Adoption and implementation of AI-driven Clinical Decision Support Systems in Cancer Care

A Case Study on Mammaprint in the Dutch Healthcare System using an Institutional Actor Analysis

EPA2942: Master Thesis

Simone Smits

Delft University of Technology

TU Delft

Adoption and implementation of AI-driven Clinical Decision Support Systems in Cancer Care

A Case Study on Mammaprint in the Dutch Healthcare System using an Institutional Actor Analysis

by

Simone Smits

to obtain the degree of Master of Science in Engineering & Policy Analysis
At the Faculty of Technology, Policy and Management
of the Delft University of Technology,
to be defended publicly on Monday November 20th, 2023 at 13:00.

Student number:	4550315	
Thesis committee:	Dr. S. Hinrichs-Krapels,	TU Delft, First supervisor
	Dr. I. Grossmann,	TU Delft, Second Supervisor
	Dr. P.W.G. Bots,	TU Delft, Chair

Cover: Generated using DALL·E 2 - OpenAI

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Acknowledgements

Simone Smits
Delft, November 2023

I would like to thank my graduation committee, dr. Saba Hinrichs-Krapels, dr. Irene Grossmann and dr. Pieter Bots. All three supervisors gave invaluable input during the thesis project. I especially would like to thank Saba for all the thesis meetings and additional support. I really appreciate that you always took the time to help where possible and give feedback. Your critical eye helped me a lot. During the writing process I learned a lot from you about how to properly structure research findings, which will also be very helpful in the future. I would like to thank Pieter for discussing insights about actors and institutions and how this can be applied to my research. You were always willing to take the time and think along for this. I would like to thank Irene for her input from a medical point of view, which was incredibly valuable to know before I started the case study interviews.

I would like to thank all the interviewees for their willingness to participate, to help me further in their networks and to provide valuable insights. Without these interviews, this research would not have been possible. In addition to being very valuable input for my research, I personally found the conversations very interesting and I really enjoyed the engagement.

I would like to express my heartfelt gratitude to my family and friends, who have been my support throughout this academic journey. It was and still is my indispensable outlet to laugh with you, to row and do sports with you, to do coffees and to catch up.

Executive Summary

The Dutch health care system is currently unsustainable. The accessibility, quality and affordability of Dutch the healthcare system is under pressure. This is particularly evident in cancer care as more and more people are getting cancer and treatment costs are rising. The current focus is therefore on adequate care to prevent stalling of cancer care. To improve this proven innovations should be quickly and effectively adopted. It is proposed to achieve this by providing care that is proven effective, preventing over-treatment, focusing on quality of life and by engaging the patient in deciding upon the treatment. On these issues, AI-driven Clinical Decision Support Systems (AI-CDSS) could potentially provide solutions. However, despite the extensive literature about development efforts of AI-CDSS, their adoption lags behind. The literature did however not explain why AI-CDSS adoption in cancer care lags behind. This leads to a main research question that will fill the gap of missing knowledge on adoption and implementation of cancer specific AI-CDSS:

Why does the adoption and implementation of AI-CDSS in cancer care fail?

The adoption and implementation of AI-CDSS in cancer care fail due to an institutional void of reimbursement by basic health insurance. Institutional voids refer to the absence or inadequacy of supportive structures, regulations, and frameworks. In the case of Mammaprint, a molecular diagnostic AI-CDSS in breast cancer, this institutional void manifested within the interpretation of an official criterion for reimbursement, the state of science and practice (SWP). The SWP criterion assesses if clinical utility, which addresses whether the patient actually experiences health benefits as a result of using the test, is proven. Mammaprint uses molecular profiling and is seen as a diagnostic test. There are no definitive requirements for inclusion of diagnostic AI-CDSS and therefore the SWP criterion is open to interpretation. Here the policy analysts and medical specialists within the reimbursement arena disagreed about the interpretation and fulfilment of the SWP criterion. Disagreements were found on the level of burden of proof and study design needed and on the trade-off between the quality of life and survival. The significant discrepancies between the medical specialists and policy analysts in case of Mammaprint are striking.

The use of AI does not play a role in the assessment for reimbursement in the specific case of Mammaprint and that Mammaprint uses AI is relatively unknown. It appears that in the medical domain, the prevailing concept of an AI-CDSS typically involves a system employing image analysis methods. AI-CDSS is also not one term when it comes to reimbursement. It was found that reimbursement by basic health insurance pathways for other AI-CDSS in the Netherlands vary significantly based on the underlying technique, namely molecular diagnostics or image analysis and based on the use in hospital or in screening. However, it has been observed that just as Mammaprint, these other AI-CDSS also encounter institutional voids in the realm of reimbursement. There is an expansive context of policy making where AI-CDSS as innovation challenges and disrupts established institutional frameworks.

In addition, several contextual factors appeared that influence the adoption and implementation of AI-CDSS in cancer care, making it likely that adoption and implementation may be stalled in a lot of other phases besides the reimbursement phase. For example, it is difficult to develop AI-CDSS within

a hospital because funding and product development expertise are lacking, there are regulatory impediments for obtaining certification, and in the event that certification is achieved, it proves insufficient for marketing the product because additional evidence seems to be considered essential for the marketing phase. It also became clear that the reimbursement arenas for Mammaprint look very different in other countries, with particular distinction between Europe with a predominantly public and the U.S. with a predominantly private healthcare market.

The research is conducted through a case study focusing on Mammaprint using interviews and grey literature as data input. The selection of Mammaprint as focal point of the case study is well-founded; its Dutch origins facilitate easier access to pertinent interviewees, and its challenges with adoption make it a compelling case. Additionally there is a very similar AI-CDSS, Oncotype, which enables some meaningful comparative analysis. An institutional actor analysis framework, employing an actor and institutions lens, has been customized to analyze the case study data. This allowed to identify critical arenas, map formal institutions, identify actors within the arena, reveal informal institutions, roles and interactions, assess power and interest of actors and synthesise the findings. A total of 18 interviews with 19 interviewees has been conducted and recorded. All interviewees were highly relevant and sampled using an institutional actor analysis and snowball sampling. The interview data were analyzed by noting relevant findings for each interview and by doing further thematic analyses of these findings. Additionally, throughout the research a sequential framework of phases of device adoption is developed and employed to categorize and structure findings from both literature and interviews.

The study is relevant as it showed that reimbursement can act as a limiting factor for the adoption and implementation of AI-CDSS in cancer care as the pathway for reimbursement by basic health insurance in the Netherlands has institutional voids for both molecular diagnostic as well as image analysis-based AI-CDSS used in-hospital.

For Mammaprint, very specific factors were identified for which consensus between the ZiN and medical specialists involved in the reimbursement can be sought in order to clarify the reimbursement pathway. Firstly, there is a need for consensus on the level of burden of proof required. This includes determining when retrospective or prospective studies are appropriate, as well as the follow-up duration. If a high level of burden of proof is required to maintain high quality standards, current policies for temporary admission could potentially offer a solution as this approach allows for data collection on clinical utility and ensures accessibility. Secondly, consensus is needed on the study design necessary and feasible to establish clinical utility. This includes discussing ethical considerations for randomization, determining requirements for measurements related to quality of life and defining outcome measures needed to proof clinical utility. Thirdly, consensus must be sought regarding the trade-off between quality of life and survival. Seeking consensus on these factors could fill existing institutional voids and thereby probably also partly clarifies the pathway for reimbursement of other AI-CDSS.

Additionally, it is found that there are current initiatives to clarify the reimbursement pathway of molecular diagnostics and image analysis-based AI-CDSS and it is advised to pursue these.

Lastly, as it becomes clear the field of AI-CDSS seems to be very much in development and that there is significant interest in employing image analysis methods for AI-CDSS in radiology and pathology, it is advised to determine a vision for the role of AI-CDSS within the current existing policy framework to make cancer care future-proof, called ’adequate cancer care’. This as neglecting to set forth a clear vision and strategy of this seemingly upcoming innovation, AI-CDSS could inadvertently impede, rather than improve, the sustainability of cancer care.

Contents

Nomenclature

Abbreviations

Abbreviation	Definition	Dutch Translation
AI-CDSS	Artificial Intelligence (driven) Clinical Decision Support System	
CieBAG	Add-on Medicines Evaluation Committee	Commissie Beoordeling Add-on Geneesmiddelen
CtP	Commit to provide	Plegen te bieden
CvZ	Board for Health Insurance	College voor Zorgverzekeringen
IAA	Institutional Actor Analysis	
NKI-AVL	Netherlands Cancer Institute-Anthonie van Leeuwenhoek	Nederland Kanker Instituut-Anthonie van Leeuwenhoek
RCT	Randomized Controlled Trial	
SWP	State of Science and Practice	Stand van Wetenschap en Praktijk
ZiN	Healthcare institute Netherlands	Zorginstituut Nederland
ZN	Health Insurers Netherlands	Zorgverzekeraars Nederland

1

Introduction

1.1. Problem Context

The Dutch health care system is currently unsustainable. The accessibility, quality and affordability of the Dutch healthcare system are under pressure [1]. Healthcare costs are currently more than 11% of GDP [2] and continue to grow rapidly over the years [1]. It is predicted that by 2040, 1 out of 4 people will have to work in healthcare to meet the increasing demand for care [1], while this is currently 1 out of 7 [3]. To raise political awareness and seek remedies, the national healthcare agreement was formulated [4]. Within this framework, both the healthcare sector and the government are collaboratively examining challenges and proposing solutions. There is special attention to four areas of care, including cancer care for which there is a special trajectory, namely 'adequate cancer care' [5, 6], to identify opportunities and impediments to future-proof care for people with cancer [3, 5, 6] as care for people with cancer is currently under pressure and in danger to get stalled [5]. Cancer care costs are a significant portion of healthcare costs, namely around 7% [7], and the treatment costs are rising [5]. The incidence of cancer is very high in the Netherlands as there is around a 40% chance to get the diagnosis of cancer during a lifetime and the prognosis is that this will increase the coming years [8]. The plan is to focus cancer care on health, functioning and quality of life and on establishing care together with patient and professional [6, 9]. Additionally, the national healthcare agreement [4] defines to disseminate proven innovations quickly and effectively and to scale down care that is not (or no longer) appropriate [4]. In terms of innovation in cancer care, there is significant potential for AI-driven Clinical Decision Support Systems (AI-CDSS) in cancer care as AI-CDSS can provide support to medical specialists in their treatment decision-making process, potentially leading to improved treatment choices. For instance, by personalizing treatment decisions, AI-CDSS could help prevent overtreatment, ultimately contributing to a higher quality of life. These potential advantages of AI-CDSS are in line with the current trajectory 'adequate cancer care' [6]. Preventing overtreatment could potentially simultaneously improve quality as well as make care more efficient thereby responding to impending financial and personnel shortages [1]. Despite extensive development efforts and substantial potential, AI-CDSS are scarcely used in practice. This raises the question of why this is the case.

1.2. Problem Statement

AI-CDSS can support medical specialists in complex decision-making processes [10]. The use of AI techniques to assist medical specialists in making rapid and precise diagnostic decisions is becoming increasingly popular [11]. An AI-CDSS takes patient data as input and produces personalized advice for each patient [12]. AI-CDSS may improve patient outcomes by providing medical specialists with real-time information and recommendations about patient care. Certainly in today's cancer care context, the AI-CDSS could be incredibly useful. Despite their potential and the amount of research being done on it (see section 1.3), AI-CDSS are still little used in cancer treatment decision-making in the Netherlands. It is unclear what the reasons are for this. There is a research gap in the factors that affect adoption and implementation of AI-CDSS in cancer care. Mohammadpour *et al.* [13] and Rahimi *et al.* [14] also indicate this gap. This leads to a main research question that will fill the gap of missing knowledge on cancer specific AI-CDSS:

Why does the adoption and implementation of AI-CDSS in cancer care fail?

As found in literate and described in chapter 3, a multi-actor perspective is often missing in literature, while the literature does show that there are many actors involved in the adoption of AI-CDSS. It is therefore expected that this multi-actor perspective is important and should be further investigated. This will be investigated by a case study while using an actor and institutions lens.

1.3. AI-CDSS Research Gap

A literature search is conducted to identify the existing research in the field of AI-CDSS in cancer care and gain an understanding of the types of studies that have been conducted to date. The search is conducted on PubMed and Scopus and is desribed in appendix A.

It is found that AI-CDSS are being developed for a variety of different cancer types. The largest proportion of applications are developed for breast cancer, which currently experiences the most benefit from AI-CDSS [15]. In terms of use case in cancer care, there are lots of different ones. AI-CDSS can predict the outcome of chemotherapy in both neoadjuvant [16–24] and adjuvant [25–27] settings, allowing medical specialists to determine the best course of treatment for individual patients [28]. It can also predict the response to immunotherapy and identify patients who are most likely to benefit from such treatment [29–31]. AI-CDSS can predict patient survival [15, 27, 32–38] and identify patients who are at high risk of mortality [39]. It can also predict the likelihood of toxicity [40] and side-effects associated with chemotherapy [41–43], helping medical specialists to adjust the dose or come up with a different treatment plan to avoid adverse events. AI-CDSS can aid in dose planning for radiation [44, 45], chemotherapy [46, 47], and other therapies [48], providing personalized treatment plans for individual patients. It can also be used to diagnose and screen for cancer [13, 24, 41, 49–55]. AI-CDSS can also play a crucial role in drug development [56], predicting the efficacy and safety of new drugs and potential drug-drug interactions [57, 58] and identifying potential drug targets by identifying target genes [59, 60] or by matching earlier found therapy options to other cancer types [56]. AI-CDSS can evaluate drug response, determining whether a full response has been achieved, and if additional surgery is needed or not [49, 51, 53, 61–66]. As for AI technology, all kinds of different types of AI are

used, from simple machine learning techniques, such as support vector machines [67] , to complex ones, such as deep learning [16, 25, 34, 45, 64, 68, 69] and reinforcement learning [70–73].

It was observed that almost all papers identified in this literature search focus on the development and performance evaluation of AI-CDSS. There are no papers focusing specifically on the adoption and implementation of AI-CDSS in cancer care. When applying the following criteria to the found papers: (1) the article examined the adoption and implementation; (2) It is mentioned that the CDSS uses AI methods; (3) The article was about supporting cancer treatment decisions. It became clear that no paper met all criteria.

1.4. AI-CDSS in Breast Cancer

The case study will focus on the Mammaprint test. However, because Oncotype is a similar test and is assessed in the same way by the Healthcare Institute Netherlands (ZiN) the Oncotype test will also be discussed.

Mammaprint

The Mammaprint is a test that can be used to predict the risk of metastases from breast cancer after the tumor has been surgically removed. The Mammaprint uses the expression profiles of 70 genes to determine whether the risk of metastasis is high or low. At high risk additional chemotherapy is needed and at low risk, it is probably unnecessary. To make the classification of high and low-risk groups, unsupervised and supervised classification algorithms were used on training data. These algorithms use a form of machine learning and are thus AI. The moment one can predict that chemotherapy is not necessary because the risk of metastases is low, toxic side effects could be avoided. The Mammaprint test is developed at the Netherlands Cancer Institute (NKI) which is affiliated with the Anthonie van Leeuwenhoek (AVL) hospital. Agendia was found in 2003 [74] as a spin-off of the NKI-AVL to further develop the Mammaprint test as a product. The test is based on groundbreaking research (11578 citations) from the NKI-AVL that was published in 2002 [75] and further validation thereof [76]. In 2004 the 70-gene test emerged as a product. Mammaprint is gaining much attention worldwide in its early stages. It receives several awards between 2005 and 2008 and is named one of the greatest healthcare innovations by Time magazine and Oprah Winfrey, among others [74]. Mammaprint also receives the health innovation award from the Dutch Minister of Health in 2008 [74]. In 2007 Mammaprint was the first diagnostic test cleared by FDA [74]. In 2009 Mammaprint is included in the St. Gallen guidelines, which are the international consensus guidelines and describe the ‘state of the art’ for early-stage breast cancer [77]. Agendia is ISO certified since 2012 and Mammaprint also received EU In Vitro Device CE mark in Europe [78–80].

Oncotype

Oncotype is a similar test to Mammaprint. Oncotype is based on a 21-gene expression profile. The test was developed by Genomic Health and the patents have been held by Exact Sciences since its acquisition and merger in 2019.

1.5. Societal Relevance

As outlined in section 1.1, the Dutch healthcare system's unsustainability, particularly in the realm of cancer care, is currently a significant challenge which is high on the political agenda [4, 6, 9]. AI-CDSS are promising in the context of 'adequate cancer care' [6] as they could potentially be part of the solution, by preventing overtreatment, but are currently rarely used in practice. Given that approximately 40% of the Dutch population receives a cancer diagnosis at some point in their lives [8], it is needles to say that ensuring accessible, affordable, and high-quality cancer care can greatly benefit many individuals. Mapping how decision-making processes influence the adoption and implementation of AI-CDSS in cancer care has relevance primarily in two areas. First, it can serve as valuable information for policy makers, enabling them to utilize the identified factors in shaping policies conducive to a more favorable environment for the adoption and implementation of AI-CDSS. Second, it might catalyze development by providing developers with a comprehensive overview, allowing them to anticipate and address challenges in decision-making processes proactively.

