyses to study patients with *de novo* and metachronous MBC separately in more detail, to test the robustness of our results and to explore hypotheses behind the differences in OS (Supplementary Methods 1, available online).

A p value < 0.05 was considered significant. Statistical analyses were performed in Stata/SE 17.0 and R version 4.0.3.

RESULTS

Clinicopathological characteristics

Between 2008 and 2018, 8656 patients were diagnosed with *de novo* MBC in the Netherlands. In addition, 2374 patients were identified with metachronous MBC, of those, 639 had a primary tumour diagnosed in 2003, 1006 in 2005, 524 in 2008 and 205 in Q1 of 2012.

Table 1 lists patient, tumour and treatment characteristics in patients with *de novo* and metachronous MBC. Most notable differences were those in T and N stages, metastasis locations and receptor status. Patients with *de novo* MBC were more likely than metachronous to have T3 or T4 tumours and positive loco-regional lymph nodes, while the majority (68%) of metachronous MBC had N0 at time of primary diagnosis. Notably, *de novo* metastases were more commonly limited to the bone with less frequent involvement of the liver or central nervous system (CNS) than in metachronous disease. However, in young *de novo* patients (≤ 40 years, n = 489) we saw more liver metastases (39%, versus 25% in *de novo* patients of all ages and 38% in young metachronous patients). CNS involvement was the same (3%) in *de novo* patients above or below 40 years of age. ER-negative/HER2-negative tumours were observed less in *de novo* MBC. Although ER positivity did not differ between *de novo* patients and the entire group of metachronous MBC, supplementary analyses showed that ER-positive (HER2 negative/unknown) tumours were more common in patients with metachronous MBC with a MFI > 60 months (83%), while patients with shorter MFI's had less ER positive tumours (39 to 65%) than patients with *de novo* MBC (67%) (Supplementary Table 1, available online). Supplementary Table 2 (available online) shows changes in patients with *de novo* MBC over time.

Treatment characteristics

Local treatment of the primary breast tumour was performed in 26% of patients with *de novo* MBC, in 43% consisting of surgery combined with radiotherapy, and in all patients with metachronous MBC at initial diagnosis. 74% of patients with metachronous metastases had received systemic treatment after primary tumour diagnosis.

Systemic therapy for metastatic disease was administered in 89% of *de novo* and in 79% of metachronous stage IV patients. Chemotherapy (without HER2 targeting agents) was administered less often to patients with *de novo* MBC (24% vs 29%) (Table 1). Meanwhile, *de novo* ER positive patients received endocrine treatment more often (67% vs 48%) and *de novo* HER2 positive patients received HER2 targeted therapy more often (75% vs

Table 1. Descriptive statistics *de novo* versus metachronous MBC, without multiple imputation

Characteristic	De novo or metachronous MBC patients				Statistics
	De novo MBC (N= 8,656)		Metachronous MBC (N=2,374)		
	N	(%)	N	(%)	P
Patient characteristics					
Sex					
Female	8,564	(99)	2,351	(99)	Pearson Chi^2 p = 0.690
Male	92	(1)	23	(1)	
Age at diagnosis of metastasis					
<50	1,774	(20)	457	(19)	Pearson Chi^2 **p<0.001**
50-69	3,658	(42)	1,135	(48)	
70+	3,224	(37)	782	(33)	
Age at diagnosis of metastasis	Mean age 63.7 (22-102)		Mean age: 63.2 yrs (24.6-97.7)		T-test p = 0.1296
Metastasis free interval (MFI)					
MFI 0 (*de novo*)	8,656		-		Not applicable
MFI 3-12 months	-		87	(3.7)	
MFI 12-24 months	-		204	(8.6)	
MFI 24-60 months	-		1,061	(44.7)	
MFI > 60 months	-		1,022	(43.0)	
Tumour characteristics					
Clinical T stage					
T0 or Tis	167	(2)	6	(0.3)	Pearson Chi^2 **p<0.001**
T1	1,226	(15)	1,022	(46)	
T2	2,889	(36)	917	(41)	
T3	1,130	(14)	151	(7)	
T4	2,675	(33)	115	(5)	
T unknown	569	-	163	-	
Multifocality					
No	5,679	(74)	1,842	(80)	Pearson Chi^2 **p<0.001**
Yes	2,037	(26)	447	(20)	
Unknown	940	-	85	-	
Clinical N stage					
N0	1,771	(22)	1,529	(68)	Pearson Chi^2 **p<0.001**
N1	4,389	(54)	689	(31)	
N2	337	(4)	14	(1)	
N3	1,594	(20)	21	(1)	
N unknown	565	-	121	-	
Tumour morphology					
Ductal carcinoma NOS	7,030	(81)	1,997	(84)	Pearson Chi^2 **p<0.001**
Ductal carcinoma in situ	33	(0.4)	31	(1)	
Invasive lobular carcinoma	1,361	(16)	309	(13)	
Low grade special types	105	(1)	29	(1)	
Other	127	(1)	8	(0.3)	
Tumour grade					

Table 1. Descriptive statistics *de novo* versus metachronous MBC, without multiple imputation (continued)

Characteristic	De novo or metachronous MBC patients				Statistics
	De novo MBC (N= 8,656)		Metachronous MBC (N=2,374)		
	N	(%)	N	(%)	P
Low grade	287	(8)	206	(10)	Too many missings, no test performed
Intermediate grade	1,865	(51)	939	(45)	
High grade	1,514	(41)	963	(46)	
Undifferentiated/anaplastic	9	(0.2)	0	(0)	
Unknown	4,981	-	266	-	
ER status					
ER negative (<10%)	1,807	(21)	512	(22)	Pearson Chi2 p = 0.643
ER positive (≥10%)	6,641	(79)	1,833	(78)	
ER unknown/not determined	208	-	29	-	
PR status					
PR negative (<10%)	3,494	(43)	928	(42)	Pearson Chi^2 p = 0.468
PR positive (≥10%)	4,700	(57)	1,293	(58)	
PR unknown/not determined	462	-	153	-	
Her2 status					
Her2 negative	5,934	(77)	1,471	(80)	Pearson Chi^2 **p = 0.002**
Her2 positive	1,767	(23)	358	(20)	
Her2 unclear	955	-	545	-	
Receptor status					
ER pos HER2 neg/unknown	5,629	(67)	1,622	(69)	Pearson Chi^2 **p<0.001**
Her2 pos	1,767	(21)	358	(15)	
ER neg Her2 neg/unknown	1,072	(13)	368	(16)	
Insufficient information	188	-	26	-	
Localization of metastasis					
Lymph nodes only	468	(5)	60	(3)	Pearson Chi^2 **p<0.001**
Bone only	2,949	(34)	510	(22)	
Liver (no CNS)	2,177	(25)	727	(31)	
CNS (with/without liver)	274	(3)	290	(12)	
All other locations	2,760	(32)	774	(33)	
Metastasis location unknown	28	-	13	-	
Treatment characteristics					
Local treatment primary tumour					
No local treatment primary	6,383	(74)	1	(0)	Fisher's exact **p<0.001**
Surgery and radiotherapy	975	(11)	1,624	(68)	
Surgery without radiotherapy	996	(12)	749	(32)	
Radiotherapy without surgery	302	(3)	0	(0)	
Any systemic therapy after previous primary tumour diagnosis					
No	Not applicable		617	(26)	Not applicable
Yes			1,757	(74)	
Unknown			0	-	

Table 1. Descriptive statistics *de novo* versus metachronous MBC, without multiple imputation (continued)

Characteristic	De novo or metachronous MBC patients				Statistics
	De novo MBC (N= 8,656)		Metachronous MBC (N=2,374)		
	N	(%)	N	(%)	P
Any systemic therapy in metastatic setting					
No	937	(11)	478	(21)	Pearson Chi^2 **p<0.001**
Yes	7,719	(89)	1,816	(79)	
Unknown	0	-	80	-	
Systemic therapy in metastatic setting contains chemotherapy					
No	5,239	(61)	1,383	(60)	Fisher's exact **p<0.001**
Yes	3,417	(39)	824	(36)	
Systemic therapy of unknown type	0	(0)	87	(4)	
Unknown	0	-	80	-	
Systemic therapy in metastatic setting contains endocrine treatment					
No	4,385	(51)	1,312	(57)	Fisher's exact **p<0.001**
Yes	4,271	(49)	895	(39)	
Systemic therapy of unknown type	0	(0)	87	(4)	
Unknown	0	-	80	-	
Systemic therapy in metastatic setting contains targeted therapy					
No	6,924	(80)	1,903	(83)	Fisher's exact **p<0.001**
Yes	1,732	(20)	304	(13)	
Systemic therapy of unknown type	0	(0)	87	(4)	
Unknown	0	-	80	-	
Systemic therapy in metastatic setting					
No systemic therapy	937	(11)	478	(21)	Pearson Chi^2 **p<0.001**
Endocrine treatment	4,119	(48)	875	(38)	
Chemotherapy (without HER2 targeted therapy)	2,116	(24)	661	(29)	
HER2 targeted therapy	1,484	(17)	191	(8)	
Systemic therapy of unknown type	0	(0)	89	(4)	
Unknown	0	-	80	-	
Any local treatment of metastasis					
No	6,497	(75)	1,803	(76)	Pearson Chi^2 p = 0.309
Yes	2,159	(25)	567	(24)	
Unknown	0		4	-	

Table 1. Descriptive statistics *de novo* versus metachronous MBC, without multiple imputation (continued)

Characteristic	De novo or metachronous MBC patients				Statistics
	De novo MBC (N= 8,656)		Metachronous MBC (N=2,374)		
	N	(%)	N	(%)	P
Surgery of metastasis					
No	8,337	(96)	2,059	(87)	Pearson Chi^2 **p<0.001**
Yes	319	(4)	311	(13)	
Unknown	0	-	4	-	
Radiotherapy of metastasis					
No	6,728	(78)	2,071	(87)	Pearson Chi^2 **p<0.001**
Yes	1,928	(22)	299	(13)	
Unknown	0	-	4		

41%). Radiotherapy was the preferred locoregional treatment in *de novo* patients while surgery was more common in metachronous patients.

Supplementary Table 2 (available online) shows changes in treatment of patients with *de novo* MBC over time, these changes could not reliably be compared between *de novo* and metachronous MBC due to the method of registration for patients with metachronous metastases (from just four primary tumour years, Figure 1).

Survival in *de novo* versus metachronous MBC

Among patients treated with systemic therapy in the metastatic setting, median OS in patients with metachronous MBC was 24.3 months (95% CI 22.5–25.5 months), compared to 34.7 months (95% CI 33.7–35.8 months) in those with *de novo* MBC (Log-rank test $p < 0.001$) (Figure 2). In the multivariable Cox PH analysis a difference in OS between *de novo* and metachronous MBC persisted (HR *de novo* MBC versus metachronous MBC 0.75, 95% CI 0.70–0.80, Likelihood ratio test $p < 0.001$) (Table 2). Sensitivity analysis supported these findings when analysing complete cases and showed that within patients with metachronous MBC, patients with a longer MFI survived longer, although still not similar to patients with *de novo* MBC (Supplementary: Figure 1, Tables 3, 4, available online). Figure 3 shows the changes in HR of patients with *de novo* versus metachronous MBC over time, while keeping all other variables constant. Although the HR varied over time, *de novo* MBC appeared to have a lower risk of death than metachronous MBC over the entire period of follow-up. The difference is most pronounced two years after diagnosis of metastatic disease. Beyond approximately eight years the HR estimate becomes increasingly imprecise.

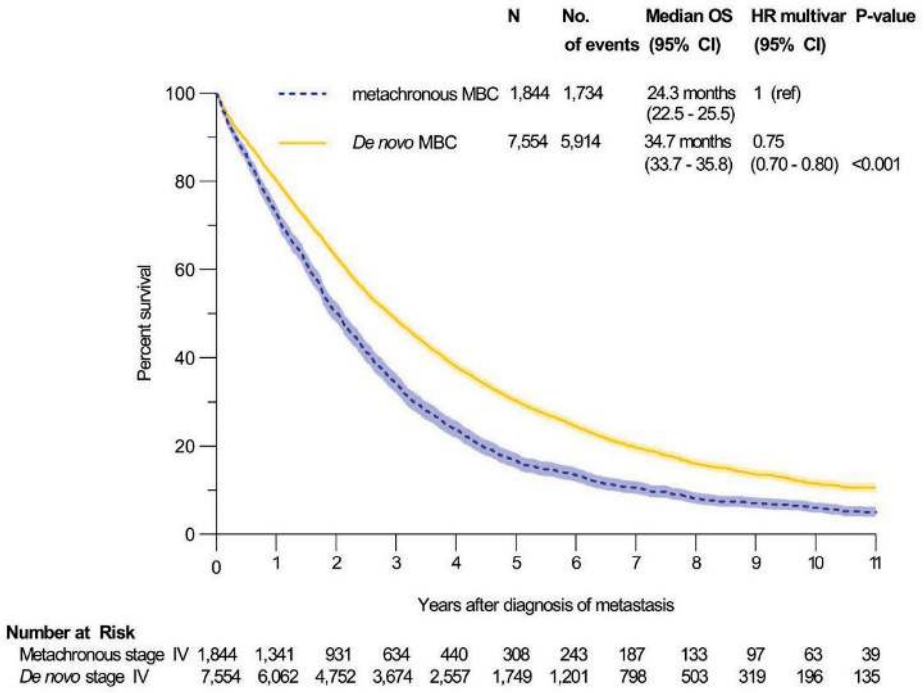

Figure 2. Overall survival in patients with versus patients metachronous metastatic breast cancer (treated with systemic therapy in the metastatic setting). The 95% confidence interval is indicated by colour around the line, number of patients at risk is noted below each year of follow-up. Overall survival is significantly longer in *de novo* MBC patients compared to metachronous MBC patients

Further exploratory analysis showed that *de novo* MBC is not always associated with longer survival. In fact, patients with metachronous MBC who had not received previous (neoadjuvant or adjuvant) systemic treatment following their primary tumour diagnosis (n = 461) survived longer than patients with *de novo* MBC (median OS 37.2 months, multivariable HR *de novo* MBC 1.13, 95% CI 1.01–1.27) (Supplementary Figure 3, Table 6, available online). Note that these were patients with favourable characteristics at primary tumour diagnosis, nearly all (97%) had T1 (75%) or T2 (22%) and 74% was ER positive and HER2 negative or unknown.

Table 2. Multivariable Cox proportional hazards analysis *de novo* versus metachronous MBC patients (treated with systemic therapy in metastatic setting)

Characteristic	Category	HR (95% CI)	Likelihood ratio test P Value
De novo versus metachronous MBC	Metachronous MBC	1 (ref)	
	De novo MBC	0.75 (0.70-0.80)	**<0.001**
Patient characteristics			
Period of metastasis diagnosis	2008-2011	1 (ref)	
	2012-2015	0.81 (0.77-0.85)	**<0.001**
	2016-2018	0.71 (0.66-0.76)	**<0.001**
Sexe	Male	1 (ref)	
	Female	0.86 (0.69-1.07)	0.173
Age at diagnosis of metastasis, restricted cubic spline (rcs))	Restricted cubic spline with four knots.		p-value omitted
Tumour characteristics			
Morphology	Ductal carcinoma NOS	1 (ref)	
	Invasive lobular carcinoma	1.12 (1.05-1.19)	**<0.001**
Multifocality	Unifocal primary tumour	1 (ref)	
	Multifocal primary tumour	0.96 (0.91-1.01)	0.149
Clinical T stage	T0 or Tis	1.11 (0.92-1.35)	0.269
	T1	1 (ref)	
	T2	1.09 (1.02-1.16)	**0.013**
	T3	1.13 (1.04 -1.24)	**0.006**
	T4	1.31 (1.22-1.42)	**<0.001**
Clinical N stage	N0	1 (ref)	
	N1	1.13 (1.06-1.20)	**<0.001**
	N2	1.20 (1.05-1.38)	**0.008**
	N3	1.20 (1.10-1.30)	**<0.001**
Receptor status	ER + HER2 - /unknown	Stratification factor	
	Her2 + (regardless of ER)		
	ER - HER2- / unknown		
Localization of metastasis	Lymph nodes only	Stratification factor	
	Bone only		
	Liver (no CNS)		
	CNS (with/without liver)		
	All other locations		

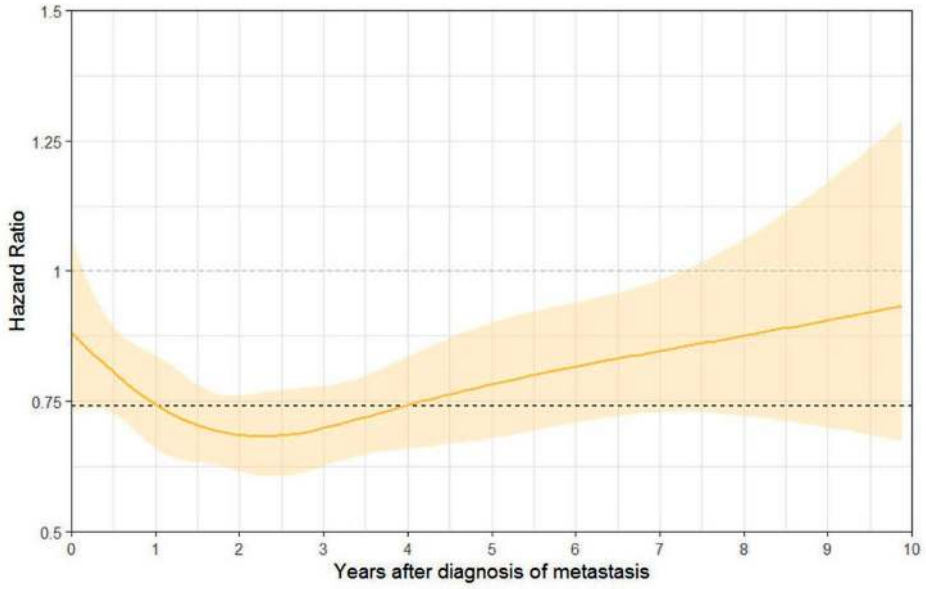

Figure 3. Overall survival in patients with *de novo* versus metachronous metastatic breast cancer (treated with systemic therapy in the metastatic setting), hazard ratio over time in multivariable analysis performed on the first multiple imputed dataset. The 95% confidence interval is indicated by colour around each line. The difference in overall survival was not proportional over time, with a lower HR in favour of *de novo* MBC in the first years of follow-up in the multivariable model, while the HR starts to rise towards 1.0 after about 5 years. Note that the confidence interval widens after about eight years

DISCUSSION

In this large population-based study, we compared clinicopathological features, treatment and OS between all patients with *de novo* MBC diagnosed in the Netherlands between 2008 and 2018 to a group of patients with metachronous MBC diagnosed in that same period. Our study shows that these are not only two very distinct groups but also that patients with *de novo* MBC survive longer.

A number of differences in characteristics between the groups are notable. Our data corroborate earlier reports of frequent bone and lymph node metastases and less involvement of viscera and brain in patients with *de novo* MBC (226, 280, 306). However, in our comparison of both groups, we did not observe more ER positive tumours in *de novo* patients (303), most likely caused by the 43% of metachronous patients with a MFI > 60 months, who are more often ER positive. Stage T3 and T4 and multifocal tumours were more often encountered in patients with *de novo* MBC, possibly reflecting delay in time to diagnosis. In *de novo* MBC, metastases limited to the bone were more common than

in metachronous MBC resulting in increased use of endocrine treatment in this group, as was local radiotherapy on painful bone metastases (170). The higher prevalence of triple negative tumours, liver and CNS metastases reflect unfavourable tumour characteristics and biology in metachronous patients. For example, triple negative tumours are known to metastasize hematogenously more often.

Regarding outcome, our data support a better OS in *de novo* MBC compared to patients with metachronous MBC, partly explained by differences in disease characteristics. Even after correction for known confounding characteristics, improved survival persists in patients with *de novo* MBC. This finding is consistent with previous literature (280, 296, 303, 305, 307, 308, 309) showing that absolute differences in median OS range from 1.5 months (when comparing *de novo* to MFI > 24 months) to 20.3 months (comparing to MFI < 24 months) (226). Without selection for MFI the reported median OS differences are similar to the 10.4 months found in our study (303, 305, 306, 307). In our sensitivity analysis we found a similar effect of MFI on OS compared to the literature: longer MFI was associated with longer OS after diagnosis of metastases. However, even MFI > 60 months did not have better survival than *de novo* MBC, which is not a consistent observation (226, 296, 308).

There are a number of possible explanations for the longer survival observed in *de novo* MBC. The most likely hypothesis is that shorter survival of patients with metachronous MBC is related to previous (neo)adjuvant systemic therapies. Recurrence despite previous therapy could reflect (1) disadvantageous tumour characteristics, (2) patient comorbidity/fitness and (3) primary or acquired therapy resistance. In our study, we could not quantify to what extent these factors played a role because we did not have data on the exact regimens and duration of systemic therapy nor could we account for fitness because data on performance score was only assessed at primary breast cancer diagnosis and contained too many missing values to include in the model. Nevertheless, the difference in use of systemic therapy in the metastatic setting (any systemic therapy in 89% of *de novo* and 79% of metachronous MBC and specifically targeted therapy in 75% of *de novo* and 41% of metachronous HER2 positive patients) may indicate that metachronous patients were less fit or had less treatment options for other reasons. As mentioned before, previous therapy can decrease treatment options in the metastatic setting due to acquired resistance to a drug, reaching a maximum tolerated cumulative dose or lasting side effects such as peripheral neuropathy or cardiotoxicity. Of note, due to the method of registration our data on systemic therapy could be an underestimation, as therapies administered not directly after metastasis diagnosis, but for example when symptoms did arise, may have been missed. In addition, the difference in targeted therapy among HER2 positive patients could be an overestimation because HER2 status

was determined on the primary tumour and (a small percentage of) metachronous patients may have converted to HER2 negative. The same might be true for endocrine treatment, as metastases of an ER positive primary may be ER negative.

Still, we did find some evidence to corroborate the hypothesis that previous systemic therapy plays a role in the survival difference between *de novo* and metachronous MBC. In a supplementary exploratory analysis we observed a longer OS among patients with metachronous MBC who were not systemically treated for their primary tumour. While this metachronous group had favourable characteristics at primary tumour diagnosis, the difference remained after correcting for baseline characteristics.

We also hypothesize that differences in metastatic burden could contribute to the observed differences in survival. Possibly, clinicians are inclined to perform more (and perhaps more sensitive) diagnostic imaging in a patient presenting with *de novo* MBC than in those diagnosed with recurrent disease. This would lead to detection of smaller, asymptomatic or oligo metastases in *de novo* MBC, associated with longer survival and possibly even curative treatment options. Although we have no data on number or volume of metastases, our data do support this hypothesis (Supplementary Table 2, available online) as we saw an increase in patients with *de novo* distant metastases limited to lymph nodes and increased use of anthracycline and taxane treatment (first choice in the neoadjuvant curative setting (170) and used for curative treatment of oligometastases).