1.6. Scientific Relevance

The scientific relevance is realized by adding insights of a comprehensive case study of adoption of an AI-CDSS in cancer care using a multi-actor perspective. This case study has unique value as a broad range of actors is interviewed who are in the core of the network of the adoption and implementation process of the AI-CDSS. Most of the interviewees have key roles in this process for both Mammaprint as well as other AI-CDSS. At the same time, it gives an insight into applying an institutional actor framework for the implementation of AI-CDSS in cancer care, which to the best of my knowledge is unique. This framework proves useful and identifies the existence of institutional voids around the reimbursement of AI-CDSS. In addition, the research extends beyond the immediate focus on this particular AI-CDSS and sheds light on whether similar challenges arise in other AI-CDSS implementations. This contextualization provides valuable insights into categorizing various AI-CDSS and the diverse pathways they involve. Consequently, this study identifies where to look further to advance AI-CDSS implementation, which also provides actionable directives for future research.

1.7. Link to the EPA program

This research fits well within the requirements of the EPA master program. The research takes a systems perspectives and focuses on the healthcare system in which AI-CDSS could be implemented. Within the healthcare system the decision-making arena for reimbursement in basic health insurance is considered. The research uses an institutional actor analysis (IAA) framework to analyse the problem situation. The problem is part of a bigger context within healthcare which aims at the grand challenge of having a sustainable healthcare system available for the public. It also touches the interface of the public and private domain, as AI-CDSS development and implementation is dependent on the collaboration between public and private parties.

2

Methodology

This chapter outlines the employed methodology for conducting the research.

2.1. Research Approach

The research flow is illustrated in figure 2.1 and further explained below.

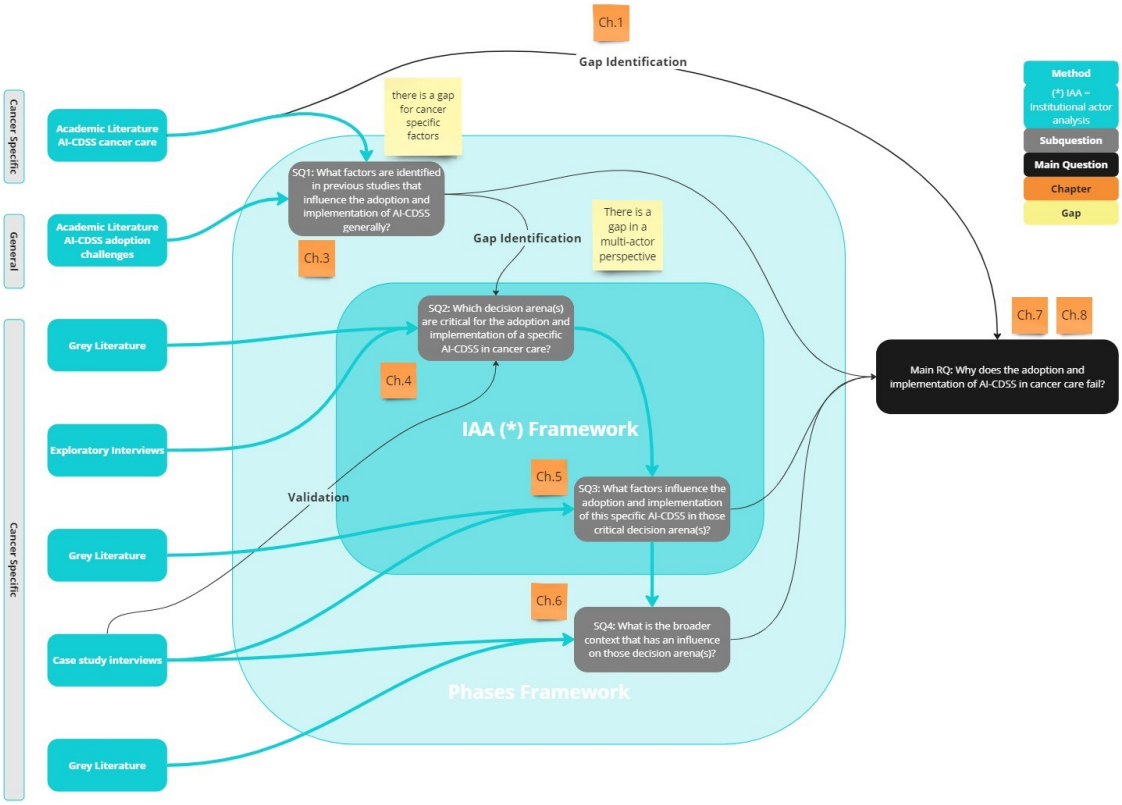

Figure 2.1: Methodology

An analysis of academic literature concerning AI-CDSS in cancer care (described in section 1.3) revealed a notable gap in understanding the factors impacting their adoption and implementation. Existing literature primarily addresses the development and accuracy testing of AI-CDSS in cancer care, with limited focus on adoption and implementation challenges. Therfore the following main research question will be answered:

Why does the adoption and implementation of AI-CDSS in cancer care fail?

The first subquestion delves into a broader examination of current literature, encompassing healthcare as a whole rather than focusing solely on cancer-specific aspects. This approach offers an initial glimpse into potential adoption and implementation challenges. The first subquestion is answered in chapter 3 and reads:

1. What factors are identified in previous studies that influence the adoption and implementation of AI-CDSS generally?

Through answering this first sub-question, it was found that there is a lack of a comprehensive multi-actor perspective in the literature regarding factors affecting adoption and implementation. This while various stakeholders, from practitioners to policy makers, are jointly shaping the technology integration journey. It is expected that no single actor can unilaterally ensure adoption and implementation. Recognizing these actors, understanding their respective roles, and acknowledging their motivations in decision-making are central aspects in understanding the pathways that facilitate AI-CDSS adoption and therefore a multi-actor perspective is deemed essential [81]. Therefore the adoption and implementation is perceived as a system where institutions and actors interact, collectively shaping the decision-making process. This is done via a case study approach while making use of data from grey literature and interviews. The case study employs an Institutional Actor Analysis (IAA) (see section 2.2.2) to acquire and structure the data. In this way the viewpoint from a multi-actor perspective can be identified, which is missing in existing literature.

To execute this analysis, identification of the critical decision-making arena(s) pivotal to the adoption and implementation of AI-CDSS in cancer care is required. Hence, the second subquestion reads:

2. Which decision arena(s) are critical for the adoption and implementation of a specific AI-CDSS in cancer care?

This subquestion is answered by grey literature and exploratory interviews and validated during the case study interviews. The answer on this subquestion is discussed in chapter 4.

Subsequently subquestion 3 aims to closely examine the critical arena(s) and identify the factors within them that contribute to or hinder the adoption of AI-CDSS. This subquestion focuses on internal dynamics and positions within the decision-making process, seeking explanations for the outcomes observed within the arena(s). The subquestion reads:

3. What factors influence the adoption and implementation of this specific AI-CDSS in those critical decision arena(s)?

This subquestion is answered in chapter 5 through grey literature and case study interviews.

However, each decision arena operates within its own contextual framework. While internal factors within the arena are significant, external factors beyond the arena's boundaries can also be influential on processes inside. By exploring the broader contextual factors, a more comprehensive understanding is gained. Subquestion 4 reads:

4. What is the broader context that has an influence on those decision arena(s)?

This is answered in chapter 6 mainly through case study interviews and supplemented in part by information obtained from grey literature.

Insights from all subquestions are then interwoven in the discussion (see chapter 7), culminating in a comprehensive answer to the primary research question as discussed in chapter 8.

2.2. Frameworks

During the research, two frameworks were employed to conduct a thorough analysis. These frameworks served as guidelines to structure and interpret the collected data. These frameworks are the 'Phases' framework and the 'Institutional Actor Analysis (IAA)' framework. The frameworks are further explained below.

2.2.1. Phases Framework

A sequential framework of phases of adoption and implementation is employed to categorize and structure the identified factors that influence the adoption and implementation of AI-CDSS. The phases are: development, regulation, sell/marketing, purchase/reimbursement and use phase and their accompanying descriptions are shown in figure 2.2. This framework is based on the life cycle of medical technology as described by the Dutch ministry of health [82], which is further explained in appendix B. Throughout the thesis, these phases served as a framework to organize and categorize the findings from both the literature review and interviews. Doing so provides a clear structure for analyzing and synthesizing existing research, as well as identifying knowledge gaps and areas where more research is needed.

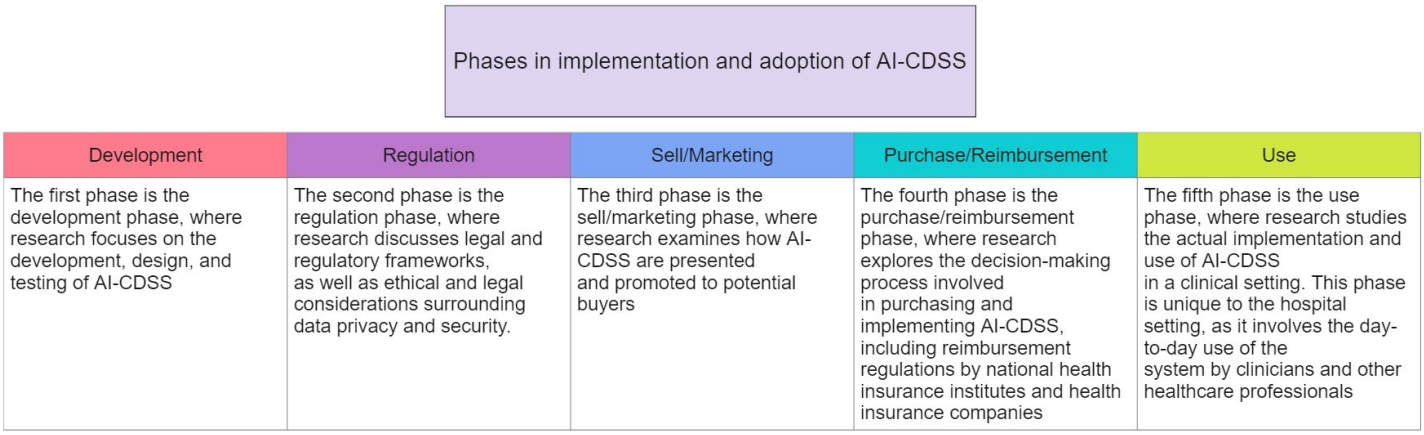

Figure 2.2: Phases in adoption and implementation of AI-CDSS

2.2.2. Institutional Actor Analysis (IAA) Framework

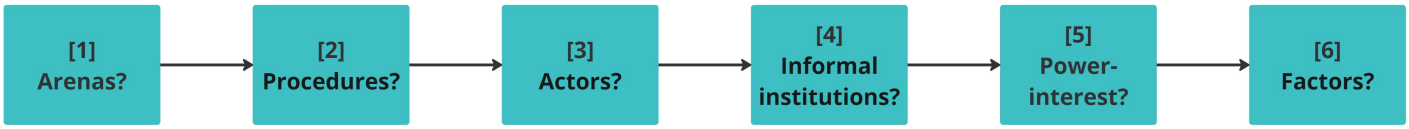

Figure 2.3: Institutional Actor Analysis

As discussed in section 3.6 a multi-actor perspective of adoption and implementation of AI-CDSS is missing in the literature, while there are many different actors involved in the adoption and implementation of AI-CDSS. Therefore in this research this perspective will be taken as it is deemed essential to have this outlook to find out why the adoption and implementation of AI-CDSS is lacking behind. Throughout this research a particular view and scientific jargon will be used. Key jargon, along with definitions and illustrative examples, will be provided in this section to ensure a clear understanding of the terminology used. Next, the steps within an institutional actor analysis (IAA) will be introduced, which is based on the introduced jargon. The IAA framework is a modified iteration of the actor analysis framework outlined in Enserink *et al.* [81], and it draws upon an extensive body of research on actors and institutions, from, among others Bryson [83], Eden & Ackermann [84], Hermans & Cunningham [85], North [86], Ostrom [87] and Teisman [88].

Decision arenas and decision-making rounds

Decision arenas are designated social spaces dedicated to strategic decision-making [85]. These can take on formal boundaries or exist virtually, where actors interact [87]. Networks and institutions within these arenas shape actor behavior significantly [85]. In the context of this research, the decision-making arena encompasses the social space where actors engage in discussions regarding the inclusion of Mammaprint in basic health insurance. Within these decision arenas, there are decision-making rounds. According to Teisman [88], a decision-round is marked by the initiation and conclusion of a specific period, during which significant decisions are made. A decision round, in the case of the case study, encompasses the initiation of the assessment process for Mammaprint's inclusion, extending up to the ultimate publication of an official position report, where a definitive decision is described regarding inclusion or exclusion. A decision-making round thus encompasses the entire process from the initial determination to address a specific question to the ultimate position about inclusion or exclusion of Mammaprint. The decision-making process regarding Mammaprint has unfolded in three distinct decision rounds. These include the initial assessment in 2010, a subsequent reassessment in 2018, and an ongoing assessment since 2023 [1].

Action situations

Within a decision-making round, there are distinct action situations [87]. An action situation refers to a particular set of circumstances or context where individuals or groups engage in decision-making or take actions. It signifies a focused moment or scenario within the broader decision-making process. In the context of a decision-making round concerning the assessment of Mammaprint, various action situations arise. For instance, the initiation of an assessment or reassessment constitutes an action situation. Similarly, when the draft position report is published, providing stakeholders with the opportunity to provide input, marks another significant action situation.

[1] Just before the publication of this report, the decision of the 2023 decision round was made. Thus, the assessment is now no longer ongoing, but has been concluded.

Formal and informal institutions

In decision-making arenas, institutions play a crucial role in shaping the decision-making processes [81, 86, 87]. An institution, in this context, encompasses established rules, norms, and practices that regulate behavior within a society, organization, or community [81, 86, 87]. Essentially, institutions serve as humanly devised constraints that govern political, economic, and social interactions. They comprise both informal constraints, such as sanctions, taboos, customs, traditions, and codes of conduct, as well as formal ones - the codified rules found in law, regulation, or custom, and those embedded in organizations or government structures [87]. Informal rules are not formally documented and can be challenging to comprehend and may not be uniformly interpreted by those involved [81].

In the context of the case study, namely inclusion in basic health insurance, there exist legally defined criteria, specifically the SWP criterion [89], which functions as a formal institution. According to this criterion, health insurers are prohibited from reimbursing care that does not meet SWP standards; otherwise, they are required to cover the cost from their own budget. Additionally, in cases where health insurers are unable to reach a consensus on whether certain intramural care aligns with SWP, the healthcare institute Netherlands (Zorginstituut Nederland, abbreviated as ZiN), must conduct an assessment.

Furthermore, informal institutions govern the arena, such as determining responsibility for reassessments. Despite lacking legal backing, there exists an unwritten rule that health insurers should refrain from conducting reassessments for care forms already evaluated by the ZiN. This unwritten norm guides the assignment of responsibility. In the case of Mammaprint, it initially came to the attention of the ZiN due to a dispute, establishing the ZiN's ongoing involvement in reassessments. Consequently, there is currently no opportunity for health insurers to conduct the assessment of Mammaprint, unless this social norm undergoes a change.

Actors and Multi-actor situation

An actor refers to an individual, group, organization, or entity that has the capacity to make decisions and take actions within a social system. Within an arena there are thus actors interacting [81, 87]. There are multiple actors involved in the decision-making processes, which is called a multi-actor situation. A multi-actor situation encompasses various actors, organized within a network rather than a traditional hierarchy. According to Enserink *et al.* [81]. It implies that no single actor can impose their preferred solution on others independently; instead, some level of cooperation among parties is essential. An institutional actor refers to an entity or organization, often with a formalized structure and recognized authority, that participates in decision-making processes within a particular institutional framework. These actors can include governmental bodies, regulatory agencies, non-governmental organizations, and other entities with a defined role in shaping policies and practices. Thus, an institutional actor is involved through the institutions in place in the arena.

In the arena concerning the inclusion of Mammaprint in basic health insurance, several institutional actors play crucial roles. One such actor is the ZiN, responsible for evaluating whether Mammaprint aligns with the SWP criterion. Additionally, medical professional groups qualify as institutional actors as they receive the assessment report and hold the opportunity to provide feedback. Therefore, individuals who do not engage in the arena through formal or informal institutional channels are not classified as institutional actors. For instance, individual medical professionals, lacking representation through insti-

tutions, do not fall under the category of institutional actors. However, if a medical professional speaks on behalf of an organization participating in decision-making through established channels, they would be deemed an institutional actor.

Power and Interest

Institutional actors have a certain power and interest within the decision-making arena. By considering both power and interest, stakeholders can be categorized into different groups, namely key players, context-setters, subjects and crowd [83, 84]. A Key player is characterized by having high levels of both power and interest in the decision-making process. They hold significant sway over the outcomes and are deeply invested in the matter at hand. A Context setter wields considerable power but exhibits low levels of personal interest in the decision. They may not be personally invested, but their influence plays a crucial role in shaping the overall context of the decision-making arena. A Subject has substantial personal interest in the decision but lacks the corresponding level of power. While they are deeply invested, their ability to directly influence outcomes may be limited. In contrast, a Crowd has neither significant power nor a substantial personal interest in the decision. They hold minimal influence over the process and may not be particularly invested in the outcomes. By mapping out these roles of actors it can help in understanding the stakeholder landscape and can help when investigating strategies to improve the decision-making processes. For example to make the decision-making processes more effective and inclusive.

Institutional Actor Analysis

The IAA framework is an adjusted version of the actor analysis framework as discussed in Enserink *et al.* [81]. The institutional approach involves examining the current policy-making structures and pinpointing actors who hold formal positions in the policy-making process. The framework consists of 6 steps, namely (1) arena identification, (2) exploring procedures, (3) identifying actors, (4) revealing informal institutions, (5) mapping the power-interest of actors and (6) subsequently synthesising these findings to identify factors. The steps of the IAA framework are visualized in figure 2.3 and the steps are mentioned below.

1. **Identification of Critical Decision-Making Arenas**
2. **Mapping Procedures/ Formal Institutions**
3. **Identify Actors**
4. **Reveal Informal Institutions, Roles and Interactions**
5. **Assess Power and Interest**
6. **Synthesise Factors**
   Here the findings of former steps are synthesised to come to a conclusion of how decision-making processes within the arena influence outcomes.

2.3. Literature Reviews

Search Protocol

This study integrated two literature reviews. The first review is about AI-CDSS in cancer care. This review is situated in the introduction (see section 1.3) and revealed a gap in the existing literature pertaining to the factors influencing the adoption and implementation of AI-CDSS in cancer care. The search criteria are described in appendix A. Titles and abstracts of the studies identified through this search were reviewed and studies were assessed for adherence to predefined inclusion criteria:

- (1) the article examined the adoption and implementation;
- (2) It is mentioned that the CDSS uses AI methods;
- (3) The article was about supporting cancer treatment decisions

The second literature review is about AI-CDSS adoption challenges. This literature search aims to acquire an initial understanding of the factors impacting the adoption and implementation of AI-CDSS in healthcare, which thus encompasses a broader context beyond cancer care. The search criteria are described in chapter 3. Titles and abstracts of the studies identified through the search were reviewed. Studies were assessed for adherence to predefined inclusion criteria:

- (1) the article examined the adoption and implementation;
- (2) It is mentioned that the CDSS uses AI methods;
- (3) The article was about supporting treatment decisions

Analysis

For the first literature review the cancer types the papers focused on and the primary use case of the AI-CDSS were identified. For the second literature review, the papers were read to identify factors that affected the adoption and implementation of those systems. This was done as per the following steps:

- Identify factors mentioned per paper
- Group factors based on the implementation phase (see figure 2.2 for the phases overview)
- Create themes per phase

If the paper included surveys and/or interviews:

- Identify actors mentioned per paper
- Group actors based on the implementation phase (see figure 2.2 for the phases overview)

2.4. Exploratory Interviews

Three exploratory interviews were done, during which valuable insights from a range of stakeholders concerning the adoption and implementation of AI-CDSS was gathered. These interviews were conducted with people who knew more about the subject and were within reach. There was input from nurses specialized in pancreatic cancer, a former medical specialist, and a developer of CDSS. These interviews were not recorded because they were of an exploratory nature. Notes were taken during the interviews for personal use. The main findings of the exploratory interviews are used to get an initial understanding of the decision arenas (step 1 of IAA) that could be critical for the adoption and implementation of Mammaprint. The outcomes are described in section 4.1.

2.5. Case Study

The case study focuses on Mammaprint and involves the data analysis of interviews and grey literature. This approach differs from the conventional method of conducting a case study as defined by Yin [90], which also includes observations for triangulation. In this study, interviews and grey literature are deemed sufficient to address the research questions, and observations were not within the scope of the study. Furthermore, conducting observations would not be possible given the historical nature of the case study and given that the current decision round is taking place in confidentiality.

2.5.1. Grey Literature

The grey literature examined in this study includes: publicly available position statements from the ZiN, documents from the Dutch house of representatives and other results from Google searches. Grey literature is used in the first five steps of the IAA framework. It is used to get an initial understanding of the decision arenas (step 1) that could be critical for the adoption and implementation of Mammaprint. Furthermore, the draft position reports of the ZiN [91–93] give insights in the subsequent steps. The procedures/formal institutions (step 2 of IAA) and the therefrom actors (step 3 of IAA) could be mapped. Also, an initial understanding of the informal institutions (step 4 of IAA) and the power-interest could be identified from these position reports. This was possible as in the appendix of the position reports, the written responses from the actors on the draft position report are included [92, 93].

2.5.2. Case Study Interviews

The case study interviews provided a validation and final conclusion for subquestion 2 (step 1 of the IAA framework). Subsequently it validated procedures (step 2 of IAA), actors (step 3 of IAA) and additional insights concerning those steps. It gave an elaborate perspective of informal institutions (step 4 of IAA) and power-interest (step 5 of IAA).

Sampling of interviewees

The selection of interviewees followed two distinct sampling methods. The initial method involved an institutional actor analysis to identify key stakeholders. Subsequently relevant stakeholders were contacted. The second method employed snowball sampling, where interviewees' recommendations led to the inclusion of additional interviewees. The process of engaging with interviewees was occasionally challenging in the beginning. Some individuals or organizations responded delayed or minimally, and

some did not respond at all. As the interview phase progressed, a network of connections was established, easing the outreach process. The snowballing approach yielded notable advantages. Firstly, it facilitated referrals to suitable interviewees, with certain names recurring as suggestions. Secondly, referred participants showed a greater willingness to contribute due to the familiarity of the referral source.

Interviewees

Nineteen participants were involved in the research, see table 2.1. Eighteen interview sessions were conducted over a 1.5-month period, including one interview where two interviewees were present in one session. All interviews were between 0,5-1,5 hour. The interviewees represented diverse backgrounds and affiliations, spanning various roles and organizational associations. Because of privacy reasons, most organizations are not coupled to the interviewees. The multifaceted composition of interviewees ensured discussions encompassing a broad spectrum of perspectives. In table 2.1 it is also shown for which arena and/or adoption phase the interviewees are relevant. It can be seen that interviewees involved in the development phase are primarily involved researchers and interviewees with ties to the commercial aspect. In the regulation phase, the interviewees are characterized as briding the intersection of research and the commercial facet of AI-CDSS. The Sell/Marketing phase engaged interviewees associated with both the commercial and health insurance domains. Finally, the use phase encompassed insights from medical specialists, who provided their perspectives on the use of AI-CDSS. It is pertinent to note that all interviewees categorized within the guideline or reimbursement subgroup were or are directly involved in this process. If an interviewee mentioned Mammaprint's reimbursement in basic health insurance but was not directly involved, they were not included in this subgroup. In section C.2 in the appendix and table C.1 additional characterization about the interviewees involved in the reimbursement and/or medical guidelines arena of Mammaprint are shown. While interactions covered discussions with most key players and context-setters, some key players have not been interviewed, for example the patients association. However, the absence of direct engagement was compensated by documented insights within the position reports of the ZiN [92, 93] as the written and verbal reactions of all stakeholders, including the patient association, on the draft position reports are included in the appendix of those reports and as insights were given by other interviewees.

Informed Consent

Before the start of each interview, participants were asked to read and sign a consent form. This was sent by email before the interview took place. This form explained the objectives of the study and also explicitly addressed potential privacy concerns. The consent form further clarified the purpose of data collection, outlined how the information collected would be used, and provided transparency about the intended methods of publication. Upon understanding and agreement, all participants signed the consent form, underscoring their voluntary and informed participation in the study. However, six interviewees made additional requests regarding their consent. Two interviewees requested the opportunity to review the report before its publication. Additionally, two interviewees expressed the desire to verify their quotes prior to publishing, while two others preferred not to have their quotes included at all. These additional requirements have been met.

Interview Questions

The topics discussed during the interviews were around three main topics, namely the medical guidelines, the reimbursement by basic health insurance and AI-CDSS. In appendix C.1 a more detailed list of topics including some of the questions asked are shown.

Given the significant diversity among interviewees and the dual roles they often held, a separate preparation was required for each interview. This had the advantage that each interview was personalized and provided a different perspective on the issue. Additionally, it allowed for further exploration of interesting findings that arose during the interview.

Before each interview, a thorough background check of the interviewee was conducted through LinkedIn and the employer's website. This helped gather relevant information about the interviewees' involvement in various aspects, ranging from medical guidelines, reimbursement by basic health insurance, Mammaprint-related matters, and broader about adoption and implementation of AI-CDSS within their field of expertise.

At the start of each interview the purpose of the research and the reason why the interviewee was contacted including their functions and potential roles, in the case of Mammaprint, was repeated. Subsequently it was asked if the interviewee could introduce him/herself and their personal involvement in the arena and/or experience with Mammaprint or other AI-CDSS. If new interesting roles of the interviewees came up, the topics to be discussed could then still be customised based on this. This was necessary in about a third of the interviews because asking this revealed that the extent and nature of their involvement in Mammaprint and/or other AI-CDSS was greater than could initially be found (online) or was heard from the referral.

The interview method employed was mainly semi-structured and for some topics unstructured. The questions pertaining to the reimbursement and guideline arena were primarily semi-structured. This meant that there were predetermined questions about procedures and involved parties, although the sequence could be adjusted and asking follow-up questions on interesting topics was possible. On the other hand, the unstructured component frequently came into play when discussing the broader context of AI-CDSS adoption and implementation. This was for example particularly evident in the case of interviewee 11, the cancer screening expert. This interview was unstructured and the general area of interest and concern was the adoption and implementation of AI-CDSS in screening, allowing the conversation to naturally evolve within this area.

Recorded

The interviews were held through video or phone calls and were recorded for further analysis. The automatic transcription feature of Microsoft Teams software was utilized to transcribe the recordings, although some errors were observed in the transcriptions. The recordings of the interviews were listened back and only relevant quotes were accurately transcribed manually. This as it was not needed and too time-consuming to manually correct all transcripts.

	Function Description	Reimbursement	Guideline	Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
1	Breast Cancer Surgeon-Oncologist		X				X	X
2	Breast Cancer Surgeon-Oncologist		X					X
3	Surgeon in Training	X						X
4	Breast Cancer Medical Oncologist	X						X
5	Clinical Pathologist and Diagnostician			X			X	X
6	Breast Cancer Pathologist and Researcher			X			X	X
7	Researcher of AI in Cancer Care			X	X			
8	Representative of Ministry of Health	X					X	
9	Cancer Care Spin-off Expert			X	X	X	X	
10	Cancer HTA Expert	X					X	
11	Cancer Screening Expert			X	X		X	X
12	Representative of the registration holder of Mammaprint	X		X	X	X	X	
13	Study Managing Director Breast Cancer Research			X				
14	Process Facilitator Breast Cancer Guidelines		X					
15	Medical Guideline Revision Specialist							
16	Medical Oncology Guideline Specialist		X					
17	Advisor for Inclusion in Basic Health Insurance	X					X	
18	Advisor at a Health Insurance Company	X				X	X	
19	Advisor at a Health Insurance Company	X					X	

Table 2.1: Overview of Interviewees and the arenas (Reimbursement and Guideline) and phases (Development, Regulation, Sell/Marketing, Purchase/Reimbursement and Use) they are involved in.

2.5.3. Case Study Analysis

The data analysis of the grey literature and the interviews followed an iterative process. As there was a significant reliance on who could be interviewed and the depth, extent and focus of their knowledge, it was challenging to establish a definitive analytical approach beforehand. The extensive diversity of interviewees and their respective specializations further complicated matters. However, a suitable approach was devised, which will be elaborated upon in this section.

Following each interview, new patterns and insights emerged from the data. Noteworthy findings were captured, and after each interview, a reassessment was conducted to identify information that was relevant to one or more of the subquestions. This involved active listening, documenting intriguing findings, comparing them with findings from other interviews, and potentially grouping quotes and formulating themes and codes within the themes. Each interview was listened to multiple times for analyses. In table 2.2 the main themes and the corresponding subquestions for which these themes are relevant are shown. In the appendix in section C.3 and in table C.2 further explanation is given about the themes and the codes identified and the relevant statements made by interviewees are included. The case study analyses can broadly be categorized into sub-questions 2, 3, and 4, which will be further explained below.

Theme	Relevant for subquestion
Guidelines	2
Reimbursement	2, 3, 4
No clear requirements for inclusion of diagnostic tests	3
(High) Level of burden of proof needed	3
Disagreement on study design between medical specialists and ZiN	3
Priorities in trade-off between quality of life and survival differ	3
Development	4
Regulation	4
Sell/Marketing	4
Purchase/Reimbursement	4
Use	4

Table 2.2: Overview of main themes created during interview analysis

Analysis for subquestion 2: Critical Arena for Mammaprint

Subquestion 2 answers the question which arenas are critical for the adoption and implementation of Mammaprint. This question matches step 1 of the IAA framework (see section 2.2.2). To answer this question, first grey literature is checked which provided a shortlist of critical arenas. Then grey literature and the interviews allowed criticality to be validated. The interviewee groups relevant for validating this were 'Reimbursement' and 'Guidelines,' (see table 2.1). The included grey literature and the relevant insights from the interviews are discussed in chapter 4 in section 4.2.

Analysis for subquestion 3: Factors influencing the Reimbursement arena of Mammaprint

Subquestion 3 answers the question which factors influence the adoption and implementation of Mammaprint within the critical arena as identified in subquestion 2. In order to answer this question steps 2-6 of the IAA framework (see section 2.2.2) are completed using grey literature and interviews. Regarding the grey literature, the responses from stakeholders on the the draft position reports of the ZiN [92, 93] which are included in the final position reports of the ZiN [92, 93] provided a wealth of information for answering this subquestion.

The interviewee group mainly relevant for answering this subquestion was the 'Reimbursement' group (see table 2.1). However more interviewees, who were not directly involved within the arena, but involved in another way, also provided relevant insights. The relevant findings of the interviews are stated in table C.3 in appendix C under the themes corresponding to subquestion 3 (see table 2.2).

The interviews were relistened, focusing on whether the formal procedures, as outlined in the grey literature, were indeed put into practice for Mammaprint. Additionally, institutional actors identified in the grey literature were corroborated. During the interviews, non-institutional actors who nonetheless influenced the arena or exerted influence also emerged (see section 5.3). The interviews and stakeholders reactions on the draft position reports [92, 93] were analyzed to understand the reasons behind Mammaprint's lack of reimbursement and the opinions surrounding this issue. This combination of grey literature and interview analysis resulted in four identified factors (see section 5.4). In table C.3 in appendix C the findings of the interviews concerning those four factors are stated. It should be mentioned, however, that the factors found in the parties' responses to the draft position reports had a large share in defining those four factors, but that those responses are not in table C.3 as they did not come from the interviews. However, quotes from parties' responses to the draft position reports can be found in section 5.4 and in the position reports themselves, which are available online [91–94].

Analysis for subquestion 4: Contextual Factors

Subquestion 4 answers the question which arenas are critical for the adoption and implementation of Mammaprint. This question matches what the broader context is that has an influence on the Reimbursement arena of Mammaprint. To answer this question, interview input was used. This exploration encompassed not only Mammaprint but also other AI-CDSS categorized as diagnostic tests and various other types of AI-CDSS. When examining the broader context, all phases (see the phases framework in section 2.2.1) were considered, rather than solely focusing on the reimbursement/purchase phase. The interviewee groups relevant for validating this were thus 'Development', 'Regulation', 'Sell/Marketing', 'Purchase/Reimbursement' and 'Use' (see table 2.1). Additionally some grey literature was checked to validate and/or elaborate on the factors that came up, especially for the 'Purchase/Reimbursement' phase.

3

Factors Influencing Adoption and Implementation of AI-CDSS

The literature review conducted in section 1.3 to identify existing research on AI-CDSS in cancer care revealed a significant lack of information regarding adoption and implementation factors. As a result, this literature review has a broader scope, extending beyond cancer, but with a sharper focus on the adoption and implementation factors. Subquestion 1 will be answered in this chapter: 'What factors are identified in previous studies that influence the adoption and implementation of AI-CDSS generally?'

For the literature review different search terms were tried. There are some variants for clinical decision support systems (CDSS), such as clinical decision aid or clinical decision support program. However, when using those terms it is noteworthy that not all the tools mentioned in the retrieved papers are computerized, and a substantial number of them are not employed for analytical purposes, but rather for databases or platforms. In case 'clinical' is not used in front of the other words, the searches are not healthcare specific, but come up with applications in many other industries. For AI a broad search is used that includes many different key terms for AI to be as complete as possible. So the search includes "simple" machine learning as well as deep learning, reinforcement learning, and text mining. To retrieve papers on adoption or implementation, the search term 'adoption OR implementation' was tested. However, it became evident that the term 'implementation' was not appropriate, as it frequently refers to the implementation of algorithms rather than implementation in the clinical care setting. Therefore, only the term 'adoption' was used as a search term. In order to specifically focus on the factors that affect the adoption and implementation, the following search terms were added: challenge, barrier, impediment, enabler, or perception. The final search term combinations for Scopus and Pubmed are shown below.