In this study we extensively studied OS of patients with *de novo* and metachronous MBC using Kaplan Meier curves and Cox PH analysis. In the literature it is seldom reported whether the PH assumption was met and time-varying effects are often overlooked (323). In our study, the variable of interest (*de novo* versus metachronous MBC) did not meet the PH assumption and therefore we additionally estimated the time-varying effects on OS. Overall, it appeared that the OS difference between *de novo* and metachronous MBC persisted over the years. The relatively small difference in OS between *de novo* and metachronous MBC in the first year could mean that a group of patients progress and die quickly despite any beneficial characteristic. Apparently, differences between *de novo* and metachronous MBC start to count after surviving longer than a year.

This study is unique as it presents a complete overview of patients with *de novo* MBC diagnosed in 2008–2018 in the Netherlands and the comparison to patients with metachronous MBC. The data convincingly shows that patients with *de novo* and metachronous MBC are distinct patient groups. However, there are some limitations of our data. It would be relevant to also study metachronous patients in more detail using nationwide

data. The MFI of our patients was probably not an accurate representation of all patients with metachronous MBC in 2008–2018 because the majority of metachronous patients in our cohort had their initial diagnosis in 2003 or 2005 (thus MFI at least 5 or 3 years, respectively) (Figure 1). Another limitation is that for patients with an initial diagnosis in 2008, distant metastases were only registered if they did not have a local or regional recurrence preceding the occurrence of metastases. Due to this registration difference, we probably missed approximately 20% of patients with metachronous MBC and an initial diagnosis in 2008 (i.e. of patients with a primary tumour in 2003/2005/Q1 2012, about 20% had a local or regional preceding their distant metastases). Additional patients were missed because recurrences from 2008 had only been registered in half of the hospitals. Nationwide data including all patients with metachronous metastases in a given period would have allowed a more accurate comparison of the two groups. Ideally, such data would also include information on metastatic burden (e.g., oligometastases), receptor and HER2 status of the metastases and information about treatment administered in the metastatic setting in more detail and beyond those given as initial therapy.

CONCLUSION

Dutch patients with *de novo* MBC survive longer compared to patients with metachronous metastases, also following correction for different clinicopathological characteristics. Our data show that *de novo* and metachronous MBC represent two distinct groups, the presence of a primary tumour being not the only difference.

The other supplemental materials can be found online at: doi: 10.1007/s10549-022-06837-4

PART 3

REAL-WORLD DATA: STATE-OF-THE-ART GUIDANCE

CHAPTER 8

ASSESSING 'FIT-FOR-PURPOSE' OF A REAL-WORLD DATASET: LESSONS LEARNED

Submitted as brief communication:
Assessing 'fit for purpose' of a real-world dataset: lessons learned. 2025

Luyendijk M, Jager A, Bennink C, Heijns J, & Blommestein HM.

ABSTRACT

Objective: To describe the initial steps of an observational study, focusing on selecting and assessing the 'fit-for-purpose' of a real-world data (RWD) set to answer a defined research question.

Materials and methods: We applied the first three steps of available RWD frameworks to our study and assessed the suitability of the selected data both before and after gaining access to the dataset.

Results: We selected hospital claims data as RWD to answer our research question on resource use and costs in early-stage breast cancer. In theory, this data appeared to be a valuable source for our study, but in practice, we encountered several challenges that were not adequately identified prior to data access**.**

Conclusion: We showed that the expected 'fit-for-purpose' of RWD does not always align with the assessment made after having access to the data. Systematic suitability assessments help to determine the value of RWD.

BACKGROUND AND SIGNIFICANCE

Real-world data (RWD) are seen as an important source for research to provide insights into care of patients delivered in real-world settings. In today's data landscape, various data sources are available, each of these with their own strengths and weaknesses regarding size, content, quality, accessibility and suitability for addressing specific research questions (55).

According to RWD guidelines of regulatory (EMA, FDA) and health technology assessment agencies (i.e., NICE), the initial step of RWD studies should be the definition of a well-defined research question and choosing data that is of sufficient quality and relevant to address the research question, i.e., selecting a 'fit-for-purpose' dataset (49, 324, 325). Generally, data are considered relevant if they cover the population and care setting of interest and include the relevant variables with sufficient detail to answer the research question. Data are of sufficient quality if they are complete (i.e., few missings) and accurate (e.g., dates are correctly reported) (49).

While straightforward in theory, in practice, it may be challenging to evaluate the suitability of a data source, especially as there is typically a lack of access to data before initiating the study. So far, reporting on 'fit-for-purpose' assessments and the challenges faced during this process has been scarce, especially for assessments ending with a negative judgement.

With this brief communication, we aim to contribute to the overall understanding of the opportunities and challenges of using RWD for research by reporting on the three initial steps mentioned in international frameworks for setting up RWD study.

MATERIALS AND METHODS

To date, several frameworks are available to guide researchers in performing RWD-studies (49, 55, 326, 327). Typically, the initial steps in these frameworks follow three main steps: 1) define the research question using the PICOTS approach (i.e., specify the population, intervention, comparator, outcome, timing, and setting); 2) identify candidate data sources potentially relevant to the research question; 3) evaluate the data source relevance to answer the defined question, as well as its quality ('fit-for-purpose'). In this study, we apply these three steps to our study. For step 3, we evaluate 4 key questions both before and after accessing the data to assess the data source's 'fit-for-purpose':

- Does the chosen data source cover the population and care setting of interest?

- Does the dataset contain relevant variables and a sufficient level of detail?
- Does the dataset contain an adequate sample size and follow-up?
- Is the data of sufficient quality to answer the research question?

RESULTS

STEP 1. Definition of the research question

Budget impact and cost-effectiveness analyses aim to improve efficient use of scarce resources but require data on effects and costs. Ideally, they are based on data that represents routine care as provided to a broad patient population. RWD can potentially provide such insights.

At the time of initiating our study, no papers were available describing real-world costs for patients with early-stage breast cancer in the Netherlands (328, 329). Using the PICOTS approach (Table 1), we defined the following question: *What is the hospital resource use and what are the associated hospital costs for adult patients with stage I-III breast cancer (diagnosed between 2015-2020) treated in routine practice in the Netherlands?* The objective of the study was mainly to describe resource use and costs (overall and by subtype and stage), but also to gain insight into the aspects of care that contributed the most to the total costs of treatment.

Table 1. PICOTS

Patients	**Patients aged 18 and older with stage I–III breast cancer**
Intervention	No defined intervention or control measures were imposed
Comparator	
Outcomes	Hospital resource utilization and associated costs, stratified by disease stage, HR/HER2 status, and cost categories
Time	Patients diagnosed between 2015 and 2020, with a follow-up period of 2 years after diagnosis
Setting	Patients who received treatment for breast cancer at a selected academic and non-academic hospital

STEP 2. Identification of data sources

For our specific research question, we required detailed information on resources use of patients with early-stage breast cancer. The literature describes two methods to collect such information: self-reported questionnaires or claims data (330). We did not prefer the former due to the considerable time it requires, the burden it places on both patients and healthcare providers and its susceptibility to recall bias (208). In contrast, claims data are generally available for large groups of patients, have a longitudinal

nature, and are available in a structured format (55). As such, we considered claims data most relevant for our study.

Common sources of claims data used by researchers include claims data from data aggregators (i.e., organizations that collect claims data from multiple organizations), insurance companies or hospitals (55). For our research question, we considered nation-wide data to be ideal, which meant that data from data aggregators would be best. Unfortunately, datasets from these organizations typically do not include information on patient and tumor characteristics, which we considered necessary for our study to select the target population and perform intended analyses.

Claims data from individual hospitals have also been used by researchers to gain insights into resource use and costs of treatment (331). In the Netherlands, hospital claims data offer extensive details about patient care delivered within hospitals. Additionally, there are (under specific conditions) opportunities to combine these data with patient and clinical characteristics, including those from the Netherlands Cancer Registry (NCR). As such, we chose to include two hospitals in our study—an academic and non-academic one—and link these with data from the NCR.

STEP 3. Fit-for-purpose of the data source

Does the chosen data source cover the population and care setting of interest?

The chosen data was limited to hospital resource use, but this was in line with the care setting of interest: our aim was to report on hospital costs. Before initiating our study, we realized that a limitation of including only two institutions in our study was that it does not allow to gain insight into resource use and costs made outside these institutions (i.e., patients visiting multiple hospitals). However, we deemed this trade-off acceptable, as it was (at that time) the only way to generate a dataset with all relevant data aspects (i.e., patient and clinical characteristics from the NCR and resource utilization data from claims).

When we gained access to the data, the limitations of including only two hospitals became evident. Based on data from the NCR, which includes initial treatment and hospital of this treatment, we were able to distinguish "complete" patients (only treated at study hospital) from "incomplete" patients (treated in both study and other hospitals). The number of patients visiting multiple hospitals for different aspects of their breast cancer care appeared beyond our expectations (Figures 1 & 2). Additionally, it became clear from the claims data that certain aspects of care (e.g., radiotherapy, surgery) were delivered only to a limited number of patients in the two hospitals, while one would expect to see them in nearly all patients (Table 2).

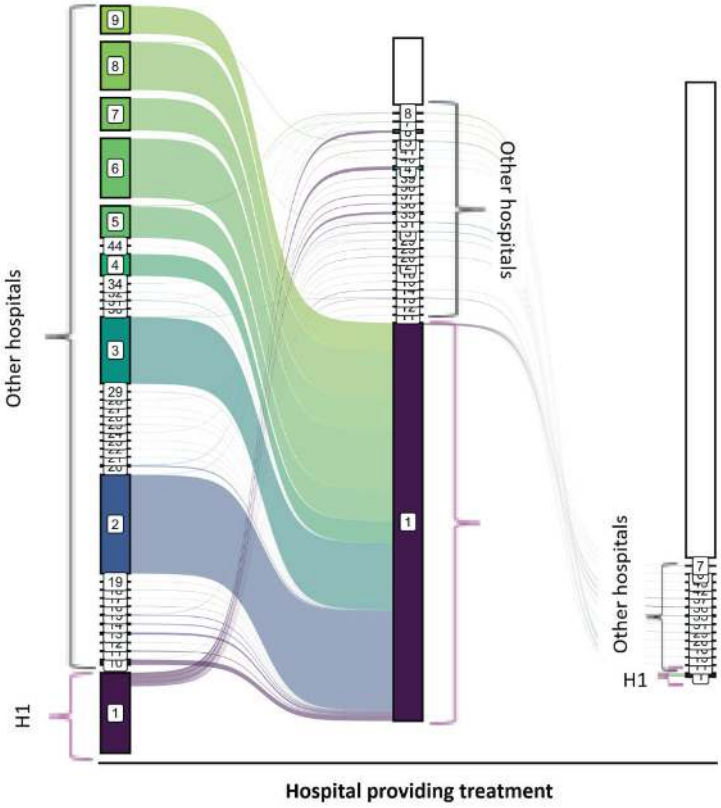

Figure 1. Sankey plot that illustrates the movement of patients across various hospitals (academic hospital)
H1: The academic hospital included in our study

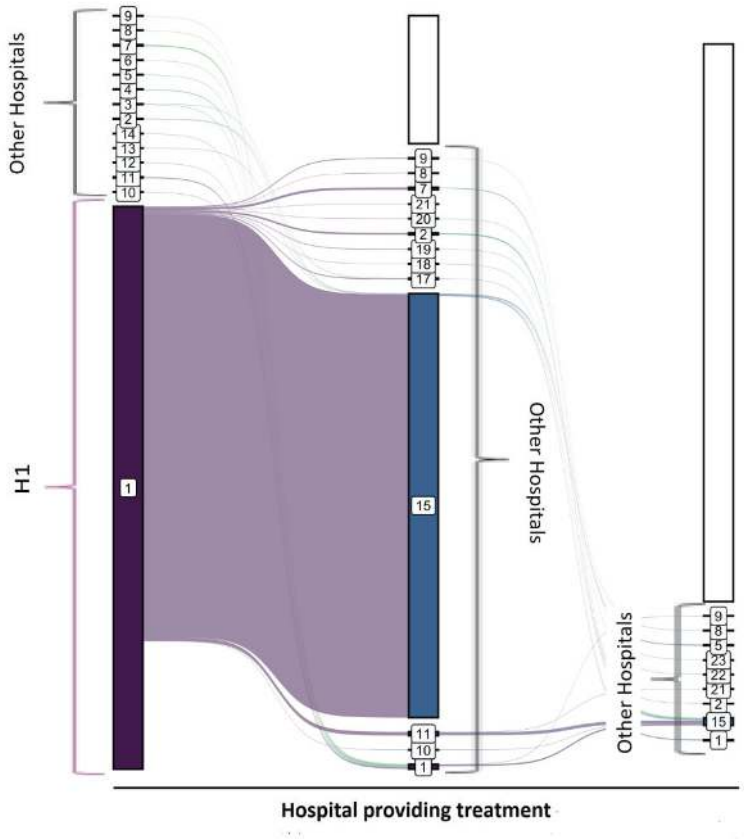

Figure 2. Sankey plot that illustrates the movement of patients across various hospitals (non-academic hospital)

H1: The academic hospital included in our study

Table 2. Number and percentage of patients who received different aspects of care in the two hospitals of our study (related and unrelated to breast cancer)

	Academic hospital (N=4,100)	Non-academic hospital (N=1,983)
Total nr. (%) of patients with:		
Outpatient visit/consult	4.097 (100%)	1,981 (100%)
Imaging	1,358 (33%)	1,969 (99%)
Radiotherapy	3.681 (90%)	5 (0%)
Diagnostic activity	1,285 (31%)	1,752 (88%)
Other	1,853 (45%)	1,738 (88%)
Hospital admission	945 (23%)	1,550 (78%)
Systemic therapy*	515 (13%)	1,547 (78%)
Lab test	1,582 (39%)	1,538 (78%)
Pathological examination	1,098 (27%)	1,459 (74%)
Surgery	741 (18%)	1,426 (72%)
Multidisciplinary consultation	177 (4%)	1,375 (69%)
Anesthesia	741 (18%)	1,366 (69%)
Molecular testing	998 (24%)	1.090 (69%)
Paramedical consultation	371 (9%)	596 (30%)
Genetic testing	954 (23%)	0 (0%)

*Chemotherapy or targeted therapy

Does the dataset contain relevant variables and a sufficient level of detail?

Measuring the costs of a treatment or disease requires the identification of healthcare resource use, as well as the volume of resource use, and related unit costs (171). For our research question, we believed that the hospital claims data would provide a sufficient amount of detail to gain insight into resource use and volumes, as all diagnoses and care activities (*zorgverrichtingen*) must be recorded for the Diagnosis Treatment Combination (DBC) in the Netherlands (332). We realized that corresponding unit costs would not be available for all procedures but considered publicly available tariffs a good alternative (171).

When accessing the data, it became clear that the data indeed contained an enormous amount of information of resource use with over 0.53 million claims in the academic hospital and 0.48 million in the non-academic hospital. For certain categories of care the data also included an incredible level of detail, for instance, there were 256 unique claims for lab tests available in the dataset of the academic hospital (Table 3). Unfortunately, corresponding tariffs were not publicly available for all aspects of care, mainly not for radiotherapy procedures (e.g., a radiation fraction). This makes it difficult to accurately estimate the costs of these aspects of care which we considered an important limitation for our study, as radiotherapy costs account for a significant portion of the total costs of early-stage breast cancer (333).

Table 3. Amount of information in hospital claims datasets (both related and unrelated to breast cancer): total number of claims and number of unique claims per care category*

	Academic hospital (n=4,100)		Non-academic hospital (n=1,983)	
	All claims (608 unique diagnoses)	**Claims related to the diagnosis of breast cancer# (% of all claims)**	**All claims (483 unique diagnoses)**	**Claims related to the diagnosis of breast cancer**
Total nr. of claims	530,622	419,414 (79%)	481,284 (74%)	354,361
Unique claims	1,256	576 (46%)	940	405 (43%)
*Unique claims in the dataset categorized by care**				
Other	434	122	474	99
Lab tests	256	182	171	135
Surgery	161	23	90	29
Systemic therapy	149	72	16	13
Imaging	98	77	87	64
Diagnostic activity	81	28	54	21
Outpatient visit/consult	22	22	14	14
Radiotherapy	11	11	0	0
Pathological examination	10	10	7	6
Paramedical consultation	11	10	11	11
Hospital admission / day care	9	6	6	4
Molecular testing	5	4	4	4
Anesthesia	4	4	4	4
Genetic testing	4	4	0	0
Multidisciplinary consultation	1	1	2	1

*care procedures are categorized; #claims with missing diagnosis code not included

According to previous literature, administrative claims data present opportunities to study treatment-related adverse events because these might lead to additional health care use (e.g., hospitalizations) (334). To include resource use for adverse events, we obtained all diagnoses and care activities recorded in the DBC-system, both those related to the diagnosis breast cancer and those unrelated (e.g., care utilization due to cardiovascular events associated with the use of trastuzumab and anthracyclines are registered under a cardiovascular disease diagnosis code (335)).

When accessing the data, we observed that our decision to include all resource use contributed significantly to the volume of our data. Specifically, >0.1 million (>20% in each hospital) of the claims in our dataset were registered under diagnosis codes other than breast cancer (Table 3). Determining which diagnosis and claims codes are truly associ-

ated with an adverse events of interest is both time-consuming and complex (336). To illustrate, manual checks might be required for all diagnosis codes or the +/- 0.1 million claims registered in the data to assess possible association with breast cancer. Besides, without detailed medical data, one can never be completely sure whether a registered diagnosis or procedure is due to the disease, or would have been there regardless of it (336).

Does the dataset contain an adequate sample size and follow-up?

Another aspect influencing data relevance is the sample size and follow-up duration (49). We did not expect any problems in this respect. Indeed, our sample contained a sufficient number of patients (4,100 and 1,983 patients in the academic and non-academic hospital). Also, the chosen 2-year follow-up period appeared adequate, as the average treatment duration (i.e., time from the first to the last claim) at both hospitals was well below 2 years.

Is the data of sufficient quality?

For our study, we did not comprehensively assess the quality (i.e., completeness and accuracy) of the claims data. However, opportunistic case verification showed logical resource use among patients (e.g., systemic therapy regimes and radiotherapy fractions appeared to be administered in accordance with the recommended protocol). Besides, the completeness of procedures (particularly costly ones) is typically high in claims data (55). Meaning that even if some codes were missed or not accurately reported, we would still obtain a reliable estimate of the total costs.

DISCUSSION

In this article we report on the initial steps of an RWD-study. We showed that the expectations of RWD 'fit-for-purpose' do not always align with reality, and this may influence the value of the data for answering a defined research question. Systematic assessments of RWD suitability, as described in RWD frameworks, help determine the value of RWD for answering a research question.

Our experience generated several insights on topics that are also addressed in current RWD guidance documents (49, 55, 326, 327). We believe our illustrative examples are valuable to report as researchers typically lack access to data prior to initiating a study.

The first insight relates to data coverage. To determine whether a data source is relevant to a research question, it is crucial to carefully assess whether it adequately captures the

relevant patient pathway. For this purpose, it is valuable to request counts and proportions of key data elements (326). For our study, we did not follow this practice, but if we had done this (e.g., by requesting proportions of patients receiving radiotherapy, surgery, and systemic therapy), we would have realized that data from one hospital would not have been sufficient for answering our research question.

Second, it is important to carefully consider what to include in an RWD dataset. Typically, such choices need to be made before researchers have access to the data and before they can fully understand the implications these choices may have on study results. In our study, for example, including diagnoses unrelated to breast cancer to gain insight into adverse effects of treatment gave us additional insights (i.e., health care utilization and costs related to side effects of treatment), but also introduced noise (many claims likely unrelated to breast cancer). This also emphasizes the importance of properly reporting the choices made in the selection of RWD, including both what is included and what is not.

CHAPTER 9

GENERAL DISCUSSION

The more scientists discovered about the biology of cancer, the more they recognize its complexity, its heterogeneity and its malignant potential. Nowadays, scientists believe that it is very unlikely that we will soon see cancer eradicated from our lives entirely. The search for effective treatments endures (2, 337). An additional challenge is the continuously rising number of patients living with cancer and the related costs of treating these patients. Even the wealthiest countries among the world have finite resources available and thus must make choices on what care to deliver and what not (59, 61). However, making such decisions in oncology is complex and involves different stakeholders with different needs for data to support their decision-making process. Clinicians, often as part of a multidisciplinary team, are involved in the decision of which treatment is the best for which patients, regulatory agencies decide which technologies (mainly medicines) become available for use on the market and policy makers aim to optimize health from a societal perspective and decide whether new technologies are reimbursed. Important consideration in these decisions is what the added value of new interventions is and what the incremental costs are. Historically, stakeholders relied mainly on results from randomized controlled trials (RCTs) to gain insights into efficacy, but at present, these are not always conducted or do not provide all answers relevant in oncology (42, 47).

Real world data (RWD), defined as routinely collected data related to a patient's health status and/or care delivery, has gained significant attention to support decision making in recent years (49, 324). In this thesis we explored the potential of different RWD sources to generate insights into cancer care and outcomes. And in this final chapter, we discuss our experiences. We reflect on the core principles of the most recent RWD guidance and tools and discuss opportunities and challenges of implementing these in RWD-studies within the current data landscape. Finally, we propose steps forward that might contribute to improving the evidence base in oncology decision making

REAL-WORLD DATA: ITS POTENTIAL

The landscape of RWD is rich, rapidly growing and includes different types of datasets, collected for various reasons and by various parties (55). Indeed, these data can provide relevant information about large groups of patients, including their diagnosis, characteristics, outcomes, treatment patterns and costs of care. In this thesis, we presented a couple of examples of how RWD can generate insights that are not gained from RCTs.

In **Chapter 4,** we used RWD in addition to results from RCTs to populate our cost-effectiveness analysis (CEA). For CEAs it is often necessary to combine data from different

sources, because information about effects and relevant costs related to the intervention is rarely ever measured in a single study. Additionally, policymakers who decide on reimbursement in a country are interested in costs that represent daily practice in that country. This cannot be obtained from multinational RCTs. In **Chapter 5**, we generate insights into the nationwide utilization of new medicines after regulatory approval is granted. Such information is not collected with RCTs, nor is it their objective. Yet, this information can be relevant for both clinicians and policymakers to identify possible delays in use, disparities in implementation and to assist improvement. Additionally, utilization of available therapies can be used to estimate market size and related costs of future medicines, which is relevant for policymakers responsible for healthcare budgets (338).