The search was conducted using the following search terms on Scopus:
(TITLE-ABS-KEY ((clinical AND decision AND support AND system) OR cdss)
AND TITLE-ABS-KEY ((artificial AND intelligence) OR 'ai' OR (machine AND learning) OR (deep

AND learning) OR (neural AND networks) OR (reinforcement AND learning) OR (neural AND language AND processing) OR 'nlp' OR (text AND mining))
AND TITLE-ABS-KEY (adoption)
AND TITLE-ABS-KEY (challenge* OR barrier* OR impediment* OR enabler* OR perception*))

The search was conducted using the following search terms on Pubmed:
((clinical decision support system) OR cdss)
AND ((artificial intelligence) OR 'AI' OR (machine learning) OR (deep learning) OR (neural networks) OR (reinforcement learning) OR (neural language processing) OR 'nlp' OR (text mining))
AND (adoption)
AND (Challenge OR barrier OR impediment OR enabler OR perception)

The search was conducted on April 2023 with no restriction on the publication date. The search was limited to English-language articles that were available in full text. A total of 129 and 123 articles were retrieved from the search on Scopus and Pubmed respectively. To avoid duplication of results, any duplicate articles were removed, which resulted in 204 unique papers. To identify relevant articles, the titles and abstracts of each article were read. Full-text articles were then reviewed for inclusion based on the following criteria: (1) the article examined the adoption and implementation; (2) It is mentioned that the CDSS uses AI methods; (3) The article was about supporting treatment decisions. A total of 37 articles met the inclusion criteria and are used in the review. Of those 37 articles, there were 9 studies (mainly) based on interviews [95–103], 6 (mainly) based on surveys [104–109] and the 22 other papers are review papers. Some additional papers are included in this literature review through snowball sampling.

3.1. Development phase

Clinician Involvement

It has been found that involving clinicians in the development of AI-CDSS can lead to more successful implementations [97, 101, 110]. This is because clinicians play a key role in identifying clinical needs and minimizing disruption to the workflow [97], which in turn increases confidence in the systems [101]. Additionally, clinicians are often curious about who developed the systems and the involvement of clinicians in the development can promote trust in the systems [95, 96, 111]. They are also curious if the data used to train and validate the system is of high quality [101, 112]. Ideally, clinicians should be involved throughout the entire development process [97, 113]. However, it is common for clinical experts to be involved only at the beginning and end of the process [113]. In a review of 80 studies on AI-CDSS development, less than a third of the studies described clinical experts being involved in the process [113]. If an expert is involved, it is often to verify the clinical correctness or relevance of the model. A study by Kashyap *et al.* [110] examined the organizational configuration of implementing AI-CDSS at 20 health sites in the U.S. and found that in 90% of the cases, clinicians and researchers from the health site were involved in developing and implementing the systems. Only in 10% of the cases was the implemented system developed by an external company. This highlights the importance of involving

clinicians in the development process and/or giving them an important role in the implementation of the system. However, it is also a risk that the CDSS will only be used by clinicians who were involved in their development [114].

Availability of Funding

The study of Watson *et al.* [97] among leaders of academic medical centers (AMCs) in the U.S. found that smaller institutions in particular had problems with funding. In some cases, this caused an ongoing development process of a model to stop temporarily or completely. In the review of [115] it is mentioned that clinical trials are cost-intensive and only a small number of AI applications can be financially supported. The costs are caused by personnel costs, technical supplies and clinical expansion of models [97]. Only large institutions with many models in development had no funding problems [97].

Availability of Data

During the development phase, the availability of data is crucial, but can often be limiting [97, 102, 116]. due to the imperfections in data sets, machine learning models in medicine cannot achieve perfection [116]. There are several data issues that can contribute to this, including insufficient, incomplete, low-quality, unreliable, and non-diverse data [97, 100, 101, 107, 115, 117]. These imperfections arise from various factors, such as noise, errors in documentation, incompleteness, and differences in documentation granularities [116]. Also the lack of adherence to standard data format guidelines in practice [100, 101, 107] plays a role. For example, obtaining high-quality cancer data sets that have undergone omics profiling is challenging in the clinical setting due to cost limitations, limited sample availability, and concerns about data quality [117]. Also, gaps in data standards in ophthalmology hinder the harmonisation of big data repositories [118].

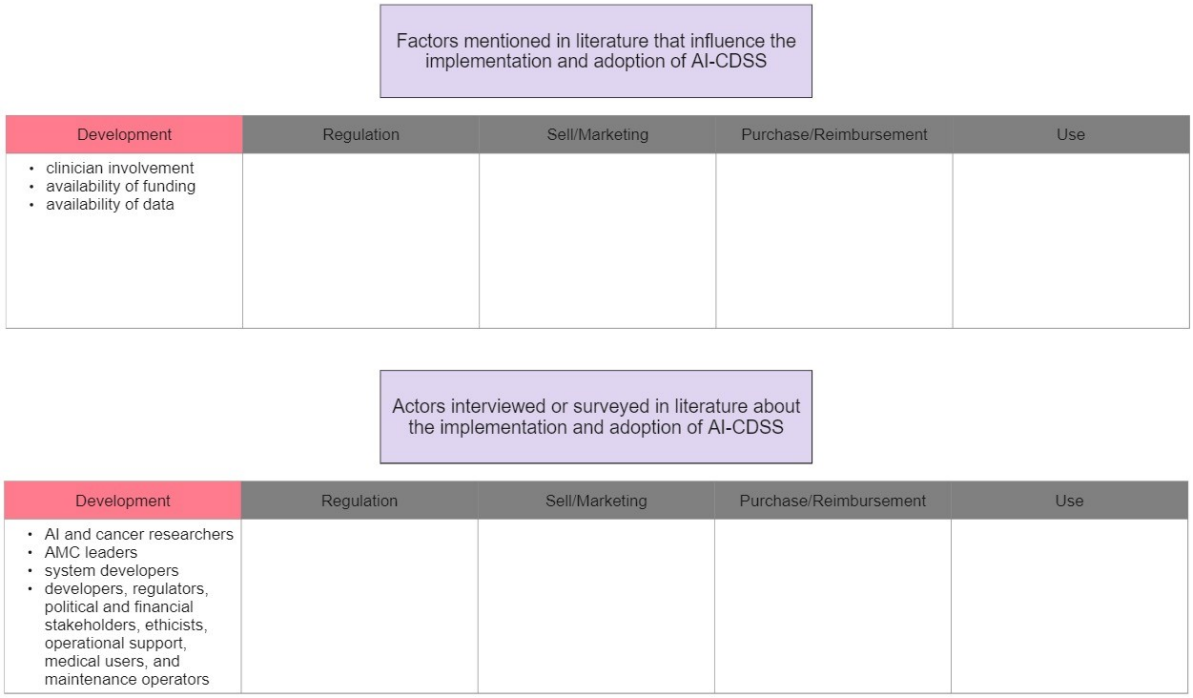

Figure 3.1: Factors and actors found in the literature that influence the adoption and implementation of AI-CDSS in the development phase

3.2. Regulation phase

CE Approval

In particular, there are challenges with obtaining CE approval for these systems. The fact that only 5% [102] of sampled systems in one study had CE approval indicates that this is not an easy task. As AI-CDSS is software designed to be used in diagnosing, treating, preventing, or monitoring diseases, it must be certified as an active medical device in the EU [101, 102]. Firstly, there are no well-established norms and laws on how to conduct a proper risk assessment for predictive algorithms like AI-CDSS [100, 101], which can make it difficult to obtain CE approval. This is also indicated in the paper of Harwich & Laycock [119], where interviewees emphasized the necessity for clarity regarding the contents of the certification, specifically outlining what it would encompass, such as testing protocols, discrimination tests, data pre-processing protocols, and training data. Secondly, gathering data becomes more complex due to the CE quality standards [101]. As an illustration, devices that generate data need to comply with CE quality standards for medical equipment. Even though they are often approved for research, they are not authorized for medical use [101]. Thirdly, for clinical evaluation, it is common to do a randomized controlled trial (RCT) [102, 115]. However, RCTs are often hardly possible for AI-CDSS because of practical and ethical concerns [101].

Data Sharing Standards

There is a lack of standards for data, model validation methods, and legal and safety regulations, as reported in [115, 120]. When there are no standards in place, achieving organizational interoperability becomes challenging [115] and data sharing between institutions becomes difficult, hindering development [102, 115]. The absence of semantic coding is identified as a root cause in long-term data integration or migration strategies [121]. Without clear regulations on model validation, it is impossible to comply in advance, which can make the validation process unnecessarily long [115, 120].

Legal factors

Legal factors are crucial in the regulation phase. Firstly because of transparency and explainability reasons. In case of any complaints, hospitals must be able to justify every step of the treatment process [101]. The European Commission and the Food and Drug Administration (FDA) have emphasized the importance of transparency when using AI in healthcare [97]. However, using black box algorithms can make this very complicated [122]. Moreover, legislation in this area still allows for multiple interpretations [123], making it even more challenging to comply with regulations. Additionally, there is a debate on what exactly explainability entails. While accuracy of a system may be sufficient for daily practice, explainability is important in order to improve the accuracy of systems and thus enhance care for future patients, as noted in a paper [124]. Secondly, the responsibility for treatment decisions made using an AI-CDSS is a topic of debate [125]. Is it solely in the hands of the clinician, or should the algorithm developer also be partially responsible [100, 101, 107]? The AI-CDSS are decision-support tools and do augment the doctor's cognitive skills, but do not substitute them [119]. However, the AI-CDSS can subsconsiously influence the decision of a doctor [126]. In case the AI-CDSS is wrong, this could be problematic. The lack of legal cases involving medical AI makes it challenging to navigate issues related to medical liability [120]. Therefore, there is a need to re-evaluate the roles and responsibilities of

both developers and medical professionals [105, 114, 119, 127, 128]. Furthermore, obtaining informed consent for AI algorithms can be difficult [107, 123, 129]. Patients can withdraw their data through the General Data Protection Regulation (GDPR), which poses a risk that the entire model would have to be re-trained and validated [123].

Factors mentioned in literature that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• clinician involvement • availability of funding • availability of data	• CE approval • data sharing standards • legal Factors			

Actors interviewed or surveyed in literature about the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• AI and cancer researchers • AMC leaders • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators			

Figure 3.2: Factors and actors found in the literature that influence the adoption and implementation of AI-CDSS in the regulation phase

3.3. Sell/Marketing phase

Acceptance of AI in Healthcare

The healthcare industry generates a large amount of data at a much faster pace and with greater variety than other fields such as education and finance [130, 131]. This data has the potential to reveal new information that could enhance work practices and yield superior outcomes [130, 131]. However, the benefits of modern analytical and data science techniques in healthcare are not fully realized [132–134]. One key reason for this is the slow adoption of artificial intelligence methods in the healthcare sector [135, 136]. A market study conducted by the McKinsey Global Institute in 2019 found that the healthcare industry has the lowest adoption rate of AI technologies [137] among all industries.

Respondence on Market Demand

There are indications that the supply and demand of AI-CDSS may not be aligned. For example, despite extensive research on applying text mining to medication errors and cancer detection and diagnosis, there is a lack of emphasis on utilizing this technology in AI-CDSS [111], while the demand for these application areas is probably high as medication errors and cancer are leading causes of death in many countries [138]. The paper of Cabral *et al.* [107] also found that early cancer detection is one of the most preferred applications of AI in cancer care according to cancer and AI researchers. This certainly opens up the opportunity to address the research of AI-CDSS on cancer screening. In addition to the gap in target applications, there is also a gap in the target group. The paper of O'Sullivan *et al.* [114] highlights that while the majority of CDSS developments are focused on clinicians, the growing engagement of patients with online medical content is creating a new market for consumer-oriented CDSSs. This could be partly attributed to the fact that most CDSS are developed in an academic environment [102]. However, marketing knowledge is often missing in such a context [102] which could hamper the marketing of those systems.

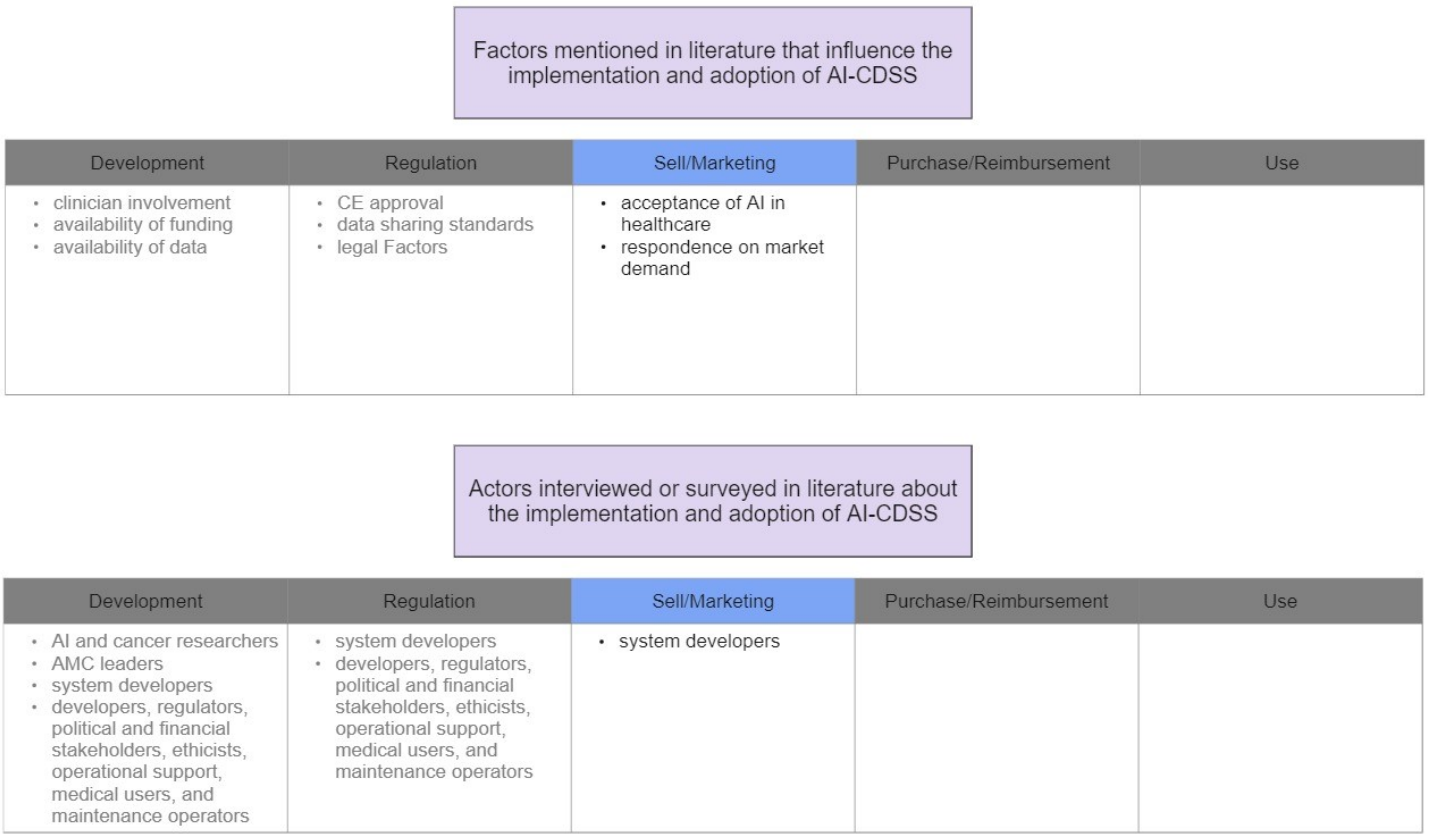

Figure 3.3: Factors and actors found in the literature that influence the adoption and implementation of AI-CDSS in the sell/marketing phase

3.4. Purchase/Reimbursement phase

Cost-effectiveness

Before making any commitment for the purchase of an AI-CDSS it is in the interest of the institute to assess the cost-effectiveness [101, 104, 129]. Realizing AI-CDSS in a healthcare institute namely requires commitment and budget [100, 101, 104]. Especially in the area of IT, more resources are needed, as data exchange protocols within the organization must be established and sufficient technical knowledge and computer systems are required [100, 101]. A systematic literature review of von Wedel & Hagist [139], which examined the economic impact of data and analytics for healthcare providers, found that CDSS and AI generally had a positive economic impact. This positive effect was achieved through direct cost effects or through efficiency gains. However, it should be noted that only 13 studies of CDSS and AI were examined in this literature review, so this could certainly turn out differently and also depends on the care context.

Installation and maintenance support

The setting in which AI-CDSS are developed poses challenges in providing complete service. The development of AI-CDSS often takes place in academic projects, which are funded for a relatively short amount of time [102, 114]. As a result, it can be challenging to offer a comprehensive service

that includes installation and maintenance support [114], which is crucial for selling it as a product. Maintenance personnel are scarce, which makes it even more significant and difficult [97]. Furthermore, deploying any CDSS on a larger scale requires tailoring the system to the local clinical setting, including the established clinical workflow, site-specific clinical vocabulary, and locally installed hardware and software IT systems. It is also essential to maintain the system in the face of quickly evolving clinical knowledge and the lack of standardized institutional guidelines on periodic reviews of CDSS [114]. In the paper of Roberts *et al.* [140] where they discuss the adoption of Oncotype, a breast cancer AI-CDSS, in the U.S. it is also mentioned that training for ordering and usage of the test, offering guidance and assistance in navigating coverage and reimbursement, and supplying educational materials and up-to-date research findings regarding its efficacy facilitate the use of the Oncotype test [140]. Furthermore, it was noted that patients belonging to groups not covered for reimbursement of the test by health insurance faced obstacles in accessing Oncotype [140].

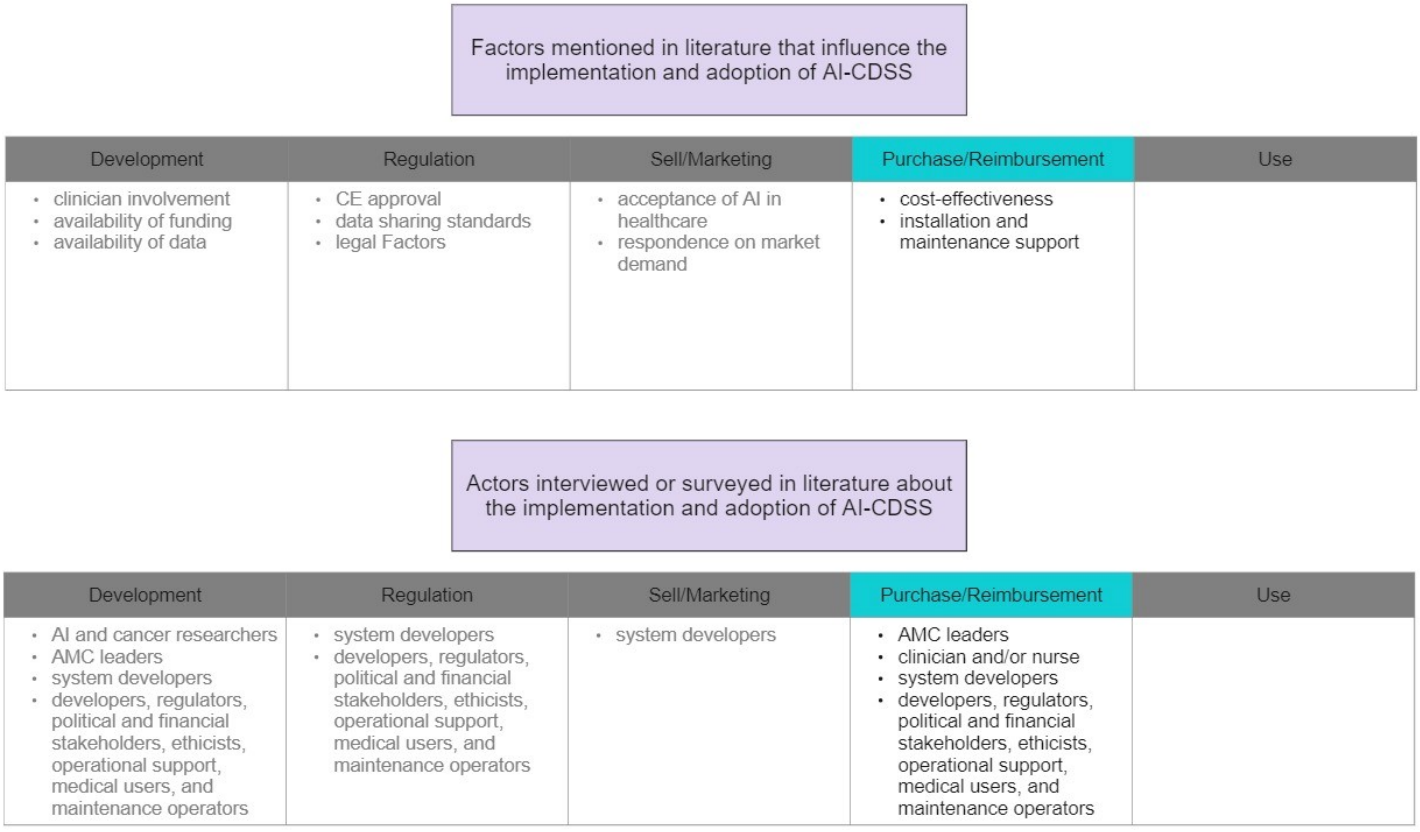

Figure 3.4: Factors and actors found in the literature that influence the adoption and implementation of AI-CDSS in the purchase phase

3.5. Use phase

Clinical validity

What comes up a lot is that clinicians need to have confidence in the systems before using them [96, 98, 112]. Clinicians find the clinical validity of the AI-CDSS very important [96, 98, 106, 111, 112]. Macro evidence such as validation studies is very important to gain this confidence [96, 98, 111]. In terms of performance, it is important that the accuracy is high and that the system is consistent [106, 107, 112]. In addition, micro evidence such as stories of experience from other clinicians can play a major role; when they are enthusiastic and share a success story, this increases confidence in the system [96, 98].

Clinician's profile

Some of the factors influencing adoption are not dependent on the AI-CDSS system itself but on the clinician's profile. In general, there is a lack of knowledge of AI among clinicians [104, 106]. Both in the use itself and also in the potential of AI [105, 106]. However, clinicians' profiles and their acceptance of ML models differ a lot. Medical expertise, technological expertise, academic background, personality, and cognitive bias [108] play a role in this [112]. In the paper of Gunasekeran *et al.* [105] they could predict with moderate to high accuracy the acceptance of ML models by clinicians with ML models itself using clinician's demographics. What feels paradoxical is that the research of Gunasekeran *et al.* [105], which is a global survey within the field of ophthalmology, found that clinicians with more medical expertise (more than 20 years) are more in favor of using AI-driven systems. However, the paper of Alsobhi *et al.* [104] indicated that seniors and juniors in rehabilitation do not significantly differ in their attitude towards AI. Although seniors often have more experience with using AI [104]. It was notable that non-academic workers in rehabilitation had greater knowledge of AI than academic ones [104]. Additionally, risk perception plays a big role in clinicians' adoption of the AI-CDSS. The studies [108, 109] investigated the perceptions of clinicians while using (or not using) an AI-CDSS to predict the appropriate number of packed red blood cells to transfuse for a specific patient. It is found that a clinician's perception of risk [108] and the trust [109] in the AI-CDSS play a mediating role in their use of an AI-CDSS. The study found that the more a clinician perceived the risk of using the AI-CDSS to be high, the less they tended to use it. Conversely, clinicians who had higher expectations for the AI-CDSS or a more positive perception of AI to positively impact care outcomes tended to perceive lower risk in using the system [108]. The disadvantage of this study is that it was only examined for 1 application of an AI-CDSS, so it is not generalizable [108].

Clinician's sense of added value

Perceived accuracy by clinicians was mainly influenced by the concordance between a clinician's impression and the AI-CDSS prediction [96]. This is also consistent with the cultural norm that physicians feel an authority in their field of experience and also base choices on their gut and intuition, which in turn is based on experience [96, 105, 107]. Clinicians feel they add value and an AI-CDSS could deprive clinicians of their sense of added value [101]. AI-CDSS can only make a decision based on the data input they receive, while important input information may then be missing, which a clinician can consider [96, 105, 106]. For example, a clinician can include a patient's facial expression, while an AI-CDSS cannot [99]. Similarly, the paper by Matthiesen *et al.* [98] with pre-implementation research

with blinded retrospective patient cases in a large hospital in Denmark mentions that clinicians did not change their decisions after seeing the output of an AI-CDSS, but most did experience it as helpful. This is also consistent with the fact that clinicians like the idea of using AI-CDSS as assistive tools in decision-making, but not as the decisive decision-maker [96, 105]. In the paper of Alsobhi *et al.* [104] where they examine the use of AI-CDSS in rehabilitation, only 4 out of 236 (1.6%) respondents had a negative opinion about using AI as an assistive technology tool in clinical practices. This picture is also confirmed in a global survey of Gunasekeran *et al.* [105] on the use of AI among ophthalmologists, where it emerges that the overall attitude of clinicians toward AI is very positive. However, from the clinicians' perspective, these systems will not replace the clinician but will support the clinician [96, 99, 105]. In case clinicians perceive information technology (IT) as a threat to their professional autonomy it has a significant, negative direct influence on their intention to use that IT [141]. This is also mentioned in focus groups with physicians in the study of [142], where it is found that medical practitioners are accustomed to a culture of autonomy, which could lead to resistance in adopting AI-CDSS within the clinical workflow as integrating such a system may disrupt their established autonomy.

Explainability

In particular, the use of artificial intelligence raises questions [97, 122, 143–145]. First, because the algorithms are often ”black box” algorithms [97, 99, 105, 122, 144, 145]. It is therefore difficult for a clinician to understand both the process leading to the outcome and the outcome itself [123], while clinicians value explainability and transparency [99, 101, 107, 112]. Understanding the algorithm creates trust among clinicians [107], which in turn supports the adoption [143]. Understandability could be influenced by training and education [95, 99, 120, 143]. After education on the operation of the AI-CDSS, few clinicians (less than 10%) were significantly hesitant to use it [95]. Second, because the absence of information on the reasons behind an AI-based decision hinders physicians from assessing the advice's reliability, posing a risk to patient safety [128]. Here the model interpretability plays a role, which is the capability of understanding how the prediction was influenced by specific aspects of the data or input features [125, 146–148].Third, because clinicians are unfamiliar with the metrics to evaluate performance of AI-CDSS [97].

Equitability

However, what is a concern of clinicians is that these AI-CDSS are not suitable for all patients, because of two main reasons. First, for very complex disease states, for example, where multiple conditions are involved at the same time or in case it is a very rare condition [96, 100, 101, 103]. Second, the systems may have a bias for certain patient groups, which may prevent other patient groups from receiving the best care [107, 120, 129]. For example Zhou *et al.* [149] suggest that the utilization of U.S.-based IBM Watson in Oncology using U.S. medical literature has resulted in misguided treatment recommendations when applied in the Chinese context. Ethically, it is important to ensure that these algorithms make care more equitable [120, 128]. A global survey among cancer and AI researchers believed that bias in the algorithms is one of the biggest factors to hamper the use of AI in cancer care [107].

Interoperability

What is a problem, however, is that many systems have different performance scores the moment they are used in another institution [98, 111, 150]. This also calls for adjustable systems that can be adapted to the current context so that they can be tuned for that context and or fit the local workflow [98, 112, 114]. That interoperability of AI-CDSS between different institutes and within an institute between different systems is important is also described in a post-implementation study of rural clinics in China [99]. According to clinicians working there, this interoperability, especially from within hospital systems, can greatly enhance use. This is also mentioned in the study of Wijnhoven [101] and Wijnhoven [100], where the use of an AI-CDSS is tested in an AMC in the Netherlands. Here the stakeholders say that data protocols are not used, so data from different systems have to be integrated manually. Additionally, the case study of AI-CDSS implementation in the emergency department of an academic healthcare centre in Canada also indicated the operability of systems within a hospital as limiting factor of the adoption in practice of AI-CDSS [103]. Additionally in case the AI-CDSS advices on treatment options which are not available in the local context or too expensive, especially in middle to low income countries [151], there will be low concordance between the advice of the system and the decision of the physician [99, 149]. The paper of van de Burgt *et al.* [111] notes that when using text mining, language is also an impediment. The moment it is developed for the English language, it is not directly applicable in another country where the data is in a different language.

Ease of use

The ease of use of AI-CDSS is crucial for their successful implementation and adoption in healthcare settings. This is especially important considering the integration with electronic health records (EHR) and the need for user-friendliness [106, 151]. Research conducted by van de Burgt *et al.* [111] emphasized the significance of familiarity with the interface. Studies revealed that using an unfamiliar interface [130] and difficult outputs [130, 152] led to participants not using the systems or making mistakes, which can ultimately hinder the adoption of CDSS. This is also mentioned in the study of Petitgand *et al.* [103], where physicians refrained from utilizing medical histories due to the challenge of comprehending patient information provided by the CDSS. However in the study of Roberts *et al.* [140] where they investigate the use of Oncotype, it becomes clear that clinicians find the interpretation of an 'intermediate risk' score difficult, but that this does not lead to less usage. It does however lead to clinicians inventing their own rules, resulting in differences of interpretation by different oncologists [140]. It is essential for AI-CDSS to have intuitive interfaces and outputs [103] that are familiar to healthcare professionals, enabling them to easily navigate and utilize the system. The outputs hold significance, and it is highly appreciated if the AI-CDSS can also provide indications of the uncertainty in its output [117]. Another aspect of ease of use is the ability of AI-CDSS to save time. Clinicians are especially positive about using AI if it causes them to do fewer monotonous tasks themselves [106]. They can then help more patients and/or further improve the quality of care [106]. The moment they can help more patients, it also increases accessibility for patients, which clinicians see as very positive [106]. These systems should not disrupt the existing clinical processes but rather augment them [120], ensuring that valuable time is saved and not wasted on complex or convoluted interfaces.

Data security and privacy

The integration of AI technology in healthcare has raised concerns regarding the security and privacy of patient data. There is a fear that healthcare may become controlled by large technology and data companies [105]. The implementation of AI in healthcare is further challenged by data breaches [128] and ransomware attacks, highlighting the need for appropriate measures to safeguard patient privacy [153]. While de-identifying data is a vital step in protecting patient privacy, the possibility of re-identification remains a significant concern [128, 154]. Researchers are studying privacy-preserving techniques and encryption processes to enhance security and data protection [129].

Factors mentioned in literature that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• clinician involvement • availability of funding • availability of data	• CE approval • data sharing standards • legal Factors	• acceptance of AI in healthcare • respondence on market demand	• cost-effectiveness • installation and maintenance support	• clinical validity • clinician's profile • clinician's sense of added value • explainability • equitability • interoperability • ease of use • data security and privacy
15 references	18 references	13 references	9 references	21 references

Actors, identified in literature, involved in implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• AI and cancer researchers • AMC leaders • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers	• AMC leaders • clinician and/or nurse • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• clinician and/or nurse • patients and care partners • AMC leaders • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators
5 interview/survey references	3 interview/survey references	1 interview/survey references	5 interview/survey references	16 interview/survey references

Figure 3.5: Factors and actors found in the literature that influence the adoption and implementation of AI-CDSS

3.6. Synthesis of Findings

The process of adopting AI-CDSS is influenced by a multitude of factors and actors. Figure 3.5 provides an illustrative overview of the factors identified in the literature review across different phases and the number of references that pertain to each phase. Also the types of actors who have been surveyed or interviewed regarding these factors are shown.

The literature focuses mainly on the use phase

From this overview it becomes clear that there is a notable emphasis on the use phase of AI-CDSS adoption, while the foregoing phases receive comparatively less attention. This trend is apparent from the number of identified factors and the number of literature references. The use phase garners more research interest, likely due to its direct impact on clinicians and patient care. In contrast, the sell/marketing and purchase/reimbursement phases are relatively underrepresented, both in terms of available sources and the number of identified factors. This suggests that the factors influencing these phases might not be as extensively studied. Further investigation into these phases could provide insights into effective strategies for adoption and implementation of AI-CDSS.

The literature lacks a multi-actor perspective

Table 3.1 shows the types of actors that are interviewed and or surveyed to understand their perspectives, needs, and challenges in adopting AI-CDSS in the found literature. From this three things can

Function Description	**References**
clinician and/or nurse	[95, 96, 98, 99, 103–105, 108, 109, 140]
system developers	[102]
AMC leaders	[97]
clinicians, patients, care partners	[106]
AI & cancer researchers	[107]
system developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, maintenance operators	[100, 101]

Table 3.1: An overview of the types of actors identified in the literature and the corresponding references

be deducted. First, studies have mainly focused on clinicians as the primary actors in the adoption of AI-CDSS. Second, a diverse array of actors have been interviewed and/or surveyed in the existing literature, highlighting the extensive involvement of various stakeholders in the adoption and implementation of AI-CDSS. Third, a limited number of papers have explored the perspectives of multiple stakeholders, indicating a lack of a comprehensive multi-actor perspective. Only the studies of Wijnhoven [100, 101] show a broad spectrum of interviewed actors, including developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators. These studies delve into the organizational learning process of AI-CDSS adoption, analyzing the reasons behind the limited adoption of AI-CDSS [100, 101]. This research emphasizes the importance of understanding AI-CDSS adoption as a complex organizational learning process, rather than a linear implementation. A case study approach is employed to analyze the adoption journey. It's remarkable that these papers are the only ones that offer a multi-actor perspective. However, this multi-actor perspective is critical for comprehensively grasping AI-CDSS adoption, capturing the intricate interplay among different actors involved in the implementation process. Therefore, this research will conceptualize the adoption and implementation of AI-CDSS as taking place within a system characterized by decision-making arenas and

the interaction between actors. The achievement of adoption and implementation cannot be attributed to a single actor alone as successful integration requires collaboration, coordination, and alignment among diverse actors [81] throughout different phases. An institutional actor analysis will therefore be used as framework to get a structured understanding of the challenges faced in implementing AI-CDSS.