In **Chapter 6**, we showed nationwide changes in survival rates over time of patients with different types of (*de novo*) metastatic cancers. Such survival changes in an unselected nationwide cohort provide insights into the benefits of all the research, therapeutic and preventive efforts for the whole population. In contrast, RCTs test a hypothesis about the effects of one single treatment in a selected patient population. A unique point of our study (**Chapter 6**) is also the long follow-up period of about 30 years, owing to the history of the Netherlands Cancer Registry (NCR). In RCTs, the follow-up is usually stopped before information about all clinically relevant outcomes is available. The results of **Chapter 7** are also based on data from the NCR. This study provides insights into the prognosis of different subgroups. Specifically, we showed that patients with *de novo* metastatic breast cancer survive longer compared to patients with recurrent metastatic disease, who have already been treated for cancer previously. Not only is this relevant for clinicians to better inform patients about prognosis, but it is also important for the design and interpretation of RCTs. So far, stratification by *de novo*/recurrent status is not always performed in RCTs evaluating the efficacy of new medicines for metastatic breast cancer (312).

REAL-WORLD DATA: A REALITY CHECK

The studies described in this thesis illustrate the potential for RWD in providing evidence *in addition* to RCTs, but many see a broader potential of RWD (85, 260). Particularly, to generate evidence into the effectiveness of medicines and technologies and to provide more detailed insights into patterns of care to support clinical and policy decision-making (49, 51). To improve RWD studies, several stakeholders have developed (or are developing) guidance and tools for planning, conducting and reporting of RWD studies in general and specifically within oncology. Core principals of these guidance documents include ensuring quality of the data, addressing the key risks of bias by us-

ing appropriate analytical techniques, and generating transparency in each phase of a RWD-study (49, 51, 324). Unfortunately, there are still multiple challenges in meeting these standards, some of which we clearly demonstrated in **Chapter 8** of this thesis. At least three areas require improvement: the data infrastructure, the transparency of RWD and the methodology of analyzing RWD.

Data Infrastructure

According to RWD guidelines, the first step in any RWD study is to identify a '*fit-for-purpose*' dataset that suits the research question (see **Chapter 8**). This means that a clear research question is defined and that the data includes relevant information and is of sufficient quality (49, 55). Typically, RWD studies require at least four types of variables: 1) variables to select eligible patients (e.g., tumor characteristics); 2) variables on patient baseline characteristics (e.g., age, sex, comorbidities, performance status, prior treatments); 3) variables on interventions or exposure (e.g., treatment); and 4) variables on outcome(s) (e.g., recurrent disease, survival, hospitalizations) (339). Using an example breast cancer patient, we illustrate the challenges of generating or finding such a *'fit for purpose'* dataset.

Imagine Yvonne, 56-year-old women who recently participated in the nationwide breast cancer screening program. A couple of days after her mammogram, she receives a call from her general practitioner. The radiologist observed anomalies in the left breast of Yvonne, and she must go to the hospital for further investigations. Less than a week later, Yvonne meets with the oncology nurse, at the breast cancer department of the nearest hospital, who explains the procedures to her. Yvonne's breast is examined, mammography is performed, and a biopsy is taken. A couple of hours later, the pathologist's report is ready, and results are discussed with Yvonne. Unfortunately, it is bad news: Yvonne has breast cancer. Treatment options are presented to her by her surgeon after these were discussed in the multidisciplinary team meeting of the hospital. Together Yvonne and her surgeon decide to start her treatment with 6 cycles of chemotherapy, to reduce the tumor size, followed by complete removal of the breast. Once these treatments are finalized, Yvonne will need to take anti-hormonal therapy for 5 years to reduce the risk of recurrence.

During the care provided to Yvonne, a large amount of data is generated and stored. For instance, breast images, laboratory results, and Yvonne's diagnosis and treatment plan are stored in the hospital's electronic health record (EHR), pathology reports are stored in the pathology system, medication dispensed is registered by the pharmacy and to receive payment for all services provided bills are recorded and submitted to Yvonne's insurance company. This information is registered not only for Yvonne, but for all patients. Together this generates a wealth of data of a broad patient population,

which is potentially relevant for research. Unfortunately, not all data is readily available for use. Besides, data from a single data source is usually not comprehensive enough to find answers to research questions, as we demonstrated in **Chapter 8**. In Yvonne's case, data on her (tumor) characteristics and treatment delivered in that hospital are available in the EHR. Treatment information is also available in billing data, as well as information that can be used as a proxy for outcomes (e.g., time on treatment is sometimes used as a proxy for disease progression). However, both data sources usually lack information on survival and quality of life outcomes. Nor do they cover the entire care pathway (e.g., care delivered outside the hospital, care that is not reimbursed).

The fragmentation of care complicates the use of RWD for research because data from different sources ideally need to be combined to generate a '*fit for purpose*' dataset (55). In the current data infrastructure and legal framework, there are two major hurdles concerning the combination of data. Firstly, patients do not have a unique person identifier within each healthcare facility. This makes it almost impossible to efficiently link data from one organization with data from another. Secondly, the EU wide General Data Protection Regulation (GDPR) states that individual data cannot be used for other purposes than those for which they were collected, especially not when they are identifiable to individual patients (99). Consequently, it is not allowed to combine data from different sources. The issue of the absence of a unique ID is sometimes addressed by using probabilistic record linkage, in which different variables (e.g., postal code, date of birth, sex) are used to match patient data in one dataset with data from the same patient in another. However, missing data or the applied linkage methodology can lead to mismatches (340). Additionally, linkages are often performed by an external organization (e.g., 'third trusted party'), and consequently, the accuracy of matches cannot always be verified, as data are anonymized for privacy reasons (55, 341).

The quality of RWD is also an important consideration when searching for a *fit-for-purpose* dataset. Quality of RWD is determined by its completeness (e.g., few missing values), accuracy (i.e., closely resembling reality) and consistency (no abnormal variability in values over time) of it (342). Due to the retrospective nature of RWD, some quality aspects are difficult to check. For instance, when using billing data, one cannot know whether a treatment was not given or whether it was not properly billed or maybe provided by another organization (55). Linkage to a gold-standard data set is proposed as a solution, but this is typically not allowed (55). Moreover, even if data linkage or other types of data validation are possible, the process of comparing data items is very time-intensive (342). Imagine having to manually check whether systemic treatment, usually consisting of several cycles of treatment and combinations with different start and stop dates, is appropriately billed on the right date. Or having to check whether all patients

in a population-based study indeed experienced an outcome that was expected based on a proxy variable in an RWD-set (e.g., cancer recurrence).

In the current data landscape, the possibilities for finding an actual '*fit for purpose*' data source are limited (**Chapter 8**). Yet, there is work in progress. The European Commission recognizes the obstacles to using routinely collected health data for research and has set up a large project to generate a European Health Data Space (343). Also in the Netherlands it is acknowledged that the current data infrastructure needs improvement to make it possible to generate relevant insights in to real-world use and (cost-) effectiveness of new medicines (344). As such, the government has initiated programs to prepare for the European Health Data Space and to shape its implementation. Specifically relevant for oncology is the recent agreement between IKNL (organization that hosts the Netherlands Cancer Registry) and Health-RI (an organization who is closely involved in the development of the European Health Data Space), which was signed to improve data accessibility for cancer research (343, 345, 346). Important objectives of all these initiatives are to provide a set-up for the secondary use and re-use of health data for research, policy making and regulatory activities (343). By emphasizing the FAIR data principles—Findable, Accessible, Interoperable, and Reusable—they aim to ensure that data adheres to governance, quality, and accessibility standards (347). These are important steps forward toward improving the potential of RWD.

Transparency

Choices made during study planning, design, cleaning, analysis and reporting can impact study outcomes and sometimes even lead to biased results. Because these concerns are well known, different tools, checklists and frameworks have been developed previously, all with the aim of improving the practice, reporting and interpretation of RWD studies. Almost a decade ago, the ISPOR (International Society for Pharmacoeconomics and Outcomes Research), GRACE (Good Research for Comparative Effectiveness) and NICE (The National Institute for Health and Care Excellence) already published guidance on how to best design RWD studies of comparative effectiveness (110, 121, 124). Key themes covered by these frameworks include selection and confounding biases, cherry-picking in RWD, uncertainty, and the definition of exposure and outcomes. Unfortunately, we observed (**Chapter 3)** that the implementation of recommendations in practice is hampered. Most RWD studies did not even report the most crucial study elements, such as choice of confounders and definition of exposure and outcome variables. This finding is broadly shared in the literature and hampers trust in RWD studies so far (260, 348).

Ever since our publication in **Chapter 3**, many more RWD tools became available, all with the objective to increasing transparency and improving the credibility of RWD

studies that deal with questions of causal inference for decision-making. To name a few: STaRT RWE (Structured Template for planning and reporting on the implementation of Real World Evidence studies), HARPER (The HARmonized Protocol Template to Enhance Reproducibility in RWE), SPIFD 1 (Structured Process to Identify Fit-For-Purpose Data), SPACE (Structured Preapproval and Postapproval Comparative Study Design Framework to Generate Valid and Transparent Real World Evidence), SPIFD 2 (connects SPIFD 1, SPACE, HARPER and STaRT RWE) and TARGET (TrAnsparent ReportinG of observational studies Emulating a Target trial) (326, 327, 349, 350, 351). These tools do complement the older once on some aspects (e.g., more detailed, focuses on reproducibility of RWD studies, more comprehensive guidance on how to deal with potential biases, emphasizes FAIR principles), but the main challenge remains to encourage their widespread implementation (350). In the context of the plans to generate a national or even European Data Space, there may be opportunities to do so. Submitting a pre-defined study protocol and using a structured way of reporting with one or more of the available tools must be made mandatory for researchers who want access to data. Relevant to note in this sense is that generating a high level of transparency in RWD is a resource-intensive process which includes the development of a detailed study protocol, a data 'fit-for-purpose' assessment (e.g., **Chapter 8**), a clear description of what is included and what is not, a comprehensive definition of variables and confounders (including their accuracy), the target trial emulation approach, and a detailed description of how was dealt with concerns related to data quality and validity, and so on (327).

Study Design and Methodology

Despite the availability of RWD guidelines, the question of when the design and methodology of RWD studies are *sufficient* to generate trustworthy results remains unanswered. This is especially a concern in RWD studies that aim to draw conclusions about causal inference. As mentioned previously, one of the most important concerns in such studies is the comparability of treatment groups in the absence of randomization (57). Design and statistical methods (e.g., propensity score matching, multivariate regression modelling, instrumental variable analyses) can partially control for this but only if there is appropriate understanding of what factors influence disease prognosis. Additionally, the relevant variables should also be available in the dataset. Usually, the most important aspects are either very difficult to quantify or simply not available in RWD which were initially not collected for research purposes (352). For instance, the reasons why physicians and their patients choose a particular treatment over another are rarely available in RWD, as well as is information on lifestyle (e.g., smoking, BMI), comorbidities, and performance status. The latter aspects can influence the choice of treatment as well as its effectiveness. As such, residual confounding will remain.

Immortal time bias (ITB) is a type of bias that is a significant concern in RWD studies comparing treatment effects. This bias can occur because treatment exposure is based on care that has already been delivered during follow-up (135, 136, 353). Previously, ITB was observed, for instance, in studies evaluating the effectiveness of resection of the primary tumor on survival in patients with *de novo* metastatic breast cancer. In some cases, the definition of exposure was chosen as mastectomy at any time during follow-up. The surgery appeared to result in a better survival rate compared to no surgery, but what really happened was that the former group had a survival advantage: they could not have died in the period from diagnosis to mastectomy (354, 355). While there are methods to control for ITB (e.g., landmark approach, time varying covariates, target trial emulation) and the problem was already recognized over a decade ago, there are still examples of studies that use approaches susceptible to this type of bias (**Chapter 3**) (356, 357, 358). Concerning in this respect is also the trend toward approving oncology medicines based on a single-arm trial with an external historical control (357). This approach leads to study designs that can be affected by ITB, especially when they involve indications that require failure of earlier lines of therapy, which is common in oncology (357).

Analytical methods to reduce the mentioned biases are being improved continuously. Most recently an approach called 'target trial emulation' has gained attention (359). In this approach, RWD is used to simulate a hypothetical RCT. While this approach certainly improves the quality and transparency of choices, it is very difficult to prove that such methods truly eliminate biases. Some authors have compared results of well-conducted RCTs with studies using health care databases with the same clinical question to understand the validity of RWD studies (360, 361). These studies showed mixed results, with some showing low and others showing high concordance. Comparability of treatment effects were typically better in studies that used the target trial emulation approach. However, this does not necessarily mean that these studies provide valid causal inference. There may be various other explanations of why similar treatment effects are observed, for instance, chance or decisions made during data analyses (57).

THE OVERARCHING CHALLENGE WITH DATA IN ONCOLOGY

Data in oncology can only contribute to better patient care and outcomes if they reduce uncertainty regarding the most pressing clinical, policy, and regulatory questions. Data from RCTs and RWD can both provide relevant information. In theory, they can perfectly complement each other, with RCTs providing unbiased estimates of the efficacy of new oncology products in controlled settings (46). And, as illustrated in the chapters of this

thesis (**Chapters 4-7**), RWD generates insights into the patterns of care, characteristics of patient groups, long term safety, and effects of treating cancer in broader patient groups. In practice, there are limitations to the optimal use of both data sources. Results from RCTs in oncology are not generalizable to the general population due to their strict inclusion criteria and because the control arm may not reflect the standard of care (46). Studies based on RWD, on the other hand, may rely on questionable data and/or are of uncertain methodological quality (260).

The challenges of RCTs and RWD mentioned above are true concerns, but most of these are not genuinely connected to the data and approach per se. If RCTs are designed appropriately, they do provide reliable and relevant estimates of effects of treatment (362). Similarly, well-designed RWD studies, that use high-quality data and appropriate methodology can provide valuable insights into cancer care and outcomes that cannot be gained from RCTs (55, 363). Unfortunately, within the current ecosystem of evidence generation in oncology, contradictory interests and goals among involved stakeholders exist. Society demands evidence for efficacy in oncology to be provided by commercial companies, which means that they take the lead in financing and initiating clinical studies. An important objective of these companies is to bring products to the market quickly, which appears influence the way studies are designed, performed and how results are communicated (364). For instance, industry-funded trials in oncology increasingly use PFS, a surrogate outcome, as the primary outcome. Results of such trials more often yield positive results, while it is (often) uncertain whether this leads to meaningful benefit to patients (44, 365). Additionally, industry funded trials more often utilize suboptimal control arms compared to publicly funded trials, which impacts the observed effect size (366). Finally, the value attached to quality-of-life outcomes might not reflect patient values. For example, increased toxicity or inferior quality of life have been observed and were framed as being 'acceptable' (367, 368). RWD-studies are then seen as a solution to address remaining uncertainties (regarding benefits versus harms) (369). However, within fragmented healthcare systems, existing datasets are often insufficient to capture the relevant variables (i.e., exposure, adverse effects and outcomes), thereby limiting their evidentiary value (260, 370).

While the above problems are long-standing, there are a couple of reasons for urgency. Firstly, because there appears to be a paradigm shift in the process of medicine approval. Whereas historically, robust evidence was required to obtain marketing authorization and lower standards were only accepted in exceptional occasions (365). Nowadays, the acceptance of evidence based on less rigorous designs, surrogate endpoints, and in some cases non-randomized (single-arm) studies (e.g., gene and cell therapies), seems to be becoming more common in oncology (58, 75, 365, 371, 372). Secondly, in the era

of big data and artificial intelligence (AI), health data is available across various settings. While the use of modern tools holds promises for improving RWD studies (e.g., replacing manual data collection from free-text notes), their use also raises additional concerns about reliability and validity because of reduced transparency (i.e., the "black box" mechanism of AI) (373, 374).

TOWARDS MEANINGFUL DATA IN ONCOLOGY

The question that remains is how we can move forward and generate data that provides the best possible evidence for all stakeholders: data that contributes to reducing uncertainties in oncology and provides reliable insights into what works and what does not, for which patients and at what costs. And most importantly, data that truly help improve care and outcomes for patients with cancer. Many directions have been proposed in the literature and based on our experiences with collecting and analyzing data for this thesis, we agree with the following.

The first is a direction for regulatory agencies. Stricter thresholds for approval should be set; this includes mandating registration studies that are designed, analyzed and reported in the most optimal way, as well as the publication of mature survival data (362, 365). In the last decades, regulatory agencies of medicines have relaxed evidentiary standards to expedite access. A displeasing side effect of this is that there are high levels of uncertainty of the clinical benefit and the economic value of medicines at the time of approval. While patients do deserve timely access to treatments, they also deserve that these treatments are supported by strong evidence of effectiveness. Moreover, society deserves to pay only for treatments that perform as promised (375). RCT design should improve by including an appropriate comparator arm that reflects the standard of care, measuring outcomes that matter to patients (in the correct way) and including all patients who are fit to receive the treatment under study and might benefit from it (44, 58, 362). RWD has potential to play a role in this, for instance by providing insights into the current standard of care in the target population, gaining better understanding of diseases including the prognosis of subgroups, or by supporting recruitment of trial participants (49, 376).

The second direction is for organizations and researchers involved in collecting, analyzing and reporting studies based on RWD. They should acknowledge that RWD can provide valuable insights but only if it is used, analyzed and reported in an appropriate way. This requires knowledge of the database and its origin (as shown in **Chapter 8**), collaboration, study planning, proper training of researchers and careful consideration

of potential biases (49, 55, 349, 350). In the era of digitalization and artificial intelligence (AI), it may be tempting to use all data for research. But the truth is that existing datasets vary substantially in terms of their completeness, accuracy and quality. Careful consideration of the suitability of RWD for specific research questions must be an part of all RWD studies (55).

Special care should be taken when attempting to answer causal questions with RWD. For this purpose, innovative research designs should be developed and improved. Promising initiatives in this respect are pragmatic trials (e.g., registry based RCTs (RRCTs) and trials within cohorts (TwiCs)) (377, 378, 379, 380). The idea behind these approaches is that they integrate aspects of RCTs (i.e., randomization) with aspects of RWD (e.g., data collection, patient recruitment) and exploit the advantages of both. Random treatment assignment is used to ensure comparability of patient groups, and the use of RWD and its infrastructure makes it possible to reduce administrative burden, include a broad patient population and measure long term outcomes (206, 377, 381). So far, pragmatic trials within oncology have been performed successfully in various clinical settings (e.g., radiotherapy, exercise therapy), but their utility to evaluate medicine effectiveness is not yet well understood (382). Further investigation is needed to better understand which research questions pragmatic trials can provide valid answers and to, and which they cannot.

The final direction is for the government and philanthropic foundations. Structural funding should be made available to set up and maintain a better data infrastructure and generate high-quality datasets that capture the entire patient trajectory (**Chapter 8**). The initiative of the European Commission to set up a European Health Data Space, along with related initiatives specifically for oncology at the national level, are important steps forward in this respect (343, 346). However, funding is also needed to improve the quality of RWD. This includes validation studies that compare data items against a chosen standard to identify possible misclassification in exposure and outcome variables, as well as covariates. Such validation studies should be an integral part of RWD-based research in order to improve confidence in the use of RWD for decision-making (383). Finally, funding is needed to better understand the validity of results of RWD studies attempting to estimate treatment effects. In other words, how can researchers know that their design is sufficient to estimate a 'true' effect. In this sense, it is not only relevant to provide insights into what can be done with RWD but also, what cannot be done, such as recently highlighted by a research team from Havard-MIT. They suggest that it is not possible to reliably estimate the effect of colonoscopy screening on mortality using RWD because of unmeasured confounding, even when employing the most up-to date

methodology (384). This is valuable information for researchers and decision-makers, but such studies will likely not be performed without incentives to do so.

CONCLUDING REMARKS

In this thesis, we explored the potential of different RWD sources to support decision making in oncology. We showed that RWD can provide valuable insights into treatment utilization, long-term outcomes and patient and disease characteristics. But we also identified complex barriers in the current data infrastructure and legal framework that hamper the optimal use of RWD. Promising initiatives to address these challenges are on their way, though they are still at early stages of development. Meanwhile, it is the responsibility of stakeholders working with RWD or using RWD in decision-making, to use RWD wisely. This includes careful consideration of the utility of RWD, identification of possible biases, transparent reporting, and critical appraisal of results.

CHAPTER 10

OTHER

REFERENCES

1. Periodical. Citizens Committee for the Conquest of Cancer. "Mr. Nixon: You Can Cure Cancer." The Washington Post. 1969 Dec 9; Sect. C9. Article. 1. Available from: https://profiles.nlm.nih.gov/101584665X20.
2. Mukherjee S, The Emperor of All Maladies. 4th ed. London: HarperCollins Publisher. 2011.
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-49.
4. Einhorn LH, Donohue J. Cis-diamminedichloroplatinum, vinblastine, and bleomycin combination chemotherapy in disseminated testicular cancer. *Ann Intern Med.* 1977;87(3):293-8..
5. Bonadonna G, Zucali R, Monfardini S, de Lena M, Uslenghi C.. Combination chemotherapy of Hodgkin's disease with adriamycin, bleomycin, vinblastine, and imidazole carboxamide versus MOPP. *Cancer*. 1975;36(1):252-9.
6. Frei E, Holland JF, Schneiderman MA, Pinkel D, Selkirk G, Freireich EJ et al. A comparative study of two regimens of combination chemotherapy in acute leukemia. *Blood.* 1958;13(12):1126-48
7. DeVita Jr VT, Chu E. A history of cancer chemotherapy. *Cancer res*. 2008;68(21):8643-53.
8. Hanna N, Einhorn LH. Testicular cancer: a reflection on 50 years of discovery. *J Clin Oncol.* 2014;32(28):3085-92.
9. Kenneth ME. Letter from Kenneth M. Endicott to Mary Lasker. 1962. Available from: http://resource.nlm.nih.gov/101584665X137.
10. Chabner BA, Roberts Jr TG. Chemotherapy and the war on cancer. *Nat Rev Cancer*. 2005;5(1):65-72.
11. Fisher B, Jeong J-H, Anderson S, et al. Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation. *N Engl J Med.* 2002;347(8):567-75.
12. Fisher B. Laboratory and clinical research in breast cancer—a personal adventure: the David A. Karnofsky memorial lecture. *Cancer res.* 1980;40(11):3863-74.
13. Fidler IJ, Poste G. The "seed and soil" hypothesis revisited. *Lancet oncol*. 2008;9(8):808.
14. Veronesi U, Banfi A, Saccozzi R, Salvadori B, Zucali R, Uslenghi C, et al. Conservative treatment of breast cancer: a trial in progress at the Cancer Institute of Milan. *Cancer*. 1977;39(6):2822-6.
15. Veronesi U, Saccozzi R, Del Vecchio M, Banfi A, Clemente C, De Lena M, et al. Comparing radical mastectomy with quadrantectomy, axillary dissection, and radiotherapy in patients with small cancers of the breast. *N Engl J Med.* 1981;305(1):6-11.
16. Patriarca C, Pini GM, Conti G. Invasion and metastasis: a historical perspective. *Pathologica*. 2020;112(4):229.
17. Hajdu SI. A note from history: landmarks in history of cancer, part 1. *Cancer*. 2011;117(5):1097-102.
18. Greene GL. In Memoriam: Elwood Jensen (1920–2012). *Endocrine Rev*. 2013;34(6):761-3.
19. Benadada F, Saad F, Delouya G, Taussky D., et al. A historical review of the Nobel laureate's pioneering discoveries. *Cancer*. 2024;130(7):1019-24.
20. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med*. 2001;344(11):783-92.
21. Kalia M. Personalized oncology: Recent advances and future challenges. *Metabolism*. 2013;62:S11-S4.