4

Identification of Critical Decision Arena

In addressing the first subquestion: 'What factors are identified in previous studies that influence the adoption and implementation of AI-CDSS generally?', three things stood out. Firstly, there is relatively much research on the use phase and relatively limited attention directed towards the preceding stages. Secondly, a diverse range of actors is involved in the adoption and implementation of AI-CDSS. Thirdly, a multi-actor perspective on the adoption and implementation is lacking. This discrepancy is noteworthy considering the numerous factors that could potentially impede successful adoption, encompassing diverse dimensions and involving multiple stakeholders. Therefore this research will approach the problem from a multi-actor perspective. In order to be able to investigate the adoption and implementation of an AI-CDSS in cancer care, a case study will be done using interviews and grey literature for data collection. Mammaprint, a breast cancer AI-CDSS, will serve as the focal point for this case study. The selection of Mammaprint is well-founded; its Dutch origins facilitate access to pertinent interviewees, and its historical challenges with adoption and implementation make it a compelling case. Despite being operational for nearly two decades, Mammaprint still lacks reimbursement under basic health insurance in the Netherlands [a]. Additionally, the recent inclusion of a similar AI-CDSS in breast cancer, Oncotype, in basic health insurance enables meaningful comparative analysis. These factors collectively position Mammaprint as an ideal subject for in-depth case study exploration. This chapter tackles the second subquestion: 'Which decision arena(s) are critical for the adoption and implementation of a specific AI-CDSS in cancer care?' This is the first step of the IAA framework as shown in figure 4.1

[a] Shortly before the publication of this report, the decision of the 2023 decision round on reimbursement of Mammaprint in basic health insurance was made and had a positive outcome. Mammaprint is thus now included in basic health insurance for a specific group of breast cancer patients.

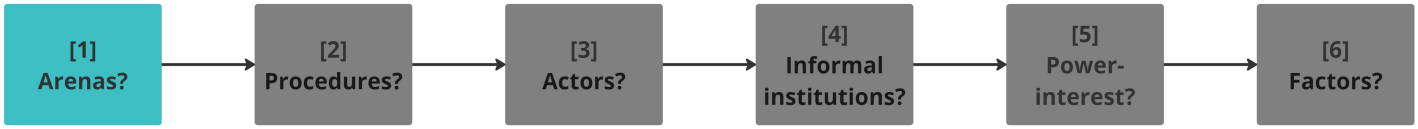

Figure 4.1: Institutional Actor Analysis

4.1. Shortlisting Critical Arenas

A shortlist of critical arenas are drawn up based on three exploratory interviews and grey literature.

Insights from Exploratory Interviews

During the exploratory interviews, insights were gathered from various stakeholders regarding the adoption and implementation of AI-CDSS. The perspectives of specialized nurses in pancreatic cancer, a former medical specialist, and a CDSS developer in cancer care were obtained. The interviewed pancreatic cancer nurses working at a Dutch academic hospital gave their opinion about the use of AI-CDSS in treatment decision-making. They were enthusiastic, however, they repeated to get into contact with the Dutch Pancreatic Cancer Group, the national opinion forming body for pancreatic cancer, as when they are on board and in favor of the use, the others will follow. Second, the medical specialist indicated that actors deciding upon adoption and implementation of AI-CDSS are mainly found on a national level and not on the level of hospitals. It was mentioned that inclusion in guidelines plays a very important role in this. Third, the interview with the CDSS researcher highlighted that the algorithms for guiding cancer chemotherapy decisions are currently in the development phase. In contrast, Mammaprint has already progressed beyond this stage and therefore the interview did not cover the pertinent decision-making contexts for Mammaprint.

These exploratory interviews thus highlighted several key insights. Firstly, it appears crucial to address adoption and implementation challenges at the national level due to the fact that decision-making processes and opinion-forming bodies operate at this level. Additionally, the adoption by medical specialists is likely contingent upon inclusion in the national medical guidelines for breast cancer.

Insights from Grey Literature

From the grey literature the following interesting findings appeared. Agendia, the company of Mammaprint, stated for achieving a grant of the European Union in 2015 that both uptake in the medical guidelines and reimbursement arrangements are important for full-scale adoption.

In order to reach this full-scale adoption, this project is focused to achieve the last clinical evidence for the clinical utility of this test needed for the uptake in EU and US clinical guidelines and reimbursement arrangements. This uptake is necessary for adoption of MammaPrint by physicians worldwide. [155]

This aligns harmoniously with earlier findings. During the selection of Mammaprint as case study, it was already evident that the endeavor to secure reimbursement under basic health insurance remained unresolved [1] [93, 94]. Moreover, the exploratory interviews underscored the pivotal role of inclusion in

[1] Shortly before the publication of this report, the decision of the 2023 decision round on reimbursement of Mammaprint in basic health insurance was made and had a positive outcome. Mammaprint is thus now included in basic health insurance for a specific group of breast cancer patients.

medical guidelines.

Critical Arenas

From the exploratory interviews and the grey literature the following two arenas are thus identified: the Guideline arena and the Reimbursement arena. In the guideline arena there is decided upon whether and in what manner Mammaprint will be integrated into the Dutch medical guidelines for breast cancer. In the reimbursement arena there is decided whether Mammaprint will be included in basic health insurance in the Netherlands and, if so, which patient groups will be covered.

From these findings, a generalization can be drawn regarding the key factors influencing the adoption and implementation of AI-CDSS. The list of factors can thus be extended to include inclusion in medical guidelines, which applies to the use phase. Additionally, reimbursement by basic health insurance (in the Netherlands) can also be considered a significant factor. The latter factor directly impacts the purchase/reimbursement phase. The expansion of factors is illustrated in figure 4.2.

Factors mentioned in literature that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• clinician involvement • availability of funding • availability of data	• CE approval • data sharing standards • legal Factors	• acceptance of AI in healthcare • respondence on market demand	• cost-effectiveness • installation and maintenance support	• clinical validity • clinician's profile • clinician's sense of added value • explainability • equitability • interoperability • ease of use • data security and privacy
			• **reimbursement by basic health insurance**	• inclusion in medical guidelines

Actors interviewed or surveyed in literature about the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• AI and cancer researchers • AMC leaders • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers	• AMC leaders • clinician and/or nurse • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• clinician and/or nurse • patients and care partners • AMC leaders • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators

Figure 4.2: Factors and actors that influence the adoption and implementation of AI-CDSS

4.2. Selection of Critical Arena Narrows down to Reimbursement

The arenas of reimbursement by basic health insurance and inclusion in clinical guidelines serve as pillars for the case study. Grey literature is checked and interviews are done with interviewees occupying various roles in both arenas (see table 2.1 and table C.1 for further characterization). These include process facilitators and individuals who were engaged in prior or current decision-making rounds. Lots of interviewees (3, 6, 7, 9, 12 and 14) underscore the importance of guideline inclusion for the acceptance of AI-CDSS, including Mammaprint, in clinical practice.

Exploration of the Guideline Arena

The guidelines themselves are publicly accessible [156–158], but meeting notes or other related documents are not available. It became evident that medical specialists hold the primary authority over the guidelines. They collaborate within the guideline committee of the Dutch Breast Cancer Society (NABON) [159], which encompasses a diverse spectrum of medical specialist associations involved in breast cancer care. The introduction of gene expression profiles in the guidelines occurred in 2008 [156], followed by updates in 2012 [157] and 2020 [158]. In 2008, gene expression profiles were recommended only for study settings. In 2012, the guidelines advised considering gene expression profiles for hormone-sensitive invasive ductal carcinoma cases where doubts arose regarding adjuvant chemotherapy after employing traditional prognostic factors. The 2020 update highlighted potential added value of the test and further specified the group for which it could be applied. For further details on appearance in the guidelines, see table D.1 in appendix D. Furthermore, it becomes clear from interviewees 14 and 15 that there is usually consideration in terms of reimbursement or non-reimbursement and inclusion in guidelines as medical specialists deemed its utilization necessary in specific cases.

Interviewees 1 and 14 were involved during all earlier decision rounds of the guideline arena and mentioned that they did not recall encountering significant challenges about the uptake of Mammaprint. Interviewee 14 also indicates that it is long ago, but as far as the interviewee can remember the medical specialist all agreed that Mammaprint should be included in the guidelines based on the outcomes of the MINDACT study [160]:

What I can still remember, because this has been going on for quite a long time of course, it seemed to be a very promising tool at first, but it was decided to wait for the results of a very large randomized study, the MINDACT, before we decided to do anything definite. And they (the results) were just reviewed by the guideline working group and they were also reviewed positively. And so I don't think there was any discussion within the guideline working group itself whether or not that was the right intervention. No, I think everyone was in agreement on that. And also when that hassle with the ZiN then came, the NABON published an additional point of view as being we as multidisciplinary working groups Netherlands also think this is very important (to include Mammaprint in basic health insurance).

It thus appears that previous decision-making rounds have been devoid of significant debate. However, it can not be ascertained that there was no significant debate because there is no documentation of meeting minutes available.

Currently, the breast cancer guidelines are under review. Interviewees 2 and 15 are involved in the current guideline arena and indicated that there is likely no alteration to the section concerning gene

expression profiles, under which Mammaprint falls as there is no disagreement at this point on how Mammaprint is included in the guidelines.

The Guideline Arena is Not Critical

Initially, this arena seemed intriguing to investigate as a critical arena. Hence, procedures of medical guideline drafting were investigated and interviews were initiated. However, it became evident that the breast cancer guideline arena is currently not a critical arena for the adoption and implementation of Mammaprint. Even while Mammaprint is not reimbursed by basic health insurance, it is in the guidelines. The guideline arena is thus not causing any hindrance in the adoption and implementation of Mammaprint. Furthermore, the investigation into previous decision-making rounds did not yield any significant insights. There is no meeting documentation available and interviewees' statements indicate the absence of significant challenges within their recollection. Therefore, it is not deemed worthwhile to further investigate this arena as the acquired knowledge did not yield any compelling answers to sub-question 3, namely what factors within this arena influence the adoption and implementation. Some additional insights on this arena are, however, provided in appendix D.

Exploration of the Reimbursement Arena

Regarding the reimbursement in basic health insurance arena, there exists a reasonably comprehensive documentation of the decision-making processes [91–94]. These final position reports also include, in the appendix, responses from key players to the draft position report. These responses already give a lot of insight into the disagreements. Furthermore, insights from the case study interviewees shed light on disagreements both in past and the current decision-making rounds (see section 5.4 and table C.2 in the appendix). Notably, Mammaprint has yet to be included in basic health insurance in the Netherlands [94] [2].

Multiple decision-making rounds have taken place, with the first occurring in 2010 [91] and the second in 2018 [92]. In both instances, Mammaprint was denied inclusion in the basic insurance package. Presently, another round of decision-making is underway [94], and it became clear from interviewees 12 and 19 that there is once again disagreement about inclusion in the basic package [2].

Moreover, various interviewees (2, 3, 6, 9), mainly medical specialists, emphasized the significance of inclusion in the basic insurance package for adoption and implementation. However, it becomes evident from interviewees 2, 3, 8, 10, 12 and 17 that Mammaprint has been extensively utilized in practice for a prolonged period, even without basic insurance coverage. Over the years, various alternative funding options have emerged and continue to exist, as elaborated further in chapter 6. Nonetheless, these funding alternatives are temporary and non-sustainable, in contrast to inclusion in basic health insurance, which offers a lasting solution.

[2] Shortly before the publication of this study, namely on October 30 2023, there was the announcement that Mammaprint is now included in basic health insurance. This position report could not be included in the analyses. However a quick check showed that the interviewees were indeed right. In the draft position report of 2023 the Mammaprint test got again a negative assessment, however after reactions of various actors on the draft position report, the ZiN reconsidered and did assessed Mammaprint positively in the final position report [161]. Mammaprint is now thus reimbursed from basic health insurance for a specific group of breast cancer patients.

The Reimbursement Arena is Critical

These findings confirm that the inclusion in basic health insurance remains a critical decision arena for the adoption and implementation of Mammaprint in the Netherlands, both in the past and at present [3]. Moreover, the case study interviews and documentation have unveiled significant factors that influence adoption and implementation. As a result, this arena is subjected to further analysis, detailed in Chapter 5.

[3] Shortly before the publication of this study, namely on October 30 2023, there was the announcement that Mammaprint is now included in basic health insurance. This position report could not be included in the analyses. So at the time of the analyses, the arena was critical because Mammaprint was not included in basic health insurance. Currently, after 13 years since the first assessment by the ZiN, this is no longer the case and so the arena is no longer critical.

5

Case study Reimbursement in Basic Health Insurance Arena

When addressing the second subquestion, 'Which decision arena(s) are critical for the adoption and implementation of a specific AI-CDSS in cancer care?' the subsequent discoveries were made. Reimbursement within basic health insurance currently stands as a critical arena. Presently, Mammaprint, the AI-CDSS in question, is not covered by basic health insurance [a], and there is considerable disagreement surrounding this matter. Moreover, it has been highlighted that inclusion in medical guidelines also holds significance for adoption and implementation. However, at present, it is not considered a critical arena, and minimal dissent has been observed in previous decision-making rounds possibly in part due to the fact that there is insufficient available documentation on this topic. Given these circumstances, the investigation will focus on the reimbursement by basic health insurance arena, serving as the subject of subquestion 3: 'What factors influence the adoption and implementation of this specific AI-CDSS in those critical decision arena(s)?' It is hypothesized that factors such as existing procedures, informal institutions, power distribution among actors, and the underlying interests and values of actors play an important role in the inclusion of Mammaprint in basic health insurance. Unraveling these aspects will provide insights into the factors contributing to the impediments in achieving Mammaprint's inclusion in basic health insurance. To comprehend these factors, the institutional actor analysis framework is further employed, as depicted in figure 5.1 The data for this analysis comes from grey literature and case study interviews.

[a] Shortly before the publication of this report, the decision of the 2023 decision round on reimbursement of Mammaprint in basic health insurance was made and had a positive outcome. Mammaprint is thus now included in basic health insurance for a specific group of breast cancer patients.

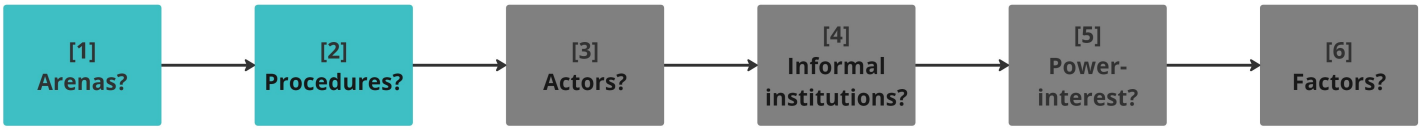

Figure 5.1: Institutional Actor Analysis

5.1. General Procedures of Reimbursement Arena

The institutional approach involves reviewing the existing policy-making structures [81]. First the general procedures for reimbursement by basic health insurance will be investigated and expained in this section.

Dutch Healthcare System Intramural vs. Extramural Care

The Dutch healthcare system is organized in such a way that everyone must have healthcare insurance. Everyone, therefore, has access to the care included in the basic package. The basic package is the same for everyone, regardless of which health insurer you are insured with. However, insurers can offer additional insurance and decide what else they offer on top of the care in the basic package. To get into the basic health insurance package in the Netherlands, there are procedures. These distinguish between extramural and intramural care. Extramural care includes drugs that are available from pharmacies on a doctor's prescription [162]. The assessments on extramural care go through the Dutch Healthcare Institute (ZiN), which in turn advises the minister of health [163]. Final decisions on extramural care are always made by the minister of health. Intramural care includes treatments and medications used in the hospital [162]. Within intramural care, there is also a 'lock of expensive drugs' [164]. Only intramural drugs or treatments that exceed legally set price limits are placed in the lock. The ZiN advises the minister of health on care that is in the 'lock of expensive drugs' and the minister of health ultimately decides on this. All intramural care that does not enter the lock is initially assessed by the health insurers according to legally established criteria determined by the ZiN and can then be reimbursed from the basic package after a decision by the health insurers. The moment it is unclear to the health insurers whether the care belongs to the basic package, the ZiN is called in [163]. The ZiN can then decide whether it should be included in the basic package or not and has the legal power here. Once the ZiN issues a position on this, it is care that is reimbursed from the basic package, so the minister of health does not decide on this and no signature is needed from the Minister.

Assessment Procedure Intramural Care

For 'normal' intramural care, so which is not in the 'lock of expensive drugs'. The assessment is thus done in the first place by the health insurance companies according to the legally established criteria of the ZiN. In case of doubt, the ZiN will assess it instead of the health insurers. The ZiN then has the final decision power without the intervention of the minister of health. In the first instance, intramural care must meet the legally established criterion of 'commit to provide' (CtP) [165] and in the second instance it must meet the 'state of science and practice' (SWP) [89] , which covers the first package principle 'effectiveness'. In addition, the advisory committee package (ACP) [166] of the ZiN can look at the 3 other package principles, cost-effectiveness, necessity and feasibility. However, this is done only if it appears that the care meets the previously mentioned criteria, CtP and SWP, and only at the time when there is reason to do so [166]. The criterion CtP means that the care is professional and is also considered as such by the relevant professional group [165]. Here the medical guidelines and stan-

dards of the professional group are primarily considered. This shows that the decision-making arena in which it is decided whether care will be included in the medical guidelines influences the decision-making arena as to whether care will be included in the basic package or not. Usually, the assessment against the criterion of CtP plays a secondary role in the assessment, because it is usually clear or beyond doubt that the intervention to be assessed falls under the domain of one of the professional groups mentioned in the health insurance decree [167]. It is then considered whether the care meets SWP meaning whether the care is effective. The principles of evidence-based medicine are used [89]. Should care meet both criteria, only then the other 3 package principles can be considered. However, this is only done if it seems necessary.