22. Sharma MR, Schilsky RL. Role of randomized phase III trials in an era of effective targeted therapies. *Nat Rev Clin Oncol*. 2012;9(4):208-14.
23. Lackner MR, Wilson TR, Settleman J. Mechanisms of acquired resistance to targeted cancer therapies. *Future Oncol*. 2012;8(8):999-1014.
24. Curigliano G, Burstein HJ, Winer EP, Gnant M, Dubsky P, Loibl S, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol.* 2017;28(8):1700-12.
25. Arpino G, Generali D, Sapino A, Del Matro L, Frassoldati A, de Laurentis M, et al. Gene expression profiling in breast cancer: a clinical perspective. *Breast.* 2013;22(2):109-20.
26. Robert C. A decade of immune-checkpoint inhibitors in cancer therapy. *Nat Commun*. 2020;11(1):3801.
27. Food and Drug Administration (FDA). FDA approval brings first gene therapy to the United States. 2017. Available from: https://www.fda.gov/news-events/press-announcements/fda-approval-brings-first-gene-therapy-united-states.
28. Wolchok JD, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Long-term outcomes with nivolumab plus ipilimumab or nivolumab alone versus ipilimumab in patients with advanced melanoma. *J Clin Oncol*. 2022;40(2):127.
29. McDermott D, Lebbé C, Hodi FS, Maio M, Weber JS, Wolchok JD, et al. Durable benefit and the potential for long-term survival with immunotherapy in advanced melanoma. *Cancer Treatment Rev*. 2014;40(9):1056-64.
30. Meet Emily Withehead. Available from: https://acgtfoundation.org/for-patients/patient-stories/emily-whitehead/.
31. Doll R et al. Smoking and carcinoma of the lung; preliminary report. *Br Med J*. 1950;2.
32. Vilos GA. The history of the Papanicolaou smear and the odyssey of George and Andromache Papanicolaou. *Obstet Gynecol.* 1998;91(3):479-83.
33. Bretthauer M, Wieszczy P, Løberg M, Kaminski MF, Werner TF, Helsingen LM, et al. Estimated lifetime gained with cancer screening tests: a meta-analysis of randomized clinical trials. *JAMA Inter Med*. 2023;183(11):1196-203.
34. Reitsma MB, Kendrick PJ, Ababneh E, Abbafati C, Abbasi-Kangevari M, Abdoli A, et al. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019. *Lancet.* 2021;397(10292):2337-60.
35. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard—lessons from the history of RCTs. *Mass Medical Soc*. 2016:2175-81.
36. Crofton J, Mitchison D. Streptomycin resistance in pulmonary tuberculosis. *Br Med J.* 1948;2(4588):1009.
37. Sackett DL, Rosenberg WM, Gray JM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996:71-2.
38. Jönsson B, Ramsey S, Wilking N. Cost effectiveness in practice and its effect on clinical outcomes. *J Cancer Policy.* 2014;2(1):12-21.
39. European Medicines Agency (EMA). Clinical trials in human medicines. Available from: https://www.ema.europa.eu/en/human-regulatory-overview/research-development/clinical-trials-human-medicines.
40. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: results of a systematic review and expert consultation across eight European countries. The *Eur J Health Econ.* 2018;19:123-52.

41. Cherny N, Dafni U, Bogaerts J, Latino N, Pentheroudakis G, Douillard J-Y, et al. ESMO-magnitude of clinical benefit scale version 1.1. *Ann Oncol*. 2017;28(10):2340-66.
42. Hulstaert F, Pouppez C, Primus-de Jong C, Harkin K & Neyt M. Evidence Gaps for Drugs and Medical Devices at Market entry in Europe and potential solutions 2021. Health Services Research (HSR). Brussels. Belgian Health Care Knowledge Centre (KCE). 2021. KCE Reports 347. DOI: 10.57598/R347C.
43. Booth C, Tannock I. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *B J Cancer.* 2014;110(3):551-5.
44. Del Paggio JC, Berry JS, Hopman WM, Eisenhauer EA, Prasad V, Gyawali B, et al. Evolution of the randomized clinical trial in the era of precision oncology. *JAMA oncol*. 2021;7(5):728-34.
45. Vera-Badillo F, Shapiro R, Ocana A, Amir E, Tannock IF. Bias in reporting of end points of efficacy and toxicity in randomized, clinical trials for women with breast cancer. *Ann Oncol*. 2013;24(5):1238-44.
46. Tannock IF, Amir E, Booth CM, Niraula S, Ocana A, Seruga B, et al. Relevance of randomised controlled trials in oncology. *Lancet Oncol*. 2016;17(12):e560-e7.
47. Hulstaert F, Pouppez C, Primus-de Jong C, Harkin K, Neyt M. Gaps in the evidence underpinning high-risk medical devices in Europe at market entry, and potential solutions. *Orphanet J Rare Dis.* 2023;18(1):212.
48. Tarricone R, Torbica A, Drummond M. Challenges in the assessment of medical devices: the MedtecHTA project. *Health Econ*. 2017;26:5-12.
49. National Institute for Health and Care Excellence (NICE). NICE real-world evidence framework. June 23, 2022. Available from: https://www.nice.org.uk/corporate/ecd9/chapter/overview.
50. Castelo-Branco L, Pellat A, Martins-Branco D, Valachis A, Derksen J, Suijkerbuijk K, et al. ESMO guidance for reporting oncology real-world evidence (GROW). *Ann Oncol*. 2023;1:100003.
51. Arlett P, Kjær J, Broich K, Cooke E. Real-world evidence in EU medicines regulation: enabling use and establishing value. *Clin Pharmacol Ther.* 2022;111(1):21.
52. Bailar III JC, Smith EM. Progress against cancer? *N Engl J Med*. 1986;314(19):1226-32.
53. Bailard JC, Smith EM. Have we reduced the risk of getting cancer or of dying from cancer? An update. *Med Oncol Tumor Pharmacother.* 1987;4(3):193-8.
54. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-world evidence—what is it and what can it tell us. *N Engl J Med.* 2016;375(23):2293-7.
55. Penberthy LT, Rivera DR, Lund JL, Bruno MA, Meyer AM. An overview of real-world data sources for oncology and considerations for research. *CA: A Cancer J Clin.* 2022;72(3):287-300.
56. Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. . *Nat Rev Clin Oncol.* 2019;16(5):312-25.
57. Groenwold RH. Trial emulation and real-world evidence. *JAMA Netw Open*. 2021;4(3):e213845-e.
58. Wieseler B, Neyt M, Kaiser T, Hulstaert F, Windeler J. Replacing RCTs with real world data for regulatory decision making: a self-fulfilling prophecy? BMJ. 2023;380.
59. Jönsson B, Hofmarcher T, Lindgren P, Wilking N. The cost and burden of cancer in the European Union 1995-2014. *Eu J Cancer*. 2016;66:162-70.
60. Hofmarcher T, Lindgren P, Wilking N, Jönsson B. The cost of cancer in Europe 2018. *Eu J Cancer.* 2020;129:41-9.
61. Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol NJ, Amir E, et al. Delivering affordable cancer care in high-income countries. *Lancet Oncol*. 2011;12(10):933-80.
62. Savage P. Development and economic trends in cancer therapeutic drugs in the UK from 1955 to 2009. *J Oncol Pharm Prac*. 2012;18(1):52-6.

63. Lythgoe MP, Krell J, Mahmoud S, Mills EC, Vasudevan A, Savage P. Development and economic trends in anticancer drugs licensed in the UK from 2015 to 2019. *Drug Discov Today.* 2021;26(2):301-7.
64. National Health Care Institute (Zorginstituut Nederland). Horizonscan Medicines – Oncology. Available from: https://www.horizonscangeneesmiddelen.nl/geneesmiddelen.
65. 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.
66. Jenei K, Raymakers AJ, Bayle A, Berger-Thürmel K, Cherla A, Honda K, et al. Health technology assessment for cancer medicines across the G7 countries and Oceania: an international, cross-sectional study. Lancet Oncol. 2023;24(6):624-35.
67. Enzing JJ, Knies S, Boer B, Brouwer WB. Broadening the application of health technology assessment in the Netherlands: a worthwhile destination but not an easy ride? *Health Econ Policy Law*. 2021;16(4):440-56.
68. Strand GC, Bonander C, Jakobsson N, Johansson N, Svensson M. Assessment of the clinical and cost-effectiveness evidence in the reimbursement decisions of new cancer drugs. *ESMO open*. 2022;7(5):100569.
69. Strom BL. What is Pharmacoepidemiology? In: Strom BL, Kimmel SE, Hennessy S, eds. Textbook of pharmacoepidemiology. 2nd ed. Oxford, UK: John Wiley & Sons Ltd; 2013.
70. European Medicines Agency (EMA). Marketing Authorization. Available from: https://www.ema.europa.eu/en/human-regulatory-overview/marketing-authorisation.
71. Allan C, Joyce TJ, Pollock AM. Europe's new device regulations fail to protect the public. *BMJ*. 2018.
72. Jardim DL, Schwaederle M, Hong DS, Kurzrock R. An appraisal of drug development timelines in the era of precision oncology. *Oncotarget*. 2016;7(33):53037.
73. European Medicines Agency (EMA). Accelerated Approval. Available from: https://www.ema.europa.eu/en/human-regulatory-overview/marketing-authorisation/accelerated-assessment.
74. Cox EM, Edmund AV, Kratz E, Lockwood SH, Shankar A. Regulatory affairs 101: Introduction to expedited regulatory pathways. *Clin Transl Sci.* 2020;13(3):451-61.
75. Mulder J, Teerenstra S, Van Hennik P, Pasmooij A, Stoyanova-Beninska V, Voest E, et al. Single-arm trials supporting the approval of anticancer medicinal products in the European Union: contextualization of trial results and observed clinical benefit. *ESMO open*. 2023;8(2):101209.
76. Thomson S, Witzke N, Gyawali B, Santos SD, Udayakumar S, Cardone C, et al. Assessing the benefit of cancer drugs approved by the European Medicines Agency using the European Society for Medical Oncology Magnitude of Clinical Benefit Scale over time. *Eur J Cancer.* 2021;150:203-10.
77. Hilal T, Gonzalez-Velez M, Prasad V. Limitations in clinical trials leading to anticancer drug approvals by the US Food and Drug Administration. *JAMA Internal Med*. 2020;180(8):1108-15.
78. Woolacott N, Corbett M, Jones-Diette J, Hodgson R. Methodological challenges for the evaluation of clinical effectiveness in the context of accelerated regulatory approval: an overview. *J Clin Epidemiol.* 2017;90:108-18.
79. Bierbaum M, Rapport F, Arnolda G, Nic Giolla Easpaig B, Lamprell K, Hutchinson K, et al. Clinicians' attitudes and perceived barriers and facilitators to cancer treatment clinical practice guideline adherence: a systematic review of qualitative and quantitative literature. *Implementation Sci.* 2020;15:1-24.
80. Kempf E, Bogaerts J, Lacombe D, Liu L. 'Mind the gap'between the development of therapeutic innovations and the clinical practice in oncology: a proposal of the European Organisation for

Research and Treatment of Cancer (EORTC) to optimise cancer clinical research. *Eur J Cancer*. 2017;86:143-9.

81. Nederlandse Vereniging voor Medische Oncologie_(NVMO). PASKWIL criteria 2023. Available from: https://www.nvmo.org/over-de-adviezen/.
82. Leeneman B, Xander NS, Fiets WE, de Jong WK, Uyl NE, Wymenga AM, et al. Assessing the clinical benefit of systemic anti-cancer treatments in the Netherlands: The impact of different thresholds for effectiveness. *Eur J Cancer*. 2024;202:114002.
83. Bakker E, Plueschke K, Jonker CJ, Kurz X, Starokozhko V, Mol PG. Contribution of real-world evidence in European medicines Agency's regulatory decision making. *Clin Pharmacol Ther*. 2023;113(1):135-51.
84. Harricharan S, Curran E, Lin HM, Walton L, Gurjar K, Nguyen K, et al. Real-world evidence in lung and hematologic oncology health technology appraisals: a review of six assessment agencies. *Future Oncol*. 2023;19(8):603-16.
85. Pellat A, Grinda T, Prelaj A, Cresta P, Valachis A, Zerdes I, et al. 1689O Comprehensive mapping review of real-world evidence publications focusing on targeted therapies in solid tumors: A collaborative work from ESMO real-world data and Digital Health Working Group. *Ann Oncol*. 2023;34:S925.
86. White R. Building trust in real world evidence (RWE): moving transparency in RWE towards the randomized controlled trial standard. *Curr Med Res Opin*. 2023;39(12):1737-41.
87. European Medicines Agency_(EMA). Real-world evidence framework to support EU regulatory decision making: Report on the experience gained with regulator-led studies from September 2021 to February 2023. 2023. Available from: https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained-regulator-led-studies-september-2021-february-2023_en.pdf.
88. 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.
89. Schouten LJ, Höppener P, van den Brandt PA, Knottnerus JA, Jager JJ. Completeness of cancer registration in Limburg, The Netherlands. *Int J Epidemiol*. 1993;22(3):369-76.
90. Integraal Kankercentrum Nederland (IKNL). Toelichting beschikbaarheid overlevingsgegevens 2022. Available from: https://iknl.nl/nkr/registratie/toelichting-beschikbaarheid-overlevingsgegevens.
91. PwC Strategy&. De maatschappelijke impact van Integraal Kankercentrum Nederland. May 23, 2023. Accessed April 3, 2024. https://iknl.nl/getmedia/f451b497-9055-48ff-b4f9-a18743a9dd7f/230523-De-maatschappelijke-impact-van-IKNL_Samenvatting.pdf
92. International Agency for Research on Cancer (IARC). Standards and Guidelines for Cancer Registration in Europe. 2003. Available from: https://publications.iarc.fr/Book-And-Report-Series/Iarc-Technical-Publications/Standards-And-Guidelines-For-Cancer-Registration-In-Europe-2003
93. Brierley JD, Gospordarowicz MK & Witteking C. TNM Classification of Malignant Tumours, 8th edition: Wiley Publications; 2016.
94. Fritz A PC, Jack A, Shanmugaratnam K, Sobbin L, Parklin DM, Whelan S. ICD-0 International Classification of Diseases for Oncology 2013.
95. Aben KK, Verhoeven R, de Vries A. De Nederlandse Kankerregistratie; een schat aan informatie. *Tijdschr Voor Urol*. 2016;7:40-6.
96. NetherlandsComprehensive Cancer Organization (IKNL). Datacatalogus. Available from: https://iknl.nl/IKNL-datacatalogus.

97. Integraal Kankercentrum Nederland (IKNL). Gegevensaanvraag 2024. Available from: https://iknl.nl/nkr/cijfers-op-maat/gegevensaanvraag.
98. Integraal Kankercentrum Nederland (IKNL). Privacyverklaring. Available from: https://iknl.nl/privacyverklaring.
99. European Commission (EC). What is the GDPR. Available from: https://gdpr.eu/what-is-gdpr/.
100. RIVM. Handreiking ontsluiting patiëntgegevens voor wetenschappelijk onderzoek. 2019.
101. Garrison Jr LP, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR real-world data task force report. *Value Health*. 2007;10(5):326-35.
102. Dutch Hospital Data (DHD). About the DHD. Available from: https://www.dhd.nl/over-ons.
103. Dutch Hospital Data (DHD). Request of hospital data. Available from: https://www.dhd.nl/
104. Privacy Officer at Erasmus Medical Centre (EMC). Personal Communication: Meeting about privacy aspects of using claims data for research. 2023.
105. Makady A, van Veelen A, Jonsson P, Moseley O, D'Andon A, de Boer A, et al. Using real-world data in health technology assessment (HTA) practice: a comparative study of five HTA agencies. *Pharmacoeconomics*. 2018;36:359-68.
106. Marshall DA, Hux M. Design and analysis issues for economic analysis alongside clinical trials. *Med care.* 2009;47:S14-S20.
107. Del Paggio J, Azariah B, Sullivan R, Hopman W, James F, Roshni S, et al. Do contemporary randomized controlled trials meet ESMO thresholds for meaningful clinical benefit? *Ann Oncol*. 2017;28(1):157-62.
108. Booth CM, Cescon DW, Wang L, Tannock IF, Krzyzanowska MK. Evolution of the randomized controlled trial in oncology over three decades. *J Clin Oncol*. 2008;26(33):5458-64.
109. Moen F, Svensson J, Carlsson KS. Assessing the value of cancer treatments from real world data—Issues, empirical examples and lessons learnt. *J Cancer Policy*. 2017;11:32-7.
110. Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part I. *Value Health*. 2009;12(8):1044-52.
111. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. JNCI: *J Natl Cancer Inst*. 2017;109(11):djx187.
112. Liang S-Y, Phillips KA, Wang G, Keohane C, Armstrong J, Morris WM, et al. Tradeoffs of using administrative claims and medical records to identify the use of personalized medicine for patients with breast cancer. *Med Care.* 2011;49(6):e1-e8.
113. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiology.* 2005;58(4):323-37.
114. Birnbaum HG, Cremieux PY, Greenberg PE, LeLorier J, Ostrander J, Venditti L. Using healthcare claims data for outcomes research and pharmacoeconomic analyses. *Pharmacoeconomics*. 1999;16:1-8.
115. Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a linked Medicare-tumor registry database. *Med Care.* 1993;31(8):749-56.
116. Ruddy KJ, Van Houten HK, Sangaralingham LR, Freedman RA, Thompson CA, Hashmi SK, et al. Impact of treatment regimen on acute care use during and after adjuvant chemotherapy for early-stage breast cancer. *Breast Cancer Res Treat*. 2017;164:515-25.

117. Dalal AA, Gagnon-Sanschagrin P, Burne R, Guérin A, Gauthier G, Small T, et al. Dosing patterns and economic burden of palbociclib drug wastage in HR+/HER2− metastatic breast cancer. *Adv Ther.* 2018;35:768-78.
118. Reeder-Hayes KE, Meyer AM, Hinton SP, Meng K, Carey LA, Dusetzina SB. Comparative toxicity and effectiveness of trastuzumab-based chemotherapy regimens in older women with early-stage breast cancer. *J Clin Oncol*. 2017;35(29):3298-305.
119. West S RM, Poole C. . Validity of pharmacoepidemiologic drug and diagnosis data. In: Wiley Blackwell Strom B KS, Hennessy S, eds., editor. Textbook of Pharmacoepidemiology Oxford: Wiley Blackwell; 2013.
120. Toh SAS AS, Raebel MA, et al. Examples of existing automated databases. In: Strom B KS, Hennessy S, editor. Textbook of Pharmacoepidemiology. Oxford: Wiley Blackwell; 2013.
121. Faria R, Alava M, Manca A, et al. NICE DSU technical support document 17: the use of observational data to inform estimates of treatment effectiveness in technology appraisal: methods for comparative individual patient data 2015. Accessed July 4, 2018. Available from: http://www.nicedsu.org.uk.
122. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report—Part II. *Value Health.* 2009;12(8):1053-61.
123. Berger ML, Martin BC, Husereau D, Worley K, Allen JD, Yang W, et al. A questionnaire to assess the relevance and credibility of observational studies to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health.* 2014;17(2):143-56.
124. Dreyer NA, Bryant A, Velentgas P. The GRACE checklist: a validated assessment tool for high quality observational studies of comparative effectiveness. *J Manag Care Sec Pharm.* 2016;22(10):1107-13.
125. Global Burden of Disease Cancer Collaboration. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the global burden of disease study. *JAMA Oncol*. 2018:1553-68.
126. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol*. 2013;14(12):1165-74.
127. Mariotto AB, Robin Yabroff K, Shao Y, Feuer EJ, Brown ML. Projections of the cost of cancer care in the United States: 2010–2020. *J Natl Cancer Inst*. 2011;103(2):117-28.
128. Li Z, Kang Y. Emerging therapeutic targets in metastatic progression: a focus on breast cancer. *Pharmacol Ther.* 2016;161:79-96.
129. QuintilesIMS. Global Oncology Trends 2017: Advances, Complexity and Costs. May 31, 2017. Available from: https://www.iqvia.com/insights/the-iqvia-institute/reports-and-publications/reports/global-oncology-trends-2017-advances-complexity-and-cost
130. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA_Group*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-9.
131. Peters BJ, Janssen VE, Schramel FM, van de Garde EM. Quantitative and qualitative assessment of real world data comparative effectiveness research of systemic therapies in lung oncology: a systematic review. *Cancer Epidemiol*. 2016;44:5-15.
132. van Walraven C, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol.* 2004;57(7):672-82.

133. Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part III. *Value Health*. 2009;12(8):1062-73.
134. Dreyer NA, Schneeweiss S, McNeil BJ, Berger ML, Walker AM, Ollendorf DA, et al. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. *The A J Man Care*. 2010;16(6):467-71.
135. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiology Drug Saf*. 2007;16(3):241-9.
136. Suissa S. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? *Diabetes Care*. 2018;41(1):6-10.
137. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JP. Current use of routinely collected health data to complement randomized controlled trials: a meta-epidemiological survey. *CMAJ open*. 2016;4(2):E132-E40.
138. Visvanathan K, Levit LA, Raghavan D, Hudis CA, Wong S, Dueck A, et al. Untapped potential of observational research to inform clinical decision making: American Society of Clinical Oncology research statement. *J Clin Oncol*. 2017;35(16):1845-54.
139. Chien H-C, Yang Y-HK, Bai JP. Trastuzumab-related cardiotoxic effects in Taiwanese women: a nationwide cohort study. *JAMA Oncol*. 2016;2(10):1317-25.
140. De Glas N, Hamaker M, Kiderlen M, De Craen A, Mooijaart S, Van De Velde C, et al. Choosing relevant endpoints for older breast cancer patients in clinical trials: an overview of all current clinical trials on breast cancer treatment. *Breast Cancer Res Treat*. 2014;146:591-7.
141. Seruga B, Templeton AJ, Badillo FEV, Ocana A, Amir E, Tannock IF. Under-reporting of harm in clinical trials. *Lancet Oncol*. 2016;17(5):e209-e19.
142. O'Sullivan AK, Thompson D, Drummond MF. Collection of health-economic data alongside clinical trials: is there a future for piggyback evaluations? *Value Health*. 2005;8(1):67-79.
143. Barcenas CH, Niu J, Zhang N, Zhang Y, Buchholz TA, Elting LS, et al. Risk of hospitalization according to chemotherapy regimen in early-stage breast cancer. *J Clin Oncol*. 2014;32(19):2010-7.
144. Schneider M, Zuckerman IH, Onukwugha E, Pandya N, Seal B, Gardner J, et al. Chemotherapy Treatment and Survival in Older Women with Estrogen Receptor–Negative Metastatic Breast Cancer: A Population-Based Analysis. *J Am Geriatr Soc*. 2011;59(4):637-46.
145. Yurkovich M, Avina-Zubieta JA, Thomas J, Gorenchtein M, Lacaille D. A systematic review identifies valid comorbidity indices derived from administrative health data. *J Clin Epidemiol*. 2015;68(1):3-14.
146. Choi JY, Cho EY, Choi YJ, Lee JH, Jung SP, Cho KR, et al. Incidence and risk factors for congestive heart failure in patients with early breast cancer who received anthracycline and/or trastuzumab: a big data analysis of the Korean Health Insurance Review and Assessment service database. *Breast Cancer Res Treat*. 2018;171:181-8.
147. Nørgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions-a primer for the clinician. *Clini Epidemiol*. 2017:185-93.
148. Iudici M, Porcher R, Riveros C, Ravaud P. Time-dependent biases in observational studies of comparative effectiveness research in rheumatology. A methodological review. *Ann Rheum Dis*. 2019;78(4):562-9.
149. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. 2014;68(3):283-7.

150. Langan SM, Schmidt SA, Wing K, Ehrenstein V, Nicholls SG, Filion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363.
151. Carnahan RM. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative data: summary of findings and suggestions for future research. *Pharmacoepidemiol Drug Saf.* 2012;21:90-9.
152. Hassett MJ, Ritzwoller DP, Taback N, Carroll N, Cronin AM, Ting GV, et al. Validating billing/encounter codes as indicators of lung, colorectal, breast, and prostate cancer recurrence using 2 large contemporary cohorts. *Med Care.* 2014;52(10):e65-e73.
153. Du XL, Osborne C, Goodwin JS. Population-based assessment of hospitalizations for toxicity from chemotherapy in older women with breast cancer. *J Clin Oncol*. 2002;20(24):4636-42.
154. Schulman KL, Berenson K, Shih Y-CT, Foley KA, Ganguli A, De Souza J, et al. A checklist for ascertaining study cohorts in oncology health services research using secondary data: report of the ISPOR oncology good outcomes research practices working group. *Value Health.* 2013;16(4):655-69.
155. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol.* 2007;36(3):666-76.
156. 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.
157. Piccart MJ, Hilbers FS, Bliss JM, Caballero C, Frank ES, Renault P, et al. Road map to safe and well-designed de-escalation trials of systemic adjuvant therapy for solid tumors. *J Clin Oncol*. 2020;38(34):4120-9.
158. Hayes JH, Ollendorf DA, Pearson SD, Barry MJ, Kantoff PW, Lee PA, et al. Observation versus initial treatment for men with localized, low-risk prostate cancer: a cost-effectiveness analysis. *Ann Intern Med*. 2013;158(12):853-60.
159. Ontario Health (Quality). Gene Expression Profiling Tests for Early-Stage Invasive Breast Cancer: A Health Technology Assessment. *Ont Health Technol Assess Ser*. 2020;20(10):1-234.
160. Retèl VP, Byng D, Linn SC, Jóźwiak K, Koffijberg H, Rutgers EJ, et al. Cost-effectiveness analysis of the 70-gene signature compared with clinical assessment in breast cancer based on a randomised controlled trial. *Eur J Cancer*. 2020;137:193-203.
161. Lopes Cardozo JM, Drukker CA, Rutgers EJ, Schmidt MK, Glas AM, Witteveen A, et al. Outcome of patients with an ultralow-risk 70-gene signature in the MINDACT trial. *J Clin Oncol*. 2022;40(12):1335-45.
162. Esserman LJ, Yau C, Thompson CK, van't Veer LJ, Borowsky AD, Hoadley KA, et al. Use of molecular tools to identify patients with indolent breast cancers with ultralow risk over 2 decades. *JAMA Oncol.* 2017;3(11):1503-10.
163. Goss PE, Ingle JN, Alés-Martínez JE, Cheung AM, Chlebowski RT, Wactawski-Wende J, et al. Exemestane for breast-cancer prevention in postmenopausal women. *N Engl J Med*. 2011;364(25):2381-91.
164. Cuzick J, Sestak I, Cawthorn S, Hamed H, Holli K, Howell A, et al. Tamoxifen for prevention of breast cancer: extended long-term follow-up of the IBIS-I breast cancer prevention trial. *Lancet Oncol*. 2015;16(1):67-75.

165. Cuzick J, Forbes JF, Sestak I, Cawthorn S, Hamed H, Holli K, et al. Long-term results of tamoxifen prophylaxis for breast cancer—96-month follow-up of the randomized IBIS-I trial. *J Natl Cancer Inst*. 2007;99(4):272-82.
166. Kleinman NL, Rohrbacker NJ, Bushmakin AG, Whiteley J, Lynch WD, Shah SN. Direct and indirect costs of women diagnosed with menopause symptoms. *J Occup Environ Med.* 2013;55(4):465-70.
167. Federatie van Medisch Specialisten. Richtlijnen Borstkanker – Genexpressie profielen. Available from: https://richtlijnendatabase.nl/richtlijn/borstkanker/risicoprofilering/genexpressie_profielen.html
168. National Health Care Institute the Netherlands (Zorginstituut Nederland). Standpunt MammaPrint® bij vrouwen met vroeg stadium borstkanker (herbeoordeling). September 27, 2018. Available from: https://www.zorginstituutnederland.nl/publicaties/standpunten/2018/09/27/standpuntmammaprintbijvrouwen-met-vroeg-stadium-borstkanker-herbeoordeling.
169. National Comprehensive Cancer Network_(NCCN). Breast Cancer: NCCN Clinical Practice Guidelines in Oncology version 2.2022 - December 20, 2021. 2022. . Available from: https://www.nccn.org/guidelines/recently-published-guidelines.
170. Federatie Medisch Specialisten (FMS). Dutch Breast Cancer Guidelines 2020. Available from: https://richtlijnendatabase.nl/richtlijn/borstkanker/startpagina_-_borstkanker.html.
171. National Health Care Institute the Netherlands (Zorginstituut Nederland). Dutch Pharmacoeconomic Guidelines. National Health Care Institute. 2016. Available from: https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg
172. National Health Care Institute the Netherlands (ZIN). Cost-effectiveness in practice. 2015. Available from: https://www.zorginstituutnederland.nl/publicaties/rapport/2015/06/26/kosteneffectiviteit-in-de-praktijk.
173. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A et all. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Oxford Centre for Evidence-Based Medicine. Accessed Jan 2, 2023. Available from: https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebmlevels-of-evidence.
174. Statistics Netherlands (CBS). Survival probabillity by gender and age 2022. Available from: https://opendata.cbs.nl/#/CBS/nl/dataset/70701ned/table.
175. van de Velde JH RD, Seynaeve C, Putter H, Hasenburg A, Vannetzel J, Paridaens R, Markopoulos C, Hozumi Y, Hille ETM, Kieback DG, Asmar L, Smeets J, Nortier JWR, Hadji R, Barlett JMS, Jones SE. Adjuvant tamoxifen and exemestane in early breast cancer (TEAM): a randomised phase 3 trial. *Lancet Oncol.* 2011;377(9762):321-31.
176. Næss IA, Christiansen S, Romundstad P, Cannegieter S, Rosendaal FR, Hammerstrøm J. Incidence and mortality of venous thrombosis: a population-based study. *J Thromb Haemost*. 2007;5(4):692-9.
177. Kadakia KC, Snyder CF, Kidwell KM, Seewald NJ, Flockhart DA, Skaar TC, et al. Patient-reported outcomes and early discontinuation in aromatase inhibitor-treated postmenopausal women with early stage breast cancer. *Oncologist*. 2016;21(5):539-46.
178. van Herk-Sukel MP, van de Poll-Franse LV, Voogd AC, Nieuwenhuijzen GA, Coebergh JWW, Herings R. Half of breast cancer patients discontinue tamoxifen and any endocrine treatment before the end of the recommended treatment period of 5 years: a population-based analysis. *Breast Cancer Res Treat*. 2010;122(3):843-51.
179. Centraal Bureau voor Statistiek (CBS). Survival probabillity by gender and age 2022. Available from: https://opendata.cbs.nl/#/CBS/nl/dataset/70701ned/table.

180. Francis PA, Pagani O, Fleming GF, Walley BA, Colleoni M, Láng I, et al. Tailoring adjuvant endocrine therapy for premenopausal breast cancer. *N Engl J Med*. 2018;379(2):122-37.
181. Federatie Medisch Specialisten (FMS). Endometriumcarcinoom. 2021. Available from: https://richtlijnendatabase.nl/richtlijn/endometriumcarcinoom/startpagina_-_endometriumcarcinoom.html.
182. Cuzick J, Forbes J, Edwards R, Baum M, Cawthorn S, Coates A, et al. First results from the International Breast Cancer Intervention Study (IBIS-I): a randomised prevention trial. *Lancet*. 2002;360(9336):817-24.
183. Goss PE IJ, Alés-Martínez JE, Cheung AM, Chlebowski RT, Wactawski-Wende J, McTiernan A, Robbins J, Johnson KC, Martin LW, Winquist E, Sarto GE, Garber JE, Fabian CJ, Pujol P, Maunsell E, Farmer P, Gelmon KA, Tu D, Richardson H, NCIC CTG MAP.3 Study Investigators. Exemestane for Breast-Cancer Prevention in Postmenopausal Women. *N Engl J Med*. 2011;23(364):2381-91.
184. Lidgren M, Wilking N, Jönsson B, Rehnberg C. Health related quality of life in different states of breast cancer. *Qual Life Res*. 2007;16(6):1073-81.
185. Janssen M, Szende A, Cabases J, Ramos-Goñi JM, Vilagut G, König H-H. Population norms for the EQ-5D-3L: a cross-country analysis of population surveys for 20 countries. *Eur J Health Econ*. 2019;20(2):205-16.
186. Lloyd AJ, Dewilde S, Noble S, Reimer E, Lee AY. What impact does venous thromboembolism and bleeding have on cancer patients' quality of life? *Value Health*. 2018;21(4):449-55.
187. Snowsill T, Coelho H, Huxley N, Jones-Hughes T, Briscoe S, Frayling IM, et al. Molecular testing for Lynch syndrome in people with colorectal cancer: systematic reviews and economic evaluation. *Health Technol Assess*. 2017;21(51):1-280.
188. Braal CL, Jager A, Hoop EO-d, Westenberg JD, Lommen KM, de Bruijn P, et al. Therapeutic drug monitoring of endoxifen for tamoxifen precision dosing: feasible in patients with hormone-sensitive breast cancer. *Clin Pharmacokinetics*. 2022;61(4):527-37.
189. Braal CL, Kleijburg A, Jager A, Koolen SL, Mathijssen RH, Corro Ramos I, et al. Therapeutic Drug Monitoring-Guided Adjuvant Tamoxifen Dosing in Patients with Early Breast Cancer: A Cost-Effectiveness Analysis from the Prospective TOTAM Trial. *Clin Drug Investig*. 2022;42(2):163-75.
190. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L. The iMTA productivity cost questionnaire: a standardized instrument for measuring and valuing health-related productivity losses. *Value Health*. 2015;18(6):753-8.
191. Versteegh MM, Vermeulen KM, Evers SM, De Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value Health.* 2016;19(4):343-52.
192. Retèl VP, Joore MA, Knauer M, Linn SC, Hauptmann M, van Harten WH. Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. *Eur J Canc*. 2010;46(8):1382-91.
193. National Health Care Institute the Netherlands (Zorginstituut Nederland). Farmacotherapeutisch Kompas. Available from: https://www.farmacotherapeutischkompas.nl/.
194. Rijksoverheid. Wanneer gaat mijn OAW in? Available from: https://www.rijksoverheid.nl/onderwerpen/algemene-ouderdomswet-aow/vraag-en-antwoord/wanneer-gaat-mijn-aow-in.
195. van Baal PH, Wong A, Slobbe LC, Polder JJ, Brouwer WB, de Wit GA. Standardizing the inclusion of indirect medical costs in economic evaluations. *Pharmacoeconomics*. 2011;29(3):175-87.
196. Verbeek JG, Atema V, Mewes JC, van Leeuwen M, Oldenburg HS, van Beurden M, et al. Cost-utility, cost-effectiveness, and budget impact of Internet-based cognitive behavioral therapy for breast cancer survivors with treatment-induced menopausal symptoms. *Breast Canc Res Treat*. 2019;178(3):573-85.

197. Lötters FJ, van den Bergh JP, de Vries F, Rutten-van Mölken MP. Current and future incidence and costs of osteoporosis-related fractures in the Netherlands: combining claims data with BMD measurements. *Calcified Tissue Int*. 2016;98(3):235-43.
198. Heisen M, Treur MJ, Heemstra HE, Giesen EB, Postma MJ. Cost-effectiveness analysis of rivaroxaban for treatment and secondary prevention of venous thromboembolism in the Netherlands. *J Med Econ*. 2017;20(8):813-24.
199. Pennington M, Gentry-Maharaj A, Karpinskyj C, Miners A, Taylor J, Manchanda R, et al. Long-term secondary care costs of endometrial cancer: a prospective cohort study nested within the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *PLoS One*. 2016;11(11):e0165539.
200. Vondeling G, Menezes G, Dvortsin E, Jansman F, Konings I, Postma M, et al. Burden of early, advanced and metastatic breast cancer in The Netherlands. *BMC Cancer*. 2018;18:1-11.
201. van der Linden N, Buter J, Pescott CP, Lalisang RI, De Boer JP, de Graeff A, et al. Treatments and costs for recurrent and/or metastatic squamous cell carcinoma of the head and neck in the Netherlands. *Eur Arch Otorhino-laryngology*. 2016;273:455-64.
202. Noels E, Hollestein L, Luijkx K, Louwman M, Carin D, van den Bos R, et al. Increasing costs of skin cancer due to increasing incidence and introduction of pharmaceuticals, 2007–2017. *Act Derm Venereol.* 2020;100(10).
203. Gout-Zwart J, Vondeling G, Rozenbaum M, Beby A, Postma M. PCN393-The burden of informal care for breast cancer patients in the Netherlands. *Value Health*. 2018;21:S81.
204. Vemer P, Corro Ramos I, Van Voorn G, Al M, Feenstra T. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. *Pharmacoeconomics*. 2016;34(4):349-61.
205. Comprehensive Cancer Centre the Netherlands (IKNL). Breast Cancer in the Netherlands: trends 1989-2020 based on data from the Netherlands Cancer Registry. 2020. Available from: https://iknl.nl/getmedia/fe459c3a-c561-40de-b740-fff15997f020/IKNL-Folder-Borstkanker-2020.pdf.
206. Condorelli R I-LI. Managing side effects in adjuvant endocrine therapy for breast cancer. *Expert Rev Anticancer Ther.* 2019;18:1101-12.
207. Ferreira A, Di Meglio A, Pistilli B, Gbenou A, El-Mouhebb M, Dauchy S, et al. Differential impact of endocrine therapy and chemotherapy on quality of life of breast cancer survivors: a prospective patient-reported outcomes analysis. *Ann Oncol*. 2019;30(11):1784-95.
208. Drummond MF, Sculpher M, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press. 2015.
209. Walraven J, Jacobs MS, Uyl-de Groot CA. Leveraging the Similarities Between Cost-Effectiveness Analysis and Value-Based Healthcare. *Value Health*. 2021;24(7):1038-44.
210. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Five-year survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med*. 2019;381(16):1535-46.
211. Wait S, Han D, Muthu V, Oliver K, Chrostowski S, Florindi F, et al. Towards sustainable cancer care: Reducing inefficiencies, improving outcomes—A policy report from the All. Can initiative. *J Cancer Policy.* 2017;13:47-64.
212. 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.
213. Uyl-de Groot CA, Heine R, Krol M, Verweij J. Unequal access to newly registered cancer drugs leads to potential loss of life-years in Europe. *Cancers*. 2020;12(8):2313.

214. Wilking N, Bucsics A, Sekulovic LK, Kobelt G, Laslop A, Makaroff L, et al. Achieving equal and timely access to innovative anticancer drugs in the European Union (EU): summary of a multidisciplinary CECOG-driven roundtable discussion with a focus on Eastern and South-Eastern EU countries. *ESMO open.* 2019;4(6):e000550.
215. Cortes J, Perez-García JM, Llombart-Cussac A, Curigliano G, El Saghir NS, Cardoso F, et al. Enhancing global access to cancer medicines. *CA: Cancer J Clin*. 2020;70(2):105-24.
216. Fundytus A, Sengar M, Lombe D, Hopman W, Jalink M, Gyawali B, et al. Access to cancer medicines deemed essential by oncologists in 82 countries: an international, cross-sectional survey. *Lancet Oncol*. 2021;22(10):1367-77.
217. Janzic U, Knez L, Janzic A, Cufer T. Time to access to novel anticancer drugs and the correlation with ESMO-Magnitude of Clinical Benefit Scale in Slovenia. *Expert Rev Pharmacoecon Outcomes Res*. 2019;19(6):717-23.
218. Downing NS, Aminawung JA, Shah ND, Braunstein JB, Krumholz HM, Ross JS. Regulatory review of novel therapeutics—comparison of three regulatory agencies. *N Engl J Med*. 2012;366(24):2284-93.
219. Cheema P, Gavura S, Migus M, Godman B, Yeung L, Trudeau M. International variability in the reimbursement of cancer drugs by publically funded drug programs. *Curr Oncol*. 2012;19(3):165-76.
220. Jönsson BH, Hofmarcher T, Lindgren P, Wilking N. A Comparator Report on Patient Access to Cancer Medicines in Europe Revisited. 2016. Available from: https://www.efpia.eu/media/412110/ihe-report-2016_4_.pdf.
221. Hortobagyi GN, Stemmer SM, Burris HA, Yap Y-S, Sonke GS, Paluch-Shimon S, et al. Updated results from MONALEESA-2, a phase III trial of first-line ribociclib plus letrozole versus placebo plus letrozole in hormone receptor-positive, HER2-negative advanced breast cancer. *Ann Oncol*. 2018;29(7):1541-7.
222. Slamon DJ, Neven P, Chia S, Fasching PA, De Laurentiis M, Im S-A, et al. Overall survival with ribociclib plus fulvestrant in advanced breast cancer. *N Engl J Med*. 2020;382(6):514-24.
223. Nederlandse Vereniging voor Medische Oncologie (NVMO). Zijn de PASKWIL-criteria anno 2007 nog up-to-date? September 14, 2007. Available from: https://www.nvmo.org/bom/zijn-de-paskwil-criteria-anno-2007-nog-up-to-date/.
224. National Health Care Institute (Zorginstituut Nederland). Beoordelingsprocedure specialistische geneesmiddelen. May 11, 2020. Accessed August 22, 2022. https://www.zorginstituutnederland.nl/over-ons/werkwijzen-en-procedures/.
225. van Der Meer DJ, Kramer I, van Maaren MC, van Diest PJ, C Linn S, Maduro JH, et al. Comprehensive trends in incidence, treatment, survival and mortality of first primary invasive breast cancer stratified by age, stage and receptor subtype in the Netherlands between 1989 and 2017. *Int J cancer.* 2021;148(9):2289-303.
226. Lobbezoo D, Van Kampen R, Voogd A, Dercksen M, Van Den Berkmortel F, Smilde T, et al. Prognosis of metastatic breast cancer: are there differences between patients with de novo and recurrent metastatic breast cancer? *Br J Cancer*. 2015;112(9):1445-51.
227. Lobbezoo D, Van Kampen R, Voogd A, Dercksen M, van den Berkmortel F, Smilde T, et al. In real life, one-quarter of patients with hormone receptor-positive metastatic breast cancer receive chemotherapy as initial palliative therapy: a study of the Southeast Netherlands Breast Cancer Consortium. *Ann Oncol*. 2016;27(2):256-62.

228. Fietz T, Tesch H, Rauh J, Boller E, Kruggel L, Jänicke M, et al. Palliative systemic therapy and overall survival of 1,395 patients with advanced breast cancer–results from the prospective German TMK cohort study. *Breast*. 2017;34:122-30.
229. Rugo H, Finn R, Diéras V, Ettl J, Lipatov O, Joy A, et al. Palbociclib plus letrozole as first-line therapy in estrogen receptor-positive/human epidermal growth factor receptor 2-negative advanced breast cancer with extended follow-up. *Breast Cancer Res Teat*. 2019;174:719-29.
230. Medische Oncologie. Plaatsbepaling NABON en NVMO: toepassing van palbociclib bij mammacarcinoom. February 13, 2017. Available from: https://medischeoncologie.nl/artikelen/2017/februari/editie-1/plaatsbepaling-nabon-en-nvmo-toepassing-van-palbociclib-bij-mammacarcinoom.
231. Finn RS, Martin M, Rugo HS, Jones S, Im S-A, Gelmon K, et al. Palbociclib and letrozole in advanced breast cancer. *N Engl J Med*. 2016;375(20):1925-36.
232. Cristofanilli M, Turner NC, Bondarenko I, Ro J, Im S-A, Masuda N, et al. Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. *Lancet Oncol*. 2016;17(4):425-39.
233. Hoyle MW, Henley W. Improved curve fits to summary survival data: application to economic evaluation of health technologies. *BMC Med Res Methodol*. 2011;11(1):1-14.
234. Latimer NR. Survival analysis for economic evaluations alongside clinical trials—extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. *Med Decis Making*. 2013;33(6):743-54.
235. Newton M, Scott K, Troein P (IQVIA). EFPIA Patients W.A.I.T. Indicator 2020 Survey. 2021. https://www.efpia.eu/media/602652/efpia-patient-wait-indicator-final-250521.pdf.
236. Jacquet E, Kerouani-Lafaye G, Grude F, Goncalves S, Lorence A, Turcry F, et al. Comparative study on anticancer drug access times between FDA, EMA and the French temporary authorisation for use program over 13 years. *Eur J Cancer*. 2021;149:82-90.
237. van Waalwijk van Doorn-Khosrovani P-vRA, Bloemendal H & Gelderblom H. Oncolytics are readily available during the 'lock' procedure. *Pharmaceutisch Weekblad*. 2021;156(38).
238. European Federation of Pharmaceutical Industries and Associations. Time to act together: patients wait more than 6 times longer for access to new medicines in neighboring European countries. July 2, 2020. Available from: https://www.efpia.eu/news-events/the-efpia-view/statements-press-releases/time-to-act-together-patients-wait-more-than-6-times-longer-for-access-to-new-medicines-in-neighboring-european-countries/.
239. Sou H. EU Compassionate Use Programmes (CUPs). *Pharmaceutical Med*. 2010;24(4):223-9.
240. Patil S. Early access programs: benefits, challenges, and key considerations for successful implementation. *Perspectives Clin Res.* 2016;7(1):4.
241. Bates AK. Implementing a pre-launch named patient programme: evidence of increased market share. *J Med Mark*. 2008;8(4):319-24.
242. Balasubramanian G, Morampudi S, Chhabra P, Gowda A, Zomorodi B. An overview of Compassionate Use Programs in the European Union member states. *Intractable Rare Dis Res*. 2016;5(4):244-54.
243. Lauenroth VD, Kesselheim AS, Sarpatwari A, Stern AD. Lessons from the impact of price regulation on the pricing of anticancer drugs in Germany: study assesses the effects of female political representation on mortality among children younger than age five in Brazil. *Health Aff*. 2020;39(7):1185-93.