Consultation of parties during the assessment procedure

Once an assessment is started by the ZiN, there is an opportunity for relevant and invited parties to have input on the PICO question [89]. The PICO describes the research question in terms of patient group, intervention type, comparison to standard of care, and the relevant outcome measures to focus on [89]. Stakeholders include the medical scientific associations, patient associations, hospital associations, registrants of the form of care under consideration and the health insurers [92, 93]. After that, the ZiN will do the assessment based on the SWP [162]. Should it become apparent in between that input from external parties is needed, it can be requested [162]. Once the draft report is ready, it is sent to the scientific advisory board (WAR) [162]. The WAR advises the ZiN and the draft position report is adjusted accordingly, the advice is almost always followed according to interviewee 17. After this, the draft report is sent to the interested parties and they can respond to it in writing. On the basis of these reactions, the ZiN may carry out further research or make adjustments and/or ask the WAR is again for its advice. After this the board of directors of the ZiN makes the final decision. Once the ZiN board has decided, a position report is published. This position report includes the reason for the assessment, the assessment itself, the reactions of interested parties and the reactions of the ZiN [92, 93], the final position report, and practical information about when the decision will take effect. The ZiN responds substantively to the parties' reactions if the final position differs greatly from the parties' opinions.

Influence of the Dutch Medical Guidelines on the Assessment Procedure for Reimbursement and vice versa

When assessing whether care meets the CtP criterion, medical guidelines can be looked at [165]. According to the procedures [165] and validated by interviewee 17, if the form of care appears in the guidelines, it appears that the professional group uses the care and it often meets the CtP criterion without discussion. In addition, according to interviewee 17, the ZiN looks at the medical guidelines to check what literature is referred to as substantiation for the use of a particular medication or test and includes this literature in its own literature analysis. It also looks at other possible treatments in the guidelines to see if the new form of care is an improvement on current (or actually former) practice.

According to interviewees 14 and 15, the ZiN itself can also provide input during the drafting of the guidelines. This is possible in the beginning at the time bottlenecks are inventoried. In addition, the ZiN is sent a draft version of the guidelines for comment. However, the guidelines really belong to the medical specialists, so what is done with those comments is up to the guideline committee. As much as possible, medical guidelines take into account which care is reimbursed and which is not. However, exceptions to this are possible in case medical specialists do consider it as best practice. The following is stated in the breast cancer guidelines about this:

A quality standard describes what constitutes good care, regardless of the source of financing (Health Insurance Act (Zvw), Long-Term Care Act (Wlz), Social Support Act (Wmo), supplementary insurance or co-payment by the client/patient). Inclusion of a quality standard in the Register, therefore, does not necessarily mean that the care described in the quality standard is insured care, nor that it meets the SWP and is therefore eligible for reimbursement by Health Insurers or Care Offices. [158]

5.2. Mammaprint Specific Procedures of Reimbursement Arena

Mammaprint falls under intramural care and specifically the 'normal' intramural one. The general procedures for the assessment of intramural care are described in the previous section (5.1). In this section, the focus will be on the specific case of Mammaprint to understand if and how the general procedures play out. Previous decision rounds and the current decision-round are studied by analyzing publicly available position reports of the ZiN [91, 92, 94], which also include responses of stakeholders on the draft position reports in the appendix, and by gathering information from case study interviews. The assessment of Oncotype in 2021 [93] is also examined because Mammaprint and Oncotype are highly similar tests and are presently being assessed in conjunction [94].

History of Assessments

RASTER Study

The RASTER study [168] prospectively assesses the feasibility of implementing Mammaprint in community-based settings. It also investigates the effects of Mammaprint on treatment decisions about adjuvant chemotherapy in combination with treatment advice in the breast cancer guidelines. The RASTER study was financially supported by the CvZ (former ZiN) [168]. Interviewee 10 and 12 indicated that little was done with the results by the CvZ, although the CvZ did pay for the study. Of the 22 authors of the study, 15 were working at the NKI-AVL and 4 were additionally or solely working at Agendia.

5.2.1. Assessment Mammaprint 2010

Reason for assessment

The inclusion of the Mammaprint and Oncotype tests in the basic health insurance package in the Netherlands is considered as decision arena. This decision was first the responsibility of health insurers, however, a dispute arose about Mammaprint, the responsibility for assessment shifted from the health insurers to the ZiN. This dispute was about reimbursement of the cost of a Mammaprint and was addressed to CvZ, the old name of the ZiN, by Health Insurance Complaints and Disputes Foundation (SKGZ) in November 2009 [91].

Procedures & Outcome of assessment

At the first assessment in 2010, assessing medical tests in the context of diagnostics and/or prognostics was something new [91]. The CvZ was used to assess therapeutic interventions through evidence-based medicine methodology, as described in the SWP [89]. During the 2010 assessment, the frameworks for assessing medical tests were set, based on the already used methods for therapeutic interventions. Just as therapeutic interventions, medical tests should be assessed by clinical utility. Clinical

utility refers to whether the patient actually experiences health benefits as a result of using the test [91–93]. As described in the SWP, the level of evidence should ideally be RCTs. However, it is possible to still allow care based on research of a lower level of evidence , but it must be well-argued. This already applied to therapeutic interventions and therefore also applies to medical tests [91].

The claim of the Mammaprint test is that it can more accurately classify patients at high risk of metastasis [91]. This could avoid the treatment, side effects, and cost of adjuvant chemotherapy for a group of patients who would otherwise be classified as high-risk using conventional methods. Only retrospective studies were available at this time, and based on these, Mammaprint appears to be a promising predictor of the development of distant metastases within 5 years, according to the CvZ [91]. The RASTER study was part of the literature review of the assessment. Mammaprint also appears from these studies to be a more accurate prognostic tool than the CDSS used at the time for risk estimates, according to the CvZ. However, according to the CvZ, no results are available that prove health benefits and thus clinical utility. For this, according to the CvZ, prospective research is needed and there is none yet. The MINDACT trial [160, 169] is underway and the CvZ will reassess when the results of that study are known. The MINDACT trial must show whether or not the choice of treatment based on the Mammaprint can be made safely and leads to health gains. Cost-effectiveness has not been considered in 2010 [91].

Interaction with stakeholders during the assessment

The CvZ was in contact with external parties. So contacted the CvZ the Dutch Association for Medical Oncology (NVMO), Dutch Internists Association (NIV), Dutch Association for Surgery (NVvH) and Dutch Association for Pathology (NVVP) to review the draft position report [91]. They suggested two experts to review the position report and the CvZ incorporated their comments. In addition, the NVMO asked to postpone the publication of the position. This as the medical breast cancer guidelines were currently in review and big changes would be made to the module about Mammaprint [91]. However, the CvZ does not comply because there is an ongoing SKGZ dispute and they are therefore obliged to publish the position report [91].

SWP For Medical Tests

Until 2011, the focus of the CvZ primarily on the assessment of therapeutic interventions and drugs. Because of a steady increase in medical tests and the questions this raised about whether or not they belong in the basic package, the CvZ developed a more explicit framework for assessment [92, 170]. CvZ published the report 'Medical tests (assessment SWP)' [170], which was an elaboration on SWP [89], the review framework for care.

MINDACT Study

The MINDACT is a multicentric, prospective RCT [160]. MINDACT stands for Microarray In Node negative Disease may Avoid Chemotherapy. It was the first prospective clinical validation of a genomic tool [169]. The trial was designed and led by the European Organisation for Research and Treatment of Cancer (EORTC). This study includes almost 7000 patients from 9 European countries in 112 institutes. The enrollment for the trial was between 2007 and 2011. In 2016 the 5-year follow-up data was published [160]. In 2021 the 9-year follow-up data was published [171]. The MINDACT study is mainly funded by subsidies from the European Union (over €4 million) [155, 160, 171] and many national and

international (breast) cancer foundations [160, 171]. Agendia provided the genome analysis without costs [160]. Some of the researchers that conducted the MINDACT study are (part-time) employee of Agendia or stockholder and/or patentholder of Mammaprint and/or employee at the NKI-AVL.

5.2.2. Reassessment Mammaprint 2018

Reason for reassessment

Mammaprint was not reassessed by health insurers, but remains with the ZiN, former CvZ, for reassessment. This is a trap in the procedure, interviewee 19 and 17 indicate that as soon as a drug or medical test is under review by the ZiN, it remains there. This is not officially established but is an informal rule. A reassessment was done in 2018 due to the availability of new study results from the MINDACT study [160]. The EUnetHTA [172] served as the basis for the ZiN's assessment. The SWP for medical tests published in 2011 [170] was used for assessment [92]. The ZiN conducts assessments to evaluate the clinical effectiveness of interventions or, in the context of medical tests, the clinical utility of the test plus treatment. This involves determining whether the combined use of a test and treatment offers greater health benefits to patients when compared to standard care. The preference is still given to RCTs, however other prospective studies can under certain conditions be used as supporting evidence [89, 92, 170].

Procedures & Outcome of assessment

ZiN assessed whether adding the Mammaprint to the standard test Adjuvant! Online (AO!) [173] in women with early-stage breast cancer leads to health gains when they forgo additional chemotherapy based on the Mammaprint result, while the standard test recommends chemotherapy. However, it has been concluded that the Mammaprint test may lead to significant additional mortality in women who decide not to take chemotherapy in response to a low-risk outcome of the Mammaprint test and still develop metastases. This possible additional mortality does not compensate for the benefits of avoiding chemotherapy according to the ZiN. This has thus led to the conclusion that the clinical utility, namely health benefit, has not been demonstrated.

Interaction with stakeholders during the assessment

The ZiN has at a number of times given the opportunity for parties to provide input. First, during the EUnetHTA assessment [172], which was actually a preliminary process of the assessment [92], the following parties attended a scoping meeting in March 2017: Professional groups: NVMO, NABON; patient associations: Breast Cancer Association of the Netherlands (BVN)/ Dutch Federation of Cancer Patient Organizations (NFK); Health Insurers Netherlands (ZN); Hospitals: Dutch Association of Hospitals (NVZ) [92, 172]. At this scoping meeting, the assessment process was explained and the PICO was discussed. According to the ZiN, the parties present agreed with the PICO [92]. Furthermore, the ZiN indicates to have asked for expertise from the other parties during the assessment procedure, when they deemed it necessary. How often and when this happened is not clear. In addition, when the draft position report was ready, the ZiN sent the draft position report to all stakeholders considered relevant by the ZiN. These stakeholders are the medical specialist associations that are in the working field where Mammaprint is applied, BVN/NFK, ZN, the Dutch Federation of University Medical Centers (NFU), NVZ and Agendia, the registration holder of Mammaprint. The following parties responded in writing in February 2018: NVMO/NABON, NVVP, Dutch Association for Surgical Oncology (NVCO), Dutch Association for Radiotherapy and Oncology (NVRO), BVN/NFK and Agendia. All parties who responded disagreed with the ZiN's position [92]. Normally, the ZiN responds to these views only in

writing. However, in this case, by exception, the ZiN held a consultation meeting in April 2018 at the request of the responding parties [92]. This meeting provided an opportunity to clarify the viewpoints of both the stakeholders and the ZiN. In May 2018 ZiN also responded in writing to all written responses. Despite all the comments from many stakeholders, the ZiN stuck to its initial position.

TAILORx

There is one study that played an important role in the assessment by the ZiN of Oncotype. This is the Trial Assigning Individualized Options for Treatment (Rx), or TAILORx trial [174]. The TAILORx is a multicentric, prospective, RCT, like the MINDACT [160, 169]. Patients were recruited from 2006 till 2010. The trial includes almost 10000 patients with around 8 years of follow-up data. The outcomes were published in 2018. The study was mainly funded by the National Cancer Institute of the U.S. Two of the authors were working for Genomic Health, the registration holder of that time. One of the authors is patentholder of the test, but all rights are transferred to Genomic Health [174].

5.2.3. Assessment Oncotype 2021

Reason for assessment

Because Mammaprint and Oncotype are very similar tests, Oncotype also followed the trap in the procedure and is directly assessed by ZiN and not by health insurers [93]. The reason for the assessment was the publication of the RCT, namely the TAILORx study [174]. This assessment again uses the SWP for medical tests framework [170] that was published in 2011. The ZiN applied the same conditions for study inclusion as applied during the 2018 review of Mammaprint [93].

Procedures & Outcome of assessment

ZiN assessed whether adding the Oncotype to the standard test AO! in women with early-stage breast cancer (ER+/HER2N-) leads to health gains when they forgo additional chemotherapy based on the Oncotype result, while the standard test recommends chemotherapy [93]. It has been concluded that there is evidence of sufficient quality that the Oncotype test leads to health benefits while weighing the positive (no side-effects of chemotherapy) and negative effects (possible lower survival). Therefore clinical utility has been proven, so the test meets the SWP and the test will be included in basic insurance. Reimbursement is only for a certain subgroup of women, namely women with early-stage breast cancer (ER+/HER2-) and a clinically high risk who meet the following conditions: - older than 50 years AND; - N0 (gland negative) AND; - grade 1 tumor with a size between 3.1 and 5 cm OR; - grade 2 with a size between 2.1 and 5 cm OR; - grade 3 with a size between 1.1 and 2 cm [93]. The ZiN decided to not reimburse the test for women younger than 50 years and women with grade 3 tumor larger than 2 cm. This is because in these cases there may well be a high probability of benefit from chemotherapy according to the ZiN.

Interaction with stakeholders during the assessment

The ZiN has at a number of times given the opportunity for parties to provide input. ZiN did not hold a meeting for the PICO, but did send the concept PICO to the same parties as during the assessment of Mammaprint in 2018 and to Exact Sciences, the registration holder [93]. It did it in this way as both PICOs show great similarities with the PICO for Mammaprint [93]. They incorporated the parties' responses or argued why they were not incorporated. Furthermore, the ZiN indicates to have asked for expertise from the other parties during the assessment procedure, when they deemed it necessary.

How often and when this happened is not clear. The draft position report was sent to all relevant stakeholders, which were the same parties as for Mammaprint, except for the registration holders. The draft position report stated that Oncotype did not meet the SWP [93]. The following parties responded in writing: BVN/NFK, professional groups: NVvH, NVMO, NVVP, Association of Clinical Genetics Netherlands (VKGN); ZN, and registration holder Exact Sciences [93]. All parties disagreed with the draft position report, except ZN indicated agreement. The ZiN modified the draft position in favor of Oncotype and submitted it to the WAR. The WAR agreed with the adjustment. The ZiN did not return letters to the parties that responded to the draft position report, but included in the final position report a section where it responded to the comments of the parties.

5.2.4. Reassessments Oncotype and Mammaprint 2023

Currently, a new assessment has been running since October 2022, the results of which are expected to be announced in the 4th quarter of this year (2023) [1]. The same kind of procedures apply as before. The draft report has already been sent out to the various parties. Based on sayings from involved parties, it seems that Mammaprint would be rejected again in the draft position report. However, during the interviews, it emerged that comments have been received from at least two key players in favor of Mammaprint. They do think the ZiN is reconsidering its draft position based on this feedback [1], but don't know what the effect will be. Also, given the fact that the release of the position paper was initially scheduled to take place in May according to the ZiN site, but this was later changed on the website to the 4th quarter of 2023, it seems that the ZiN is considering the comments. What the draft position report says about Oncotype did not emerge during the interviews.

[1] Shortly before the publication of this study, namely on October 30 2023, there was the announcement that Mammaprint is now included in basic health insurance. This position report could not be included in the analyses. However a quick check showed that the interviewees were indeed right. In the draft position report of 2023 the Mammaprint test got again a negative assessment, however after reactions of various actors on the draft position report, the ZiN reconsidered and did assessed Mammaprint positively in the final position report [161]. Mammaprint is now thus reimbursed from basic health insurance for a specific group of breast cancer patients.

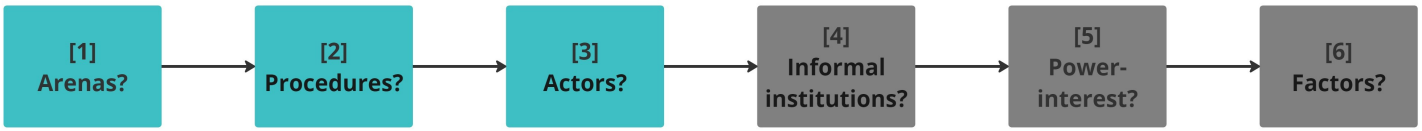

Figure 5.2: Institutional Actor Analysis

5.3. Actors

With this institutional approach an initial list of actors could be identified. The ZiN is identified as a key player as it makes the final decision about reimbursement. In the first instance, the actors that have a say in the decision making-process are seen as other key players. Actors have a say in case they are able to give input to the ZiN. This means they either have a say during the drafting of the PICO and/or they received the draft position report and gave feedback on it in at least one of the decision rounds. This leads to a list with actors as shown in table 5.1.