244. Burne R, Balu S, Guérin A, Bungay R, Sin R, Paul ML. Comparison of healthcare resource utilization and costs of patients with HR+/HER2-advanced breast cancer treated with ribociclib versus other CDK4/6 inhibitors. *J Med Econ*. 2021;24(1):806-15.
245. Harbeck N, Bartlett M, Spurden D, Hooper B, Zhan L, Rosta E, et al. CDK4/6 inhibitors in HR+/HER2-advanced/metastatic breast cancer: a systematic literature review of real-world evidence studies. Future Oncol. 2021;17(16):2107-22.
246. Fountzilas E, Koliou G-A, Vozikis A, Rapti V, Nikolakopoulos A, Boutis A, et al. Real-world clinical outcome and toxicity data and economic aspects in patients with advanced breast cancer treated with cyclin-dependent kinase 4/6 (CDK4/6) inhibitors combined with endocrine therapy: the experience of the Hellenic Cooperative Oncology Group. *ESMO open*. 2020;5(4):e000774.
247. Marineau A, St-Pierre C, Lessard-Hurtubise R, David M-È, Adam J-P, Chabot I. Cyclin-dependent kinase 4/6 inhibitor treatment use in women treated for advanced breast cancer: Integrating ASCO/NCODA patient-centered standards in a community pharmacy. *J Oncol Pharm Pract*. 2022:10781552221102884.
248. West MT, Smith CE, Kaempf A, Kohs TC, Amirsoltani R, Ribkoff J, et al. CDK 4/6 inhibitors are associated with a high incidence of thrombotic events in women with breast cancer in real-world practice. *Eur J Haematol*. 2021;106(5):634-42.
249. Burstein HJ, Somerfield MR, Barton DL, Dorris A, Fallowfield LJ, Jain D, et al. Endocrine treatment and targeted therapy for hormone receptor–positive, human epidermal growth factor receptor 2–negative metastatic breast cancer: ASCO guideline update. *J Clin Oncol*. 2021;39(35):3959-77.
250. van Ommen-Nijhof A, Konings I, Van Zeijl C, Uyl-de Groot C, van der Noort V, Jager A, et al. S electing the optimal positio n of CDK4/6 i nhibitors in hormone receptor-positive a dvanced breast cancer–the SONIA study: study protocol for a randomized controlled trial. *BMC Cancer*. 2018;18(1):1-7.
251. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. *J R Soc Med*. 2011;104(12):510-20.
252. Westerling R, Westin M, McKee M, Hoffmann R, Plug I, Rey G, et al. The timing of introduction of pharmaceutical innovations in seven E uropean countries. Journal of evaluation in clinical practice. 2014;20(4):301-10.
253. Dietz JR, Moran MS, Isakoff SJ, Kurtzman SH, Willey SC, Burstein HJ, et al. Recommendations for prioritization, treatment, and triage of breast cancer patients during the COVID-19 pandemic. the COVID-19 pandemic breast cancer consortium. *Breast Cancer Res Treat*. 2020;181(3):487-97.
254. Verhoeven RH, Karim-Kos HE, Coebergh JW, Brink M, Horenblas S, de Wit R, et al. Markedly increased incidence and improved survival of testicular cancer in the Netherlands. *Acta Oncol*. 2014;53(3):342-50.
255. Pavlik T, Majek O, Buchler T, Vyzula R, Petera J, Ryska M, et al. Trends in stage-specific population-based survival of cancer patients in the Czech Republic in the period 2000-2008. *Cancer Epidemiol.* 2014;38(1):28-34.
256. Gapstur SM, Thun MJ. Progress in the war on cancer. *JAMA*. 2010;303(11):1084-5.
257. 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*. 2018;391(10125):1023-75.
258. Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, et al. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer

Benchmarking Partnership): an analysis of population-based cancer registry data. *Lancet*. 2011;377(9760):127-38.

259. Siesling S, Kwast A, Gavin A, Baili P, Otter R. Availability of stage at diagnosis, cancer treatment delay and compliance with cancer guidelines as cancer registry indicators for cancer care in Europe: Results of EUROCHIP-3 survey. *Int J Cancer.* 2013;132(12):2910-7.
260. Boyle JM, Hegarty G, Frampton C, Harvey-Jones E, Dodkins J, Beyer K, et al. Real-world outcomes associated with new cancer medicines approved by the Food and Drug Administration and European Medicines Agency: a retrospective cohort study. *Eur J Cancer*. 2021;155:136-44.
261. Westgeest HM, Uyl-de Groot CA, van Moorselaar RJ, de Wit R, van den Bergh AC, Coenen JL, et al. Differences in trial and real-world populations in the Dutch Castration-resistant Prostate Cancer Registry. *Eur Urol Focus*. 2018;4(5):694-701.
262. Verver D, Grünhagen DJ, van Akkooi AC, Aarts MJ, van den Berkmortel FW, van den Eertwegh AJ, et al. Clinical outcome of patients with metastatic melanoma of unknown primary in the era of novel therapy. *Cancer Immunol Immunother*. 2021:1-13.
263. Dickman PW. Estimating and modeling relative survival. *Stata J*. 2015.
264. Perme MP, Stare J, Estève J. On estimation in relative survival. *Biometrics*. 2012;68(1):113-20.
265. Jensen OP, MacLenna R, Muir CS, Skeet RG. Cancer registration: principles and methods: International Agency for Research on Cancer. 1991. Accessed November 10, 2021. Available from: https://publications.iarc.fr/Book-And-Report-Series/Iarc-Scientific-Publications/Cancer-Registration-Principles-AndMethods-1991
266. Verweij J, Casali PG, Zalcberg J, LeCesne A, Reichardt P, Blay J-Y, et al. Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. *Lancet*. 2004;364(9440):1127-34.
267. Patel S. Long-term efficacy of imatinib for treatment of metastatic GIST. *Cancer Chemother Pharmacol*. 2013;72(2):277-86.
268. Chia SK, Speers CH, D'yachkova Y, Kang A, Malfair-Taylor S, Barnett J, et al. The impact of new chemotherapeutic and hormone agents on survival in a population-based cohort of women with metastatic breast cancer. *Cancer*. 2007;110(5):973-9.
269. Quaedvlieg PF, Visser O, Lamers CB, Janssen-Heijen M, Taal BG. Epidemiology and survival in patients with carcinoid disease in the Netherlands: An epidemiological study with 2391 patients. *Ann Oncol*. 2001;12(9):1295-300.
270. Dasari A, Shen C, Halperin D, Zhao B, Zhou S, Xu Y, et al. Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States. *JAMA Oncol*. 2017;3(10):1335-42.
271. Yao JC, Hassan M, Phan A, Dagohoy C, Leary C, Mares JE, et al. One hundred years after "carcinoid": epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. *J Clin Oncol*. 2008;26(18):3063-72.
272. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Five-year survival and correlates among patients with advanced melanoma, renal cell carcinoma, or non–small cell lung cancer treated with nivolumab. *JAMA Oncol*. 2019;5(10):1411-20.
273. Asher N, Ben-Betzalel G, Lev-Ari S, Shapira-Frommer R, Steinberg-Silman Y, Gochman N, et al. Real world outcomes of ipilimumab and nivolumab in patients with metastatic melanoma. *Cancers*. 2020;12(8):2329.
274. Franken M, Leeneman B, Aarts M, van Akkooi A, van den Berkmortel F, Boers-Sonderen M, et al. Trends in survival and costs in metastatic melanoma in the era of novel targeted and immunotherapeutic drugs. *ESMO open*. 2021;6(6):100320.

275. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Five-year outcomes with pembrolizumab versus chemotherapy for metastatic non–small-cell lung cancer with PD-L1 tumor proportion score≥ 50%. *J Clin Oncol*. 2021:JCO. 21.00174.
276. Vokinger KN, Hwang TJ, Grischott T, Reichert S, Tibau A, Rosemann T, et al. Prices and clinical benefit of cancer drugs in the USA and Europe: a cost–benefit analysis. *Lancet Oncol*. 2020;21(5):664-70.
277. Salas-Vega S, Iliopoulos O, Mossialos E. Assessment of overall survival, quality of life, and safety benefits associated with new cancer medicines. *JAMA Oncol*. 2017;3(3):382-90.
278. Ruiterkamp J, Ernst M, De Munck L, van der Heiden-van der Loo M, Bastiaannet E, van de Poll-Franse L, et al. Improved survival of patients with primary distant metastatic breast cancer in the period of 1995–2008. A nationwide population-based study in the Netherlands. *Breast Cancer Res Treat*. 2011;128:495-503.
279. van den Hurk CJ, Eckel R, van de Poll-Franse LV, Coebergh JWW, Nortier JW, Hölzel D, et al. Unfavourable pattern of metastases in M0 breast cancer patients during 1978-2008: a population-based analysis of the Munich Cancer Registry. *Breast Cancer Res Treat*. 2011;128:795-805.
280. Malmgren JA, Mayer M, Atwood MK, Kaplan HG. Differential presentation and survival of de novo and recurrent metastatic breast cancer over time: 1990–2010. *Breast Cancer Res Treat*. 2018;167:579-90.
281. Gennari A, André F, Barrios C, Cortes J, de Azambuja E, DeMichele A, et al. ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer. Ann Oncol. 2021;32(12):1475-95.
282. Khan SA, editor Surgical management of de novo stage IV breast cancer. *Sem Radiat Oncol*; 2016: Elsevier.
283. Headon H, Wazir U, Kasem A, Mokbel K. Surgical treatment of the primary tumour improves the overall survival in patients with metastatic breast cancer: A systematic review and meta-analysis. *Mol Clin Oncol*. 2016;4(5):863-7.
284. Soran A, Ozmen V, Ozbas S, Karanlik H, Muslumanoglu M, Igci A, et al. Randomized trial comparing resection of primary tumor with no surgery in stage IV breast cancer at presentation: protocol MF07-01. *Ann Surg Oncol*. 2018;25:3141-9.
285. Tsukioki T, Shien T, Doihara H. Effect of local surgery on outcomes of stage IV breast cancer. *Transl Cancer Res*. 2020;9(8):5102.
286. Khan SA, Zhao F, Goldstein LJ, Cella D, Basik M, Golshan M, et al. Early local therapy for the primary site in de novo stage IV breast cancer: results of a randomized clinical trial (E2108). *J Clin Oncol*. 2022;40(9):978-87.
287. Miao H, Hartman M, Bhoo-Pathy N, Lee S-C, Taib NA, Tan E-Y, et al. Predicting survival of de novo metastatic breast cancer in Asian women: systematic review and validation study. *PLoS One*. 2014;9(4):e93755.
288. Zeichner SB, Herna S, Mani A, Ambros T, Montero AJ, Mahtani RL, et al. Survival of patients with de-novo metastatic breast cancer: analysis of data from a large breast cancer-specific private practice, a university-based cancer center and review of the literature. *Breast Cancer Res Treat*. 2015;153:617-24.
289. Bertaut A, Mounier M, Desmoulins I, Guiu S, Beltjens F, Darut-Jouve A, et al. Stage IV breast cancer: A population-based study about prognostic factors according to HER 2 and HR status. *Eur J Cancer Care*. 2015;24(6):920-8.
290. Eng LG, Dawood S, Sopik V, Haaland B, Tan PS, Bhoo-Pathy N, et al. Ten-year survival in women with primary stage IV breast cancer. *Breast Cancer Res Treat*. 2016;160:145-52.

291. Tao L, Chu L, Wang LI, Moy L, Brammer M, Song C, et al. Occurrence and outcome of de novo metastatic breast cancer by subtype in a large, diverse population. *Cancer Causes Control*. 2016;27:1127-38.
292. Leone BA, Vallejo CT, Romero AO, Machiavelli MR, Pérez JE, Leone J, et al. Prognostic impact of metastatic pattern in stage IV breast cancer at initial diagnosis. *Breast Cancer Res Treat*. 2017;161:537-48.
293. Press DJ, Miller ME, Liederbach E, Yao K, Huo D. De novo metastasis in breast cancer: occurrence and overall survival stratified by molecular subtype. *Clin ExpMetastasis*. 2017;34:457-65.
294. Vaz-Luis I, Lin NU, Keating NL, Barry WT, Winer EP, Freedman RA. Factors Associated with Early Mortality Among Patients with De Novo Metastatic Breast Cancer: A Population-Based Study. *Oncologist*. 2017;22(4):386-93.
295. Ogiya R, Sagara Y, Niikura N, Freedman RA. Impact of subtype on survival of young patients with stage IV breast cancer. *Clin Breast Cancer*. 2019;19(3):200-7. e1.
296. McKenzie HS, Maishman T, Simmonds P, Durcan L, Fiona J. 11 Hamed Hisham 12 Hodgson Shirley 13 Simmonds Peter 14 Eccles Diana 15 PSGADGJLEGTAMPPHALSERG, Eccles D, et al. Survival and disease characteristics of de novo versus recurrent metastatic breast cancer in a cohort of young patients. *Br J Cancer.* 2020;122(11):1618-29.
297. He Z-Y, Lian C-L, Wang J, Lei J, Hua L, Zhou J, et al. Incorporation of biologic factors for the staging of de novo stage IV breast cancer. *NPJ breast cancer*. 2020;6(1):43.
298. Plichta JK, Thomas SM, Sergesketter AR, Greenup RA, Rosenberger LH, Fayanju OM, et al. A novel staging system for de novo metastatic breast cancer refines prognostic estimates. *Ann Surg*. 2022;275(4):784-92.
299. Gu Y, Wu G, Zou X, Huang P, Yi L. Prognostic value of site-specific metastases and surgery in de novo stage IV triple-negative breast cancer: a population-based analysis. *Med Sci monit*. 2020;26:e920432-1.
300. Li Y, Wang S, Yang W, Liu H. Prognostic significance of molecular subtype, metastatic site and primary tumor surgery for survival in primary metastatic breast cancer: a SEER-based study. *Medicine*. 2021;100(27):e26619.
301. Zheng A, Guo B-L, Zhang J-G, Jin F, Surgery CSoB. Clinical information and management status of de novo stage IV breast cancer patients: a Chinese multicenter investigation (CSBrS-002). *Chin Med Journal*. 2021;134(13):1569-75.
302. Leone JP, Leone BA, Zwenger AO, Vallejo CT, Romero AO, Machiavelli MR, et al. The prognostic significance of metastatic pattern in stage IV male breast cancer at initial diagnosis: a population-based study. *Breast Cancer Res Treat*. 2021;187:237-44.
303. Dawood S, Broglio K, Ensor J, Hortobagyi G, Giordano S. Survival differences among women with de novo stage IV and relapsed breast cancer. *Ann Oncol*. 2010;21(11):2169-74.
304. Rossi V, Nolè F, Redana S, Adamoli L, Martinello R, Aurilio G, et al. Clinical outcome in women with HER2-positive de novo or recurring stage IV breast cancer receiving trastuzumab-based therapy. *Breast*. 2014;23(1):44-9.
305. Yardley DA, Kaufman PA, Brufsky A, Yood MU, Rugo H, Mayer M, et al. Treatment patterns and clinical outcomes for patients with de novo versus recurrent HER2-positive metastatic breast cancer. *Breast Cancer Res Treat*. 2014;145:725-34.
306. Lambertini M, Ferreira AR, Di Meglio A, Poggio F, Puglisi F, Sottotetti F, et al. Patterns of care and clinical outcomes of HER2-positive metastatic breast cancer patients with newly diagnosed stage IV or recurrent disease undergoing first-line trastuzumab-based therapy: a multicenter retrospective cohort study. *Clin Breast Cancer.* 2017;17(8):601-10. e2.

307. Den Brok WD, Speers CH, Gondara L, Baxter E, Tyldesley SK, Lohrisch CA. Survival with metastatic breast cancer based on initial presentation, de novo versus relapsed. *Breast Cancer Res Treat*. 2017;161:549-56.
308. Yamamura J, Kamigaki S, Fujita J, Osato H, Komoike Y. The difference in prognostic outcomes between de novo stage IV and recurrent metastatic patients with hormone receptor-positive, HER2-negative breast cancer. *In Vivo*. 2018;32(2):353-8.
309. Hassett MJ, Uno H, Cronin AM, Carroll NM, Hornbrook MC, Ritzwoller DP. Comparing survival after recurrent vs de novo stage IV advanced breast, lung, and colorectal cancer. *JNCI Cancer Spect*. 2018;2(2):pky024.
310. Seltzer S, Corrigan M, O'Reilly S. The clinicomolecular landscape of de novo versus relapsed stage IV metastatic breast cancer. *Exp Mol Pathol*. 2020;114:104404.
311. Ernst MF, van de Poll-Franse LV, Roukema JA, Coebergh JWW, van Gestel CM, Vreugdenhil G, et al. Trends in the prognosis of patients with primary metastatic breast cancer diagnosed between 1975 and 2002. *Breast*. 2007;16(4):344-51.
312. Lord SJ, Bahlmann K, O'Connell DL, Kiely BE, Daniels B, Pearson S-A, et al. De novo and recurrent metastatic breast cancer–A systematic review of population-level changes in survival since 1995. EClinicalMedicine. 2022;44.
313. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.
314. Bodner TE. What improves with increased missing data imputations? *Struc Equ Model.* 2008;15(4):651-75.
315. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30(4):377-99.
316. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med*. 2009;28(15):1982-98.
317. Nelson W. Hazard plotting for incomplete failure data. *J Qual Technol*. 1969;1(1):27-52.
318. Nelson W. Theory and applications of hazard plotting for censored failure data. *Technometrics*. 1972;14(4):945-66.
319. Aalen O. Nonparametric inference for a family of counting processes. *Ann Stat*. 1978:701-26.
320. Cox DR. Regression models and life tables. *J Stat Soc*. 1972;34:248-75.
321. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982;69(1):239-41.
322. Heinzl H, Kaider A. Gaining more flexibility in Cox proportional hazards regression models with cubic spline functions. *Computer Methods Programs Biomed.* 1997;54(3):201-8.
323. Rulli E, Ghilotti F, Biagioli E, Porcu L, Marabese M, D'Incalci M, et al. Assessment of proportional hazard assumption in aggregate data: a systematic review on statistical methodology in clinical trials using time-to-event endpoint. *Br J Cancer*. 2018;119(12):1456-63.
324. Food and Drug Administration (FDA). Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products. August 30, 2023. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug.
325. European Medicine Association (EMA). Reflection paper on use of real-world data in non interventional studies to generate real-world evidence. 2024. Available from: https://www.ema.europa.eu/en/reflection-paper-use-real-world-data-non-interventional-studies-generate-real-world-evidence-scientific-guideline

326. Gatto NM, Reynolds RF, Campbell UB. A structured preapproval and postapproval comparative study design framework to generate valid and transparent real-world evidence for regulatory decisions. *Clin Pharmacol Ther*. 2019;106(1):103-15.
327. Gatto NM, Vititoe SE, Rubinstein E, Reynolds RF, Campbell UB. A Structured Process to Identify Fit-for-Purpose Study Design and Data to Generate Valid and Transparent Real-World Evidence for Regulatory Uses. Clin Pharmacol Ther. 2023;113(6):1235-9.
328. Schneider PP, Ramaekers BL, Pouwels X, Geurts S, Ibragimova K, de Boer M, et al. Direct medical costs of advanced breast cancer treatment: A real-world study in the Southeast of The Netherlands. *Value Health*. 2021;24(5):668-75.
329. Frederix G, Severens J, Hövels A, Van Hasselt J, Hooiveld M, Neven P, et al. Real world cost of human epidermal receptor 2-positive metastatic breast cancer patients: a longitudinal incidence-based observational costing study in the N etherlands and B elgium. *Eur J Cancer Care*. 2015;24(3):340-54.
330. Noben CY, de Rijk A, Nijhuis F, Kottner J, Evers S. The exchangeability of self-reports and administrative health care resource use measurements: assessement of the methodological reporting quality. *J Clin Epidemiol*. 2016;74:93-106. e2.
331. Bennink C, Westgeest H, Schoonen D, Boersen F, Sonneveld P, Hazelzet J, et al. High Hospital-related Costs at the End-of-life in Patients With Multiple Myeloma: A Single-center Study. *HemaSphere*. 2023;7(6):e913.
332. Dutch Healthcare Authority (NZa). Handleiding DBC systematiek. 2024. Available from: https://www.nvog.nl/wp-content/uploads/2024/03/20240101-Handleiding-dbc-systematiek-v20230928.pdf.
333. Brandão M, Morais S, Lopes-Conceição L, Fontes F, Araújo N, Dias T, et al. Healthcare use and costs in early breast cancer: a patient-level data analysis according to stage and breast cancer subtype. *ESMO open*. 2020;5(6):e000984.
334. Pichardo D, Michael R, Mercer M, Korina N, Onukwugha E. Utility of a Clinically Guided Data-Driven Approach for Predicting Breast Cancer Complications: An Application Using a Population-Based Claims Data Set. *JCO Clin Cancer Inform*. 2022;6:e2100191.
335. Vo JB, Ramin C, Veiga LH, Brandt C, Curtis RE, Bodelon C, et al. Long-term cardiovascular disease risk after anthracycline and trastuzumab treatments in US breast cancer survivors. *J Natl Cancer Instit*. 2024:djae107.
336. Lambert J, Sandhu H, Kean E, Xavier T, Brokman A, Steckler Z, et al. A strategy to identify event specific hospitalizations in large health claims databases. *BMC Health Ser Res*. 2022;22(1):705.
337. European Commission (EC). Will we ever cure cancer(2019 [Available from: https://projects.research-and-innovation.ec.europa.eu/en/horizon-magazine/will-we-ever-cure-cancer.
338. Linnér L, Eriksson I, Persson M, Wettermark B. Forecasting drug utilization and expenditure: ten years of experience in Stockholm. *BMC Health Ser Res*. 2020;20:1-11.
339. Dai WF, de Oliveira C, Blommaert S, Pataky RE, Tran D, Aurangzeb Z, et al. Mapping Canadian data assets to generate real-world evidence: lessons learned from Canadian real-world evidence for value of cancer drugs (CanREValue) collaboration's RWE data working group. *Cur Oncol*. 2022;29(3):2046-63.
340. Padmanabhan S, Carty L, Cameron E, Ghosh RE, Williams R, Strongman H. Approach to record linkage of primary care data from Clinical Practice Research Datalink to other health-related patient data: overview and implications. *Eur J Epidemiol*. 2019;34:91-9.

341. Rivera DR, Gokhale MN, Reynolds MW, Andrews EB, Chun D, Haynes K, et al. Linking electronic health data in pharmacoepidemiology: appropriateness and feasibility. *Pharmacoepidemiol Drug Saf*. 2020;29(1):18-29.
342. Castellanos EH, Wittmershaus BK, Chandwani S. Raising the bar for real-world data in oncology: Approaches to quality across multiple dimensions. *JCO Clin Cancer Infor*. 2024;8:e2300046.
343. European Commission (EC). European Health Data Space. Available from: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en#:~:text=The%20European%20Health%20Data%20Space%20%28EHDS%29%20will%20be,agreement%20on%20the%20Commission%20proposal%20for%20the%20EHDS.
344. National Health Care Institute (Zorginstituut Nederland). Regie op Registers voor Dure Geneesmiddelen (ROR DGM) - casestudies. 2023. Available from: https://www.zorginstituutnederland.nl/over-ons/programmas-en-samenwerkingsverbanden/regie-op-registers-voor-dure-geneesmiddelen.
345. Health RI. Program HDAB-NL 2024. Available from: https://www.health-ri.nl/node/862.
346. Netherlands Comprehensive Cancer Organization (IKNL). Health-RI en IKNL bundelen krachten voor betere databeschikbaarheid bij kankeronderzoek. 2024. Available from: https://iknl.nl/nieuws/2024/mou_health-ri_iknl#:~:text=Health%2DRI%20en%20Stichting%20Integraal,onderzoek%2C%20innovatie%20en%20optimale%20zorg.
347. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3(1):1-9.
348. Khambholja K, Gehani M. Use of structured template and reporting tool for real-world evidence for critical appraisal of the quality of reporting of real-world evidence studies: a systematic review. *Value Health*. 2023;26(3):427-34.
349. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ*. 2021;372.
350. Wang SV, Pottegård A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, et al. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: a good practices report of a joint ISPE/ISPOR task force. *Value Health*. 2022;25(10):1663-72.
351. Hansford HJ, Cashin AG, Jones MD, Swanson SA, Islam N, Dahabreh IJ, et al. Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. *BMJ open*. 2023;13(9):e074626.
352. Banerjee R, Prasad V. Are observational, real-world studies suitable to make cancer treatment recommendations? JAMA Netw Open. 2020;3(7):e2012119-e.
353. Dekkers OM, Groenwold RH. When observational studies can give wrong answers: the potential of immortal time bias. *Eur J Endocrinol*. 2021;184(1):E1-E4
354. Jackson BE, Greenup RA, Strassle PD, Deal AM, Baggett CD, Lund JL, et al. Understanding and identifying immortal-time bias in surgical health services research: an example using surgical resection of stage IV breast cancer. *Surg Oncol*. 2021;37:101539.
355. Duchesneau ED, Jackson BE, Webster-Clark M, Lund JL, Reeder-Hayes KE, Nápoles AM, et al. The timing, the treatment, the question: comparison of epidemiologic approaches to minimize immortal time bias in real-world data using a surgical oncology example. *Cancer Epidemiology Biomarkers Prev*. 2022;31(11):2079-86.
356. Yu OHY, Suissa S. Metformin and cancer: solutions to a real-world evidence failure. Diabetes care. 2023;46(5):904-12.

357. Suissa S. Single-arm trials with historical controls: study designs to avoid time-related biases. *Epidemiology*. 2021;32(1):94-100.
358. Hernán MA, Sterne JA, Higgins JP, Shrier I, Hernández-Díaz S. A structural description of biases that generate immortal time. *Epidemiology*. 2025;36(1):107-14.
359. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *JAMA*. 2022;328(24):2446-7.
360. Wang SV, Schneeweiss S, Franklin JM, Desai RJ, Feldman W, Garry EM, et al. Emulation of randomized clinical trials with nonrandomized database analyses: results of 32 clinical trials. *JAMA*. 2023;329(16):1376-85.
361. Soni PD, Hartman HE, Dess RT, Abugharib A, Allen SG, Feng FY, et al. Comparison of population-based observational studies with randomized trials in oncology. *J Clin Oncol*. 2019;37(14):1209-16.
362. Powell K, Prasad V. Common misconceptions of randomized controlled trials in oncology. *Eur J Clin Invest*. 2022;52(11):e13832.
363. Tang M, Pearson S-A, Simes RJ, Chua BH. Harnessing real-world evidence to advance cancer research. *Cur Oncol*. 2023;30(2):1844-59.
364. Booth CM, Sengar M, Goodman A, Wilson B, Aggarwal A, Berry S, et al. Common Sense Oncology: outcomes that matter. *Lancet Onc*. 2023;24(8):833-5.
365. Koole SN, Huisman AH, Timmers L, Westgeest HM, van Breugel E, Sonke GS, et al. Lessons learned from postmarketing withdrawals of expedited approvals for oncology drug indications. *Lancet Oncol*. 2024;25(3):e126-e35.
366. Rossi A, Aimar G, Audisio M, Bungaro M, Caglio A, Di Liello R, et al. Analysis of the adequacy of control arms in oncology randomised clinical trials published between 2017 and 2021: a meta-research study. *Eur J Cancer*. 2023;189:112920.
367. Samuel JN, Booth CM, Eisenhauer E, Brundage M, Berry SR, Gyawali B. Association of quality-of-life outcomes in cancer drug trials with survival outcomes and drug class. *JAMA Oncol*. 2022;8(6):879-86.
368. Gyawali B, Shimokata T, Honda K, Ando Y. Reporting harms more transparently in trials of cancer drugs. *BMJ*. 2018;363.
369. Sullivan R. Cancer medicines: a private vice for public benefit? *ecancermedicalscience*. 2024;18.
370. Aggarwal A, Nossiter J, Parry M, Sujenthiran A, Zietman A, Clarke N, et al. Public reporting of outcomes in radiation oncology: the National Prostate Cancer Audit. *Lancet Oncol*. 2021;22(5):e207-e15.
371. Tenhunen O, Lasch F, Schiel A, Turpeinen M. Single-arm clinical trials as pivotal evidence for cancer drug approval: a retrospective cohort study of centralized European marketing authorizations between 2010 and 2019. *Clin Pharmac Ther*. 2020;108(3):653-60.
372. Agrawal S, Arora S, Amiri-Kordestani L, de Claro RA, Fashoyin-Aje L, Gormley N, et al. Use of single-arm trials for us food and drug administration drug approval in oncology, 2002-2021. *JAMA Oncol*. 2023;9(2):266-72.
373. Van Stokkum RM, Bouwman J & Kamstra RJM. Generatieve AI in de Nederlandse zorg. 2024. Available from: https://publications.tno.nl/publication/34643183/U5tb8oyL/TNO-2024-R10662.pdf
374. Bosch D, Kuppen MC, Tascilar M, Smilde TJ, Mulders PF, Uyl-de Groot CA, et al. Reliability and efficiency of the CAPRI-3 metastatic prostate cancer registry driven by artificial intelligence. *Cancers*. 2023;15(15):3808.
375. Booth CM, Detsky AS. Why patients receive treatments that are minimally effective? *Nat Rev Clin Oncol*. 2019;16(1):3-4.

376. Tan MH, Thomas M, MacEachern MP. Using registries to recruit subjects for clinical trials. *Contemp Clin Trials*. 2015;41:31-8.
377. Patel TH, Rivera DR, Singh H, Kluetz PG. When less is more—reducing complexity in cancer trials. The *Lancet Oncol*. 2024;25(1):10-2.
378. McCall SJ, Imran M, Hemkens LG, Mc Cord K, Kwakkenbos L, Sampson M, et al. Reporting transparency and completeness in trials: Paper 4-reporting of randomised controlled trials conducted using routinely collected electronic records–room for improvement. *J Clin Epidemiol*. 2022;141:198-209.
379. Imran M, Mc Cord K, McCall SJ, Kwakkenbos L, Sampson M, Fröbert O, et al. Reporting transparency and completeness in trials: Paper 3-trials conducted using administrative databases do not adequately report elements related to use of databases. *J Clin Epidemiol*. 2022;141:187-97.
380. Derksen JW, May AM, Koopman M. The era of alternative designs to connect randomized clinical trials and real-world data. *Nat Rev Clin Oncol*. 2019;16(9):589-.
381. Vazquez MA, Oliver G, Amarasingham R, Sundaram V, Chan K, Ahn C, et al. Pragmatic Trial of Hospitalization Rate in Chronic Kidney Disease. *N Engl J Med*. 2024;390(13):1196-206.
382. Bibby AC, Walker S, Maskell NA. Are intra-pleural bacterial products associated with longer survival in adults with malignant pleural effusions? A systematic review. *Lung Cancer*. 2018;122:249-56.
383. Ehrenstein V, Hellfritzsch M, Kahlert J, Langan SM, Urushihara H, Marinac-Dabic D, et al. Validation of algorithms in studies based on routinely collected health data: general principles. *Am J Epidemiol*. 2024:kwae071.
384. García-Albéniz X, Hsu J, Bretthauer M, Hernán MA. Estimating the effect of preventive services with databases of administrative claims: reasons to be concerned. *Am J Epidemiol*. 2019;188(10):1764-7.

SUMMARY

General introduction

The prognosis of patients with cancer has improved the past decades. Yet, with more than 18 million new cases and 10 million deaths worldwide each year, the cancer epidemic remains a significant challenge. To further improve cancer outcomes, it is important to allocate adequate resources to prevention, early detection, and the improvement of treatment. This is not always straightforward, as pressure on healthcare and research budgets is increasing worldwide.

The costs of cancer care are rising at a rapid pace, driven partly by the growth in the number of patients in need of treatment, but also by the introduction of new health technologies (including medicines and medical devices) and the rising costs of these technologies. One approach to controlling health care costs without compromising outcomes is by critically assessing whether the costs of certain aspects of care proportionate to the value they deliver. In this context, medicines (and other technologies) have drawn significant attention, as there has been a notable increase in their number. However, their value is often uncertain at the time (reimbursement)decisions have to be made, and their costs are high.

Decision-making in oncology

The added value of health technologies is a concern for various stakeholders. Added value, in terms of efficacy and safety, is relevant for regulatory agencies that decide whether medicines and technologies can be allowed on the market. Value, in terms of the added effect and costs, is relevant to policymakers who need to decide on reimbursement. The value of a technology for individual patients is relevant for healthcare professionals who must decide which treatment fits a patient best. All these stakeholders have to make choices and they need reliable data to support their decision-making processes.

Traditionally, data from randomized controlled trials (RCTs) have served as the primary source of evidence for decision-making in oncology. However, evidence gaps often remain after these RCTs are conducted, particularly regarding the effects of a technology in broader patient populations and its effect on survival and quality of life. Moreover, in oncology there is an increasing number of technologies and medicines for which randomized studies have never been performed—and likely never will be. To address remaining questions, there is growing interest in using data from everyday clinical practice—real-world data (RWD). This is also driven by the increasing availability of such data. While RWD have been used alongside RCTs to provide insights into long-term safety

and to contextualize findings from RCTs, the different stakeholders believe these data offer greater potential for informing health care decisions. Yet, there is still uncertainty about how different types of RWD can be optimally used to generate reliable insights to support decision-making.

In this thesis, we explored the potential of different RWD sources to generate insights into cancer care and outcomes to support decision-making. We focused on retrospectively collected data, meaning data that were collected without the initial goal of answering a specific research question.

Real-World Data Sources

This thesis begins by introducing two important RWD sources in oncology that we used in this thesis: the Netherlands Cancer Registry (NCR) and administrative claims data (**Part 1 - Chapter 2**). The NCR is a nationwide registry that includes retrospectively collected data (patient, tumor, and treatment during the first year) from electronic health records (EHRs) of newly diagnosed patients with cancer in the Netherlands (coverage >95%). Vital status is obtained through linkage with the municipal personal records data. It offers several strengths that make it a valuable resource for research: an extensive history spanning over 30 years, rigorous quality ensured by trained data managers, and accessibility, as the data are available to researchers upon request. Administrative claims data originate from the financing of care: provided care is recorded by healthcare providers and billed to health insurers. Claims data offer strengths in their comprehensiveness regarding treatments and procedures delivered within institutions (**Chapters 2 and 3**), as well as their coded format, which facilitates detailed analysis. However, both data sources also have their limitations. For instance, the NCR contains limited information on cancer recurrences or treatments beyond the first line. Similarly, claims data typically lack patient, tumor, and outcome information.

To guide research based on RWD, and to assist decision-makers in judging the validity of such studies, several good research practice guidelines have been developed since 2009. Key themes covered by these guidelines are on information, selection and immortal time bias, as well as on confounding and cherry-picking in RWD. In **Chapter 3**, we assessed the extent to which such recommendations were implemented by researchers in studies evaluating the cost, or (cost-)effectiveness of systemic therapies for breast cancer using claims data. We found that many studies had methodological shortcomings. Specifically, very few studies provided clear rationales for the choice of potential confounding variables, validity of outcomes was almost never provided and there appeared to be a high risk of immortal time bias in many studies. Notably, the quality of studies remained unchanged after the guidelines were published in 2009.

Opportunities of Real-World Data

In **Part 2**, we present examples of studies based on RWD. Each of these studies provides evidence that cannot be obtained through RCTs, but can be relevant for decision-making. Specifically, we used RWD in a modeling study to assess the cost-effectiveness of the MammaPrint® (**Chapter 4**), a gene expression profiling test. This test was initially designed to tailor chemotherapy decisions, but recent evidence suggests that it may also guide choices regarding the use of endocrine therapy in patients with early-stage breast cancer. In this cost-effectiveness analysis, RWD from a cohort study were reused and provided insights into the quality of life and productivity losses of patients with breast cancer who suffer side effects from endocrine therapy. Despite the substantial literature on endocrine therapy, no studies were available that quantified these effects previously.

In **Chapter 5**, we used administrative claims data and data from an early access program to evaluate the utilization and access to new medicines for breast cancer. Our study demonstrated that these medicines (i.e., CDK4/6 inhibitors) were gradually adopted in clinical practice in the Netherlands, both before and after reimbursement became available. These findings expand upon existing literature, as previous studies on access to medicines have conducted less detailed analyses. For example, earlier studies have only looked at the reimbursement status (i.e., whether or not medicines are available in a country) or the 'waiting time' between EMA (European Medicines Agency) or FDA (Food and Drug Administration) approval and the first sales of a medicine. Our study suggests that results from such studies do not provide a complete picture of access and should therefore be interpreted with caution.

In **Chapters 6 and 7**, we used data from the NCR to evaluate the survival of patients with cancer. Both studies are unique in terms of sample sizes: nearly 2 million patients with solid cancers and over 11,000 patients with metastatic breast cancer; and follow-up periods of ±30 years and up to 14 years, respectively. In **Chapter 6**, we demonstrated that while the 1-year survival rate increased for most patients with *de novo* metastatic cancer, the 5-year rate showed gains for only a small number of them. Throughout the study period of almost 30 years, many novel medicines and new indications for existing medicines to treat metastatic disease were approved. Although causation cannot be assumed, as we did not have patient-level information regarding the use of medicines, there seems to be no correlation between the changes in survival rates and the number of new medicines approved. These findings may add to the discussion regarding the value of new medicines: are treatments worth their costs?

In **Chapter 7**, we showed that patients with *de novo* metastases survived longer than patients with metachronous metastases. This finding is relevant not only for clinicians to better inform patients about prognosis but also for the design and interpretation of RCTs. So far, stratification based on synchronous or metachronous metastasis has not been consistently performed in RCTs evaluating the efficacy of new medicines for metastatic breast cancer.

Challenges of Real-World Data

In **Chapter 8**, we reported the initial steps of an RWD study, guided by recently developed RWD guidelines from regulatory authorities and health technology assessment organizations. A key aspect of this study was assessing whether a given RWD set was 'fit for purpose' for a specified research question. We showed that there are several challenges in using hospital claims data for research, such as the fact that these data do not cover the entire patient pathway. In our experience, such challenges may not be fully identified before initiating a study. This study adds to the literature by providing insight into the content of hospital claims data and its (im)possibilities for future research.

In the discussing, we further reflect on the challenges of using RWD for research and decision-making, which we have (partially) identified during the processes of data collection, analysis, and reporting. One of the main challenge in using RWD for oncology research is the current data infrastructure. Due to the fragmented delivery of care, it is almost impossible to generate a 'fit-for-purpose' dataset that covers the entire patient care pathway (including patient and tumor characteristics, treatments delivered and outcomes), as shown in **Chapter 8**. Fortunately, several initiatives have been launched with the aim of improving data availability and its reuse for research and policy purposes.

A second challenge of RWD studies is that they are often limited in transparency regarding choices made during study design, data cleaning and analyses. This reduces their value for decision-making, as understanding such choices is important for the proper interpretation of the results. In response to these concerns, a range of tools, checklists, and frameworks have been developed recently to improve practice, reporting and interpretation of RWD studies. Our research (**Chapter 3**) showed that implementing these can be challenging, and future studies will need to determine whether these tools and checklists actually lead to improvements.

A third challenge concerns the methodology used in analyzing RWD. Currently, the question of when design and methodology of RWD-studies are sufficient to generate trustworthy results, remains unanswered. This is particularly concerning for studies aiming to assess the effectiveness of interventions using RWD and evaluate whether

the costs are proportionate to the effects. Achieving comparability between treatment groups is challenging without randomization; consequently, it remains uncertain whether the effects of different technologies can be adequately compared.

Overarching Challenges

Data in oncology can only contribute to better patient care and outcomes if they help reduce uncertainty around the most pressing clinical, policy, and regulatory questions. Unfortunately, the current evidence-generation landscape in oncology appears to be suboptimal. Society demands that evidence of efficacy in oncology be provided by commercial companies, which appears to influence the way studies are designed, performed and how results are communicated. Additionally, regulatory agencies have relaxed evidentiary standards to expedite access. As a result, a high level of uncertainty often remains regarding the added clinical benefits and economic value of technologies at the time of approval.

Towards meaningful Data

Many directions to improve evidence in oncology have been proposed in the literature. Based on our experiences with collecting and analyzing RWD for this thesis, we agree to a couple of these. The first direction focuses on raising evidentiary standards of regulatory agencies and improve the certainty of the effectiveness and economic value of technologies at the time of approval. This includes mandating registration studies that include appropriate comparator arms and measure outcomes that are relevant to patients. The second direction focuses on the collection, analysis, and reporting of RWD. Transparency in this process is essential, as is compliance with RWD guidelines. Finally, it is necessary to ensure that structural funding is available to improve and maintain the data infrastructure and to generate datasets that provide a comprehensive overview of the entire care pathway and are of sufficient quality to obtain valid results.

Improvements in these areas will hopefully lead to stronger evidence in oncology, which can ultimately contribute to better care for patients.

Concluding remarks

In this thesis, we showed that RWD can provide valuable insights into treatment utilization, long-term outcomes and patient and disease characteristics. But we also identified complex barriers within the current data infrastructure and legal framework that hamper the optimal use of RWD. Promising initiatives to address these challenges are on their way, though they are still at early stages of development. Meanwhile, it is the responsibility of stakeholders working with RWD or using RWD in decision-making to use RWD wisely.

SAMENVATTING (IN DUTCH)

Algemene introductie

De prognose van kanker is de afgelopen decennia verbeterd. Toch blijft kanker, met jaarlijks meer dan 18 miljoen nieuwe diagnoses en 10 miljoen sterfgevallen wereldwijd, een belangrijk probleem voor de volksgezondheid. Om de uitkomsten van patiënten met kanker verder te verbeteren, is het belangrijk om te blijven investeren in preventie, vroege detectie en verbetering van de behandeling. Dit is niet altijd gemakkelijk, gezien de toenemende druk op de zorg- en onderzoeksbudgetten wereldwijd.

De kosten van de zorg voor kanker stijgen snel, enerzijds door de groei van het aantal patiënten dat behandeling nodig heeft, en anderzijds door de introductie van nieuwe gezondheidstechnologieën (inclusief medicijnen). Een manier om deze kosten te beheersen zonder dat dit ten koste gaat van de gezondheid, is door kritisch te kijken of de kosten van bepaalde zorg in verhouding staan tot de waarde die het oplevert. Binnen de oncologie hebben medicijnen en andere gezondheidstechnologieën in dit opzicht veel aandacht getrokken, doordat het aantal sterk is toegenomen, de toegevoegde waarde voor patiënten op het moment van een (vergoedings)beslissingen vaak onzeker is, terwijl de kosten hoog zijn.

Besluitvorming in de oncologie

De toegevoegde waarde van gezondheidstechnologieën (inclusief medicijnen) is belangrijk voor verschillende *stakeholders* binnen de oncologie, die beslissingen moeten nemen over de inzet van deze technologieën. Zo is de toegevoegde waarde, in termen van effectiviteit en veiligheid, relevant voor regelgevende instanties die bepalen of technologieën op de markt mogen worden toegelaten. De toegevoegde waarde, in termen van de effecten, kosten en kosteneffectiviteit, is van belang voor beleidsmakers die moeten beslissen of technologieën vanuit de basisverzekering vergoed kunnen worden. De waarde van een technologie voor een individuele patiënt is relevant voor zorgverleners en de desbetreffende patiënt om samen te kunnen beslissen welke behandeling het beste is. Al deze *stakeholders* moeten keuzes maken, en hiervoor hebben zij betrouwbare gegevens nodig.