Actor	Representing
ZiN	All health insured persons
Medical Specialist Associations: NVMO, NVVP, NVCO, NVRO, NVvH, VKGN	Medical specialists in breast cancer care
NABON	Medical specialist associations in breast cancer care
BVN/NFK	Breast Cancer Patients
ZN	Health Insurers
Agendia	Registration holder
Hospital societies: NFU, NVZ	Hospitals

Table 5.1: Actors identified with an institutional analysis and the corresponding groups they represent

Subsequently, an examination is conducted into the functioning of the ZiN, which operates on behalf of the Ministry of Health while retaining its own statutory authority. Once the ZiN makes a decision, it stands without requiring intervention from the Minister. Nevertheless, it is the minister of health who formally represents the ZiN's stance in the event of parliamentary inquiries. Additionaly, a former minister of health did handed out the Health Innovation Award to Mammaprint in 2008 [74]. Additionally, A search through parliamentary records reveals that various Ministers of Health have received questions from the House of Representatives regarding Mammaprint [175–178]. These inquiries primarily revolved around the reasons for Mammaprint's exclusion from basic health insurance, with questioners expressing dissent and concern over the prolonged process. Thus, it is evident that both the minister of health and the House of Representatives play a role in the arena.

Furthermore, the NKI-AVL's involvement comes to light. As previously mentioned, the Mammaprint test was invented at the NKI-AVL. Interviews also reveal that the NKI-AVL is a highly influential hospital and research institute in the field of cancer care in the Netherlands. Additionally, it is found that many individual researchers also have a direct or indirect role in the arena. More than half of the spokespersons representing medical specialist associations during the 2018 consultation round [92] are affiliated with the NKI-AVL and 3 of the 10 are employees at the NKI-AVL. Moreover, the study coordinator of the

MINDACT study is also associated with the NKI-AVL [160], and many researchers from the RASTER study [168] and the MINDACT study are also affiliated with the NKI-AVL. While these researchers may not have direct influence within the arena, their studies are used to inform decisions within the arena.

Based on these findings, the minister of health, the House of Representatives, and the NKI-AVL are also considered actors within the arena. Therefore the elaborated list of actors is shown in table 5.2.

Actor	**Representing**
ZiN	All health insured persons
Medical Specialist Associations: NVMO, NVVP, NVCO, NVRO, NVvH, VKGN	Medical specialists in breast cancer care
NABON	Medical specialist associations in breast cancer care
BVN/NFK	Breast Cancer Patients
ZN	Health Insurers
Agendia	Registration holder
Hospital societies: NFU, NVZ	Hospitals
Minister of Health	Ministry of Health
House of Representatives	Voting-Age Dutch population
NKI-AVL	Management, Researchers and Medical Specialists working at NKI-AVL

Table 5.2: Identified Actors and the corresponding groups they represent

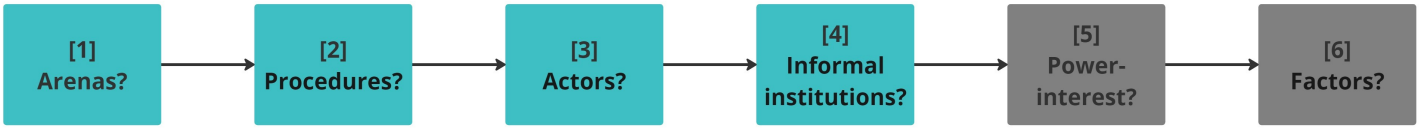

Figure 5.3: Institutional Actor Analysis

5.4. Informal Institutions

It is clear from section 5.2 that complying with SWP is the impeding factor for Mammaprint in the reimbursement arena. Furthermore, it became clear that medical specialist associations, the NABON and Agendia disagree with the ZiN's decisions [92]. To find out what exactly are the underlying factors that cause Mammaprint not to comply with SWP, this was discussed with all interviewees that were in a way involved in the reimbursement decision-making process of Mammaprint. In addition, the ZiN reports and key players' responses to the draft proposition reports, which are in the appendix of those reports [92, 93], were carefully reviewed. From these interviews and reactions on the draft position reports, the following factors came up frequently:

5.4.1. No clear requirements for inclusion of diagnostic tests

There are no clear requirements for the inclusion of diagnostic tests in basic health insurance. According to many interviewees (3, 5, 6, 10 and 12) it is not clear what the ZiN wants to see as evidence of clinical utility for diagnostic tests in general. Interviewees 4 and 12 mention that the requirements for Mammaprints are changing between the decision rounds. Furthermore it is mentioned by interviewees 6 and 19 that many steps in the assessments are debatable and not objective. It is indicated that the ZiN does not properly explain the subjective choices they make, including the statistical cut-off values. All in all, the procedure is seen as unclear.

According to interviewee 6, who himself was not directly involved in the approval of the Mammaprint, but did follow it

Well, so there is some subjectivity of interpretation of the assessments I feel, so yes, it depends on who you ask, who thinks that there's enough added value to include a Mammaprint in the healthcare system, in basic insurance.

And that of course is actually crazy that there are no hard end lines or points or that those are then interpretable again.

But if it's clear how that path develops and what you have to do to do that and if admission to care is even more transparent, which we talked about earlier, then that can help.

According to interviewee 12 it is a moving-goal-post.

A moving-goal-post, so every time we had to comply with something different because it was new, the government didn't really know what kind of requirements they wanted to put on it either.

The same is actually said by interviewee 4:

The ZiN initially raised reasons for not including it, and each time, those arguments were addressed. The first time, there was a small study where everyone received Mammaprint and they said, "Well, that's great, we'll evaluate it." Then it was stated that a randomized study was needed. So, a randomized study was conducted, and then they said, "Your follow-up is too short, and quality of life hasn't been measured." Now, with longer follow-up and quality of life data, they're saying, "But we're still going to separate lymph node positive and negative cases." It seems like there's a new argument each time. The issue of lymph node positivity and negativity played a role in the latest evaluation, and there has been a lot of criticism about it. They are re-evaluating it now and considering them together.

There was some inconsistency in the way Mammaprint was assessed in 2018 and Oncotype in 2021. For example, MINDACT study was not downgraded due to changing medication regimes in chemotherapy [92], while TAILORx was [93]. Medication for breast cancer is changing rapidly and there is a difference in the types of chemotherapy administered to the patients during the 10-year follow-up. Therefore, according to ZiN, there is a bias in the study for certain chemotherapy, which is used less nowadays. Because of this reason, in the draft assessment, they rated the quality of evidence on survival low for the TAILORx study. However, they had not done this in the case of the MINDACT study, even though the same thing was at play here. There were a lot of comments from other key players on the way of assessment [93], but on this particular aspect, the ZiN finally indicated to reverse this and therefore deemed that the clinical utility was proven for Oncotype [93].

The response to the draft position report from Exact Sciences, the registration holder of Oncotype, was the following:

First, we note in this context that it is simply impossible to provide data with 9 years of follow-up (patients in TAILORx were enrolled between April 2006 and October 2010) while ensuring that the chemotherapy treatment in the study is in line with the most up-to-date recommendations in guidelines. The ZiN indicated at the time that this was ultimately the deciding factor in allowing Oncotype to be included in the basic health insurance package.[93]

the ZiN does give in on this argument and therefore changes its position after the draft position report and writes the following in its report [93]:

Confidence in data on survival was low in the draft assessment. When confidence in the data on survival is low, the ZiN is of the opinion that data on quality of life cannot be decisive. Based on the parties' responses and reweighing the arguments, the ZiN concludes that the earlier GRADE assessment was too stringent. As confidence in non-inferiority data has increased, quality of life data play a more important role.[93]

It did become clear from the interviews that there are also explanations as to why there are no clear requirements for diagnostic tests. Assessing diagnostic tests is something very new and interviewees (1, 5, 12) indicate there is missing expertise in this area. The ZiN is specialized in assessing medication and this is something where several interviewees (1, 3, 5, 12) indicate they do a good job. However,

assessing diagnostic tests is something very new for the ZiN. Mammaprint was the first diagnostic test they reviewed (confirmed by interviewee 17). The ZiN did issue an SWP for medical tests in 2011 [170], which is the same method as for medication [89], but made more explicit for medical tests, and that document is still used today. Little has changed over the years in the assessment of medical tests by the ZiN. However, the ZiN regularly updates the general SWP [89] (which also applies to the assessment of medical tests) to make the procedure more transparent.

The ZiN (interviewee 17) states the following about changes in their procedures:

We did recently release a new version of the SWP and we do always keep track of if there are new insights, like over time GRADE has come as a way to assess quality of studies. We didn't have that before either. At this moment there are no adjustments in assessing tests. If new insights appear, we will assess if our methodology needs adjustments.

The update of the update of SWP report is mostly about how we get from evidence to conclusion. It is not a different way of assessing but more transparency about how the weighing of positive and negative effects, the quality of the evidence, arguments of appropriate evidence and medical arguments.

In the consultation meeting with the ZiN in 2018, the following was said by Gabe Sonke on behalf of the NVMO/NABON about the experience of the ZiN with drugs and with diagnostics [92] :

The MammaPrint is now well-established in practice. It is being used properly, not for everyone. The practice contrasts with the report. The practice is clear. There is experience with drug reviews but hardly any with diagnostics. This requires a different way of assessment. There is not only one way that leads to Rome.

In the reaction of NVvH on the draft proposition report in 2021 of Oncotype (which had a negative outcome in the first place), the following was stated about the way of assessing diagnostic tests [93]:

This letter also points out that diagnostic tests have difficulty passing ZiN assessments. We endorse this. It seems that evidence is only looked at from an epidemiological point of view and that medical knowledge regarding content is lacking. The latter is then leaned on the input of the scientific associations, but this is almost always only at the end of a pathway (ductoscopy excepted). Moreover, it often seems that this input ultimately no longer has much influence on the final advice. Naturally, we find this regrettable and are happy to discuss a better and optimal study design for such tests.

Interviewee 12 indicates that the ZiN has no expertise in diagnostic testing, but that they are also not changing in this area.:

And so, it's, it's just very difficult, especially if you're very far ahead like we are or you're the first then the government also doesn't have the expertise to just properly understand how that should be evaluated.

They just don't understand the diagnostics. And, That's a bit of a stepchild

You'll also see almost no diagnostic tests are evaluated by the ZiN. just with that medication cap on you just can't evaluate a diagnostic test that's just yes it lacks that expertise in that regard so to speak. But the fact that so in the medicine area they just make sure that the care remains affordable in the Netherlands.

Yes, we are just a small drop in that bucket, so why should they change anything based on that?

But again, it's diagnostics, it's a stepchild. It gets limited attention in the scale of things that they're doing. Yeah, you know, all those pharma companies that are screaming blue murder at the ZiN, yeah, they're generally taken less seriously and rightly so, because I just know that generally, the effect of most drugs is just pretty small at a price tag that's pretty high and that's where the ZiN just does a good job. It's the diagnostic side for which there are not quite the right procedures and that's getting a little bit better by now, just richly too late for Mammaprint.

Interviewee 1, who involved in the assessment procedure of another diagnostic test, says the following:

My own experience was with one particular procedure, also a diagnostic procedure, that they didn't understand what it was about at all, and only when we brought in someone who started kind of lobbying, then they turned the cart and said maybe we should look at it from a different angle.

We have learned that they (the ZiN) think fairly rigidly and are particularly focused on the value of treatment, of drugs, there are very tight agreements about that, which studies have what cut-off points and how do you deal with that, but when it comes to diagnostics it becomes a lot harder.

The ZiN (interviewee 17) indicates that the procedure of assessing medicines is somewhat ahead in terms of the documents to be provided by manufacturers. Other than that, however, the ZiN says rather little about how advanced the way of assessing of medical tests is.

With pharmaceuticals, there is always a manufacturer who needs to provide a complete file including a model for assessing cost-effectiveness model. In case of medical devices and tests is that not standard yet, but steps are made to assess cost-effectiveness in medical devices and medical tests as well. In that respect, of course, pharmaceuticals are a little further ahead than the other branches

5.4.2. (High) Level of burden of proof needed

A high burden of proof is preferred by the ZiN. A 10-year follow-up and an RCT, which is the highest level of clinical evidence, is preferred [92, 93]. However, according to the medical scientific associations in the arena, BVN/NFK and Agendia, 5 years is enough [92]. Since only the 5-year data were available [160], the ZiN's assessment in 2018 was based on surrogate outcomes for a 10-year follow-up [92], which were also surrogate outcomes where the medical scientific associations, BVN/NFK and Agendia disagreed with [92]. The surrogate outcomes subsequently gave less confidence and were not in the preferred range according to the ZiN [92]. For Oncotype 9-year follow-up data was available [174], however, that this was 9 years instead of 10 years eventually did not affect the assessment in 2021 [93]. The ZiN only included RCTs in both the 2018 (MINDACT [160]) and 2021 (TAILORX [174]) assessments. However, after comments on the draft position report, it did also include a prospective study about the quality of life in 2018 [92, 179], but this did not change the final conclusion of the ZiN.

From the interviews, there appeared to be a number of underlying explanations for the different opinions on the required follow-up duration.

Interviewee 12 mentions that a business case does become very complicated with such a long study duration. If investors would know this beforehand, they would not invest:

If you have to go wait 10 years for a prospective study and it takes you 4 years to enroll those patients, so that means if you develop a new test that the earliest you can possibly get reimbursed in the Netherlands is 14 years after the development of that test. Yes, of course, you can't build a business there. So if that really is the requirement, then that is, then I would advise everyone against ever doing anything, apart from the fact that such a prospective randomized study is extremely expensive, so yes someone has to pay. In our case, we just make use of investors, but when they see this, they also say this proposition is not fundable. I'm not going to give you money now because you might get reimbursement in 14 years. Yes, that's not how that works.

So that gives you a time frame in which it's just unrealistic to say develop an essay from a for profit setting look you can always say, we do it academically, miserable academically is that there isn't the money there either to do this so thoroughly. There's a reason why the NKI said, we'll make a spin-off of this. We're not going to do this ourselves. But so, as the ZiN requirement remains that you have to have prospective validation, then I see it bleak for the Dutch patient when it comes to access of advanced diagnostics.

Additionally, interviewee 12 also indicates it is an unreasonably high requirement to have to provide a prospective study as evidence and indicates that in U.S. this is not required.

Look, in the U.S., we were already reimbursed in 2009 And Oncotype also based on retrospective studies. In the Netherlands, it could only be with prospective studies and I think that is an unreasonably high requirement.

The medical specialists in the arena indicate that it does not match the clinical practice, they want to use new innovation quickly in practice. The following was stated in the report of the 2018 consulta-

tion meeting [92]:

Mr. van Diest (NVVP) concludes that there is a discrepancy between the field and the ZiN when one can speak of sufficient evidence. The field thinks it is sufficient and the ZiN wants the best evidence, but that is not realistic. He is somber about the future of biomarkers if the ZiN does not adjust its requirements. Mr. Rutgers also gets the feeling that there is a big difference in the principles of package management and practice. [92]

Gabe Sonke on behalf of the NVMO/NABON: Based on the results of the MINDACT, the NVMO/NABON believes there is sufficient scientific evidence.[92]

In response to the draft position report in 2021 the NVvP replies the following:

The choice to speak of mature data at a follow-up of 10 years does not match the clinical practice in which new treatments become available more quickly anyway. A 5-year follow-up would be more appropriate.[93]

In 2021, NVMO indicates that even larger and longer studies of gene expression profiles are not going to take place.

We further argue that there will be no randomized study of the value of gene expression profiling with larger numbers of patients and longer follow-up, and thus no data with a smaller confidence interval than that of the TAILORx study (the same is true of the Mammaprint in the MINDACT study).[93]

According to interviewee 3 and 12 these studies cannot be done again because of several reasons, besides simply being very expensive, taking a very long time, and being very difficult in practical and logistical terms, it is now also unethical to randomize because there is already enough evidence for the tests according to the interviewees.

5.4.3. Disagreement on study design between medical specialists and ZiN

There is disagreement in the arena on the study design between medical specialists and the ZiN. For both the MINDACT [160] and the TAILORx [174] study, only a subgroup of the patients was an appropriate RCT according to the ZiN [92][93]. Creating those subgroups results in a smaller patient population and therefore a bigger confidence interval, which results in both cases in non-inferiority. However, this has a lesser impact on the TAILORx study than on the MINDACT study, in part due to the reason that the TAILORx study included more patients (more than 10,000) than the MINDACT (almost 7000). For the MINDACT study, the primary part of the patient population was not included because according to the ZiN the MINDACT study compares survival rates in a specific group based on MammaPrint results and a non-inferiority threshold of 92% [92]. However, this analysis is one-sided and leaves according to the ZiN the question unanswered whether patients with high clinical risk and low genetic risk treated with MammaPrint have comparable survival to those treated based on AO! [92].

Interviewee 12 indicated the following during an interview: *Then they (the ZiN) said, yes, we actually don't think the design of that trial was good enough, even though the design was put together by the leading oncologists in Europe.*

The ZiN indicates that with the study design of the MINDACT, the research question of the ZiN cannot be answered [92]. It seems that here the medical