Traditioneel gezien vormden gerandomiseerde studies (RCT's) de meest betrouwbare bron van bewijs voor besluitvorming binnen de oncologie. Echter, steeds vaker blijven er na het uitvoeren van deze studies vragen over. Zo is het vaak onduidelijk wat het effect van een medicijn of andere technologie is in een bredere, meer heterogene patiëntpopulatie. Bovendien is het effect op overleving en kwaliteit van leven niet altijd aangetoond. Daarnaast zijn er binnen de oncologie steeds meer voorbeelden van

technologieën waarvoor nooit een gerandomiseerde studie is uitgevoerd is of zal worden uitgevoerd. Om overgebleven vragen te beantwoorden is er groeiende interesse in het gebruik van gegevens uit de dagelijkse klinische praktijk-*real-world data* (RWD). Dit wordt ook gestimuleerd door de toenemende beschikbaarheid van deze gegevens. Tot op heden werden RWD al gebruikt om inzichten te verschaffen in bijvoorbeeld de veiligheid op de lange termijn en om context te bieden ter aanvulling op de resultaten van RCT's. Tegenwoordig zien de verschillende *stakeholders* een bredere rol voor RWD. Echter, bestaat er nog onduidelijkheid over hoe verschillende typen RWD optimaal kunnen worden benut om betrouwbare inzichten te genereren ter ondersteuning van besluitvorming.

In dit proefschrift hebben we onderzocht of en op welke manier verschillende bronnen van RWD waardevolle inzichten kunnen bieden ter ondersteuning van besluitvorming in de oncologie. We hebben ons gericht op retrospectief verzamelde gegevens, dat wil zeggen gegevens die niet verzameld zijn om een gespecificeerde onderzoeksvraag te beantwoorden.

Real-world data bronnen

Dit proefschrift begint met de introductie van twee belangrijke bronnen van RWD in de oncologie: de Nederlandse Kanker Registratie (NKR) en declaratiedata (**Deel 1 – Hoofdstuk 2**). De NKR is een landelijk register die retrospectief verzamelde gegevens (patiënt, tumor gegevens, en behandeling in het eerste jaar) bevat geregistreerd vanuit elektronische patiëntendossiers (EPD's) van nieuw gediagnosticeerde patiënten met kanker (dekking >95%). Overlijdens gegevens worden verkregen door koppeling met de gemeentelijke basisadministratie. Het is een waardevolle bron voor onderzoek doordat het in 1989 is opgezet, wat betekent dat er lange follow-up beschikbaar is. Bovendien zijn de gegevens van goede kwaliteit en toegankelijk voor onderzoekers. Zorgdeclaratiedata kennen hun oorsprong in de bekostiging van zorg: verleende zorg wordt door zorgverleners geregistreerd en gedeclareerd bij zorgverzekeraars. Sterke punten van deze data voor onderzoek zijn dat ze erg compleet zijn wat betreft gedeclareerde zorgverrichtingen in de betreffende instelling (**Hoofdstuk 2 en 3**), dat ze reeds beschikbaar zijn in gecodeerde vorm en dat ze vaak grote groepen patiënten omvatten. Beide gegevensbronnen hebben ook hun beperkingen. Zo bevat de NKR beperkte informatie over kankerrecidieven of behandelingen na de eerste lijn. Evenzo ontbreekt in declaratiedata vaak informatie over patiënt- en tumorkarakteristieken en uitkomsten.

Sinds 2009, zijn er verschillende richtlijnen ontwikkeld ter ondersteuning van onderzoek op basis van RWD. Belangrijke onderwerpen die aan de orde komen in deze richtlijnen zijn *information bias*, *selectiebias, immortal time bias*, *confounding* en *cherry picking*. In

Hoofdstuk 3 hebben we onderzocht in welke mate dergelijke aanbevelingen werden opgevolgd door onderzoekers die gebruik maakten van declaratiedata om de (kosten) effectiviteit van systemische therapieën voor borstkanker te onderzoeken. In vele studies zagen we methodologische tekortkomingen. Zo werden er vaak geen duidelijke redenen genoemd voor de keuze van mogelijke *confounding* variabelen, was de validiteit van uitkomsten bijna nooit gerapporteerd en leek er een hoog risico op *immortal time bias*. Opvallend was dat de kwaliteit van studies niet verbeterde na de publicatie van de (in 2009 gepubliceerde) richtlijnen.

Mogelijkheden van real-world data

In **Deel 2** van dit proefschrift presenteren we verschillende voorbeelden van studies gebaseerd op RWD. Elke van deze studies levert bewijs dat niet verkregen wordt met RCTs, maar wel relevant kan zijn voor besluitvorming. In **Hoofdstuk 4** hebben we RWD gebruikt in een modelmatige kosteneffectiviteitsanalyse van de MammaPrint® (een genexpressieprofileringstest). Deze test werd oorspronkelijk ontworpen om keuzes te maken over de inzet van chemotherapie. Recent bewijs toont aan dat deze test mogelijk ook kan worden ingezet voor de keuze met betrekking tot de inzet van endocriene therapie bij patiënten met vroegstadium borstkanker. In onze studie hebben we RWD uit een cohortstudie hergebruikt om de impact van bijwerkingen door endocriene therapie op de kwaliteit van leven en productiviteit te bepalen. Dergelijke gegevens waren niet beschikbaar in (klinische) studies over endocriene therapie, ondanks dat deze middelen al decennialang worden ingezet.

In **Hoofdstuk 5** hebben we declaratiedata en gegevens uit een *early access*-programma gebruikt om de inzet en de toegang tot nieuwe medicijnen voor borstkanker te evalueren. Onze studie toonde aan dat deze medicijnen (CDK4/6-remmers) geleidelijk in de klinische praktijk in Nederland werden geïmplementeerd, zowel vóór als nadat er vergoeding beschikbaar kwam. Dit vormt een aanvulling op de literatuur, aangezien eerdere studies naar de toegankelijkheid van medicijnen aanzienlijk minder gedetailleerde analyses hebben uitgevoerd. Zo is in eerder studies alleen gekeken naar de status van vergoeding (d.w.z. het wel of niet beschikbaar zijn van geneesmiddelen in een land) of naar de 'wachttijd' tussen EMA (*European Medicines Agency*)- of FDA (*Food and Drug Administration*)-registratie en de eerste uitgifte van een geneesmiddel. Onze studie suggereert dat dergelijke studies niet het volledige beeld geven van toegang en dus voorzichtig moeten worden geïnterpreteerd.

In **Hoofdstukken 6 en 7** hebben we gegevens uit de NKR gebruikt om de overleving van patiënten met kanker te evalueren. Beide studies zijn uniek vanwege hun omvang, dat wil zeggen: bijna 2 miljoen patiënten met solide tumoren (**Hoofdstuk 6**) en meer

dan 11.000 patiënten met uitgezaaide borstkanker (**Hoofdstuk 7**). Bovendien waren in beide studies de follow-upperioden lang, respectievelijk ~30 jaar en ~14 jaar. In **Hoofdstuk 6** hebben we laten zien dat de 1-jaarsoverleving voor vrijwel alle typen *de novo* gemetastaseerde kanker is toegenomen, terwijl de 5-jaarsoverleving slechts bij een klein aantal verbeterde. Gedurende de onderzoeksperiode van circa 30 jaar werden veel nieuwe medicijnen en indicaties voor bestaande medicijnen voor uitgezaaide kanker goedgekeurd. Ondanks dat er geen informatie op patiënt niveau was ten aanzien van het gebruik van geneesmiddelen, en met de beschikbare data geen uitspraak kan worden gedaan over een mogelijk oorzaak-gevolgverband, lijkt het erop dat er geen relatie bestaat tussen het aantal nieuwe geneesmiddelen dat beschikbaar is gekomen en de veranderingen in overleving. Deze bevindingen kunnen bijdragen aan de discussie over de waarde van nieuwe medicijnen op brede schaal: zijn de behandelingen hun kosten waard?

In **Hoofdstuk 7** hebben we laten zien dat patiënten met *de novo* metastasen langer leefden dan patiënten met metachrone metastasen. Deze bevinding is zowel relevant voor clinici, om patiënten beter te informeren over hun prognose, als voor het ontwerp en de interpretatie van RCT's. Tot nu toe wordt stratificatie op basis van *de novo*- of metachrone metastasering niet vaak toegepast in RCT's die de effectiviteit van nieuwe medicijnen voor uitgezaaide borstkanker evalueren.

Belemmeringen bij het gebruik van real-world data

In **Hoofdstuk 8** hebben we de eerste stappen van een RWD-studie beschreven aan de hand van recent ontwikkelde RWD-richtlijnen van regelgevende instanties en *health technology assessment* (HTA) organisaties. Een belangrijk onderdeel van deze studie was het beoordelen van de geschiktheid van een RWD-set ('*fit-for-purpose*') voor een specifieke onderzoeksvraag. We hebben aangetoond dat er tekortkomingen zijn aan het gebruik van declaratiedata van ziekenhuizen voor onderzoek, omdat deze niet het volledige zorgtraject van een patiënt bestrijken. Dergelijke tekortkomingen zijn niet altijd volledig te overzien voordat een studie van start gaat. Onze studie biedt inzicht in de inhoud van ziekenhuis declaratiedata en de (on)mogelijkheden hiervan voor toekomstig onderzoek.

In de discussie gaan we dieper in op belemmeringen bij het gebruik van RWD voor onderzoek en besluitvorming, die we (deels) geïdentificeerd hebben bij het verzamelen, analyseren en rapporteren. De huidige data-infrastructuur is een van de grootste tekortkomingen wat dit betreft. Doordat zorg wordt geleverd in verschillende instellingen, is het vrijwel onmogelijk om een dataset te creëren die het volledige zorgtraject van de patiënt omvat (inclusief patiënt- en tumorkarakteristieken, behandelingen en uitkom-

sten) en *'fit-for-purpose'* is, zoals beschreven in **Hoofdstuk 8**. Gelukkig zijn er verschillende initiatieven gestart die als doel hebben om de beschikbaarheid van data en het hergebruik ervan voor onderzoek en beleid te verbeteren.

Een tweede belemmering van huidige RWD-studies is dat ze vaak onvoldoende transparantie bieden over de gemaakte keuzes tijdens de studieopzet, dataschoning en analyses. Dit vermindert hun waarde voor besluitvorming, aangezien inzicht in dergelijke keuzes belangrijk is voor een juiste interpretatie van de resultaten. Recentelijk zijn verschillende tools en checklists ontwikkeld om de uitvoering, rapportage en interpretatie van RWD-studies te verbeteren. Ons onderzoek (**Hoofdstuk 3**) liet zien dat de implementatie hiervan een uitdaging kan zijn en toekomstig onderzoek zal moeten laten zien of de tools en checklists tot daadwerkelijke verbeteringen leiden.

Een derde belemmering betreft de methodologie die wordt gebruikt bij de analyse van RWD. Momenteel is onduidelijk wanneer een studieopzet en methodologie van RWD-studies *goed genoeg* zijn om betrouwbare resultaten te genereren. Dit is met name van belang bij studies die de effectiviteit van interventies proberen vast te stellen met RWD en beoordelen in hoeverre kosten in verhouding staan tot de effecten. Zonder randomisatie kan de vergelijkbaarheid van behandelgroepen niet gegarandeerd worden en is het onzeker of de effecten van technologieën goed vergeleken kunnen worden.

Overkoepelende belemmeringen

Data in de oncologie kunnen alleen bijdragen aan betere patiëntenzorg en -uitkomsten als ze waardevolle klinische en beleidsmatige inzichten bieden. Helaas lijkt de huidige wijze waarop bewijs wordt gegenereerd in de oncologie niet optimaal te zijn. De maatschappij verlangt dat bewijs naar de effectiviteit van technologieën (inclusief medicijnen) wordt geleverd door de ontwikkelaars en hiermee commerciële bedrijven wat invloed lijkt te hebben op de wijze waarop studies worden ontworpen, uitgevoerd en hoe de resultaten worden gecommuniceerd. Daarnaast hebben regelgevende instanties in de afgelopen decennia de bewijsstandaarden versoepeld om de toegang tot geneesmiddelen te versnellen. Een gevolg hiervan is dat er vaak een hoge mate van onzekerheid bestaat over de toegevoegde waarde van technologieën op het moment dat deze toegelaten worden tot de markt.

Van data naar betekenis

In de literatuur zijn verschillende oplossingen voorgesteld om de kwaliteit van bewijs binnen de oncologie te verbeteren. Op basis van onze ervaringen met het verzamelen en analyseren van RWD voor dit proefschrift ondersteunen wij een aantal hiervan.

De eerste oplossing richt zich op het aanscherpen van de criteria van regelgevende instanties voor bewijs van toegevoegde waarde. Bij voorkeur zou bewijs gebaseerd moeten zijn op gerandomiseerde studies, met een controlegroep die de standaardzorg in de praktijk weerspiegelt en uitkomstmaten meet die relevant zijn voor patiënten. De tweede oplossing richt zich op het verzamelen, analyseren en rapporteren van RWD. Transparantie in dit proces is noodzakelijk, evenals de naleving van de richtlijnen voor RWD. Tot slot is het noodzakelijk dat structurele financiering beschikbaar komt om de datainfrastructuur te verbeteren en te onderhouden, en om datasets te genereren die het volledige zorgtraject inzichtelijk maken en van voldoende kwaliteit zijn om valide resultaten te verkrijgen.

Verbeteringen op deze gebieden zullen hopelijk leiden tot sterker bewijs in de oncologie, wat zo uiteindelijk bij kan dragen en aan betere zorg voor patiënten.

Conclusie

In dit proefschrift hebben we de waarde van RWD voor besluitvorming in de oncologie onderzocht. Wij hebben laten zien dat RWD waardevolle inzichten kan bieden, onder andere over de inzet van behandelingen in de praktijk, langetermijnresultaten en patiënt- en ziektekarakteristieken. Maar we hebben ook belemmeringen geïdentificeerd, met name in de huidige data infrastructuur. Veelbelovende initiatieven om deze uitdagingen aan te pakken zijn in ontwikkeling. Ondertussen is het de verantwoordelijkheid van degene die met RWD werken en deze inzetten in de besluitvorming, om RWD verstandig in te zetten.

PHD PORTFOLIO

PhD Training

C1 Business English. The Square Mile Language Training. Interactive online course. 2022.

Survival Analysis for Clinicians. Netherlands Institute for Health Sciences Rotterdam (NIHES), the Netherlands. 2020.

Advanced Methods in Epidemiology. Leiden University Medical Centre. Leiden, The Netherlands. 2019.

Advanced Methods for Addressing Selection Bias in Real-World Effectiveness and Cost-Effectiveness Studies. ISPOR short course: Copenhagen, Denmark. 2019.

Data visualization. Erasmus School of Health Policy & Management. Rotterdam, the Netherlands. 2019.

Introduction to the clinical and fundamental oncology. Netherlands Association for Medical Oncology (NVvO). 2019.

Meta-analysis and network meta-analysis. Institute for Medical Technology Assessment (iMTA) & University of Ioannina (Greece). Rotterdam, the Netherlands. 2018.

Various short courses provided by IKNL: introduction to IKNL, data registration, data safety, introductory course to cancer (general) and specific for breast cancer (Oncologisch spectrum). Utrecht, the Netherlands. 2018-2022.

Basic didactics. Risbo. Rotterdam, the Netherlands. 2018.

Teaching

Health Technology Assessment, Master program Health Economics Policy and Law, Erasmus University Rotterdam. Instructor computer lab. 2018, 2019 & 2021.

Bachelor and master theses, Bachelor and master program Health Economics Policy and Law, Erasmus University Rotterdam. Supervisor, 2018-2021.

How do we keep healthcare affordable? Bachelor program Medicine, Erasmus Medical Centre Rotterdam. Tutor. 2020.

Quality and efficiency in health care, bachelor program Health Economics Policy and Law, Erasmus University Rotterdam. Tutor. 2018.

Presentations

Podium presentation: *Challenges and opportunities with acquiring and analyzing claims data*. BOOG symposium: Real world data. Utrecht, the Netherlands. 2022.

Online interview: *Access to CDK4/6 inhibitors seems good in the Netherlands.* Oncologie.nu. 2021.

Poster: *Accessibility of CDK4/6 inhibitors for breast cancer patients in the Netherlands.* San Antonio Breast Cancer Symposium (SABCS), San Antonio, United States. 2021.

Poster: *The Novo Metastatic Cancer: An overview of changes in survival and novel therapies.* European Network of Cancer Registries (ENCR) conference. Hybrid. 2020.

Poster: *29 Years of breast cancer: A timeline of progress in the Netherlands*. San Antonio Breast Cancer Symposium (SABCS), San Antonio, United States. 2019. (no attendance)

Poster: *Quality of studies based on claims data: A systematic review of (cost)effectiveness studies of systemic therapies in breast cancer.* ISPOR annual European congress: Copenhagen, Denmark. 2019.

In the Media

Front page article in de Volkskrant and an article in Trouw following the publication of our paper 'Changes in survival in *de novo* metastatic cancer in an era of new medicines'

"Nog maar weinig patiëntengroepen hebben baat bij nieuwe dure kankermedicijnen." De Volkskrant. 29-03-2023

"Dure kankermedicijnen hebben een beperkt effect op de levensverwachting." Trouw. 30-03-2023.

LIST OF PUBLICATIONS

Scientific publications

Luyendijk M, Vernooij RWM, Blommestein HM, Siesling S, Uyl-de Groot CA. Assessment of Studies Evaluating Incremental Costs, Effectiveness, or Cost-Effectiveness of Systemic Therapies in Breast Cancer Based on Claims Data: A Systematic Review. Value in Health. 2020; 23(11):1497–1508

Luyendijk M, Blommestein HM, Uyl-de Groot CA, Siesling S, Jager A. Regulatory Approval, Reimbursement, and Clinical Use of Cyclin-Dependent Kinase 4/6 Inhibitors in Metastatic Breast Cancer in the Netherlands. JAMA Network Open. 2023;6(2):e2256170.

Luyendijk M, Visser O, Blommestein HM, de Hingh IHJT, Hoebers FJP, Jager A, Sonke GS, de Vries EGE, Uyl-de Groot CA, Siesling S. Changes in survival in de novo metastatic cancer in an era of new medicines. JNCI: Journal of the National Cancer Institute, 2023(0;0)

Luyendijk M, Jager A, Buijs SM, Siesling S, Uyl de Groot CA, Blommestein HM. Cost Effectiveness Analysis of MammaPrint® to Guide the Use of Endocrine Therapy in Patients with Early Stage Breast Cancer. PharmacoEconomics

de Maar*, **Luyendijk M***, Suelmann M, van der Kruijssen DEW, Elias SG, Siesling S, van der Wall E. Comparison between de novo and metachronous metastatic breast cancer: the presence of a primary tumour is not the only difference —a Dutch population-based study from 2008 to 2018. Breast Cancer Research and Treatment (2023) 198:253–264.
*contributed equally

Westwood M, Corro Ramos I, Lang S, **Luyendijk M**, Zaim R, Stirk L, Al M, Armstrong N, Kleijnen J. 2017. Faecal immunochemical tests to triage patients with lower abdominal symptoms for suspected colorectal cancer referrals in primary care: a systematic review and cost-effectiveness analysis. Health Technol Assessment.

Redekop WK*, Lenk EJ*, **Luyendijk M**, Fitzpatrick C, Niessen L, Stolk WA, Tediosi F, Rijnsburger AJ, Bakker R, Hontelez JAC, Richardus JH, Jacobson J, de Vlas SJ & Severens JL 2017. The Socioeconomic Benefit to Individuals of Achieving the 2020 Targets for 5 Preventive Chemotherapy Neglected Tropical Diseases. PLoS Negl Trop Dis.
*Contributed equally

Lenk EJ*, Redekop WK*, **Luyendijk M***, Rijnsburger AJ and Severens JL 2016. Productivity Loss Related to Neglected Tropical Diseases Eligible for Preventive Chemotherapy: A Systematic Literature Review. PLoS Negl Trop Dis.
#Contributed equally

de Vlas SJ, Stolk WA, le Rutte EA, Hontelez JA, Bakker R, Blok DJ, Cai R, Houweling TA, Kulik MC, Lenk EJ, **Luyendijk M**, Matthijsse SM, Redekop WK, Wagenaar I, Jacobson J, Nagelkerke NJ, Richardus JH 2016. Concerted Efforts to Control or Eliminate Neglected Tropical Diseases: How Much Health Will Be Gained? PLoS Negl Trop Di

Other publications

Eijkelboom A, Luyendijk M, van Maaren M, de Munck L, Schreuder K, Siesling S, Verloop J. 2020. Borstkanker in Nederland: trends 1989-2019 gebaseerd op cijfers uit de Nederlandse Kankerregistratie.

ABOUT THE AUTHOR

Marianne Luyendijk (1985) holds a master's degree in Health economics, Policy, and Law, which she completed in 2013. She has worked as a researcher at the Erasmus School of Health Policy & Management (ESHPM) and the Institute for Medical Technology Assessment (iMTA). In 2018, she began her PhD, focusing on the use of real-world data to inform health decisions, in collaboration with the Netherlands Comprehensive Cancer Organization (IKNL). Currently, she works as a pharmacoeconomic adviser at Zorginstituut Nederland (Healthcare Institute Netherlands), where she continues to contribute to health policy.

DANKWOORD

Data kan pas bijdragen aan het verbeteren van de zorg voor patiënten met kanker als het wordt vertaald naar waardevolle informatie en inzichten. Om dit te kunnen doen, moet data worden begrepen, geschoond, geanalyseerd, geïnterpreteerd en uiteindelijk gevisualiseerd en gecommuniceerd. Dit is een proces wat geen enkele onderzoeker alleen kan, maar waarbij hulp en inzichten nodig zijn van vele anderen.

De personen die ik wil bedanken voor hun hulp bij de totstandkoming van dit proefschrift zijn mijn (co)promoteren: Hedwig, Agnes, Sabine en Carin. Dank voor jullie enthousiasme, inspiratie, wetenschappelijke input op verschillende vlakken (medisch, epidemiologisch en gezondheidseconomisch) en jullie kritische blik. Maar dank ook voor de vrijheid die jullie mij gaven in het schrijven van artikelen, en voor het vertrouwen en de steun dat ik hierbij kreeg – om soms toch door te gaan als anderen niets in mijn plannen zagen, of juist te stoppen als het goed genoeg was. Tot slot ook: dank voor jullie geduld!

Dank ook aan het Integraal Kankercentrum Nederland (IKNL) en het Erasmus School of Health Policy & Management (ESPHM) dat ik deel uit mocht maken van deze mooie organisaties en gebruik mocht maken van de mooie data uit de NKR, en aan alle coauteurs die meeschreven aan mijn artikelen, aan alle collega's binnen en buiten deze organisaties met wie ik heb samengewerkt, gebrainstormd of juist alleen maar heb gekletst over alles behalve werk.

Als laatste dank aan de liefste mannen in mijn leven: Ricardo, Jurre & Len; mijn familie; vrienden en vriendinnen. Bedankt dat jullie er zijn! Samen met jullie is alles leuker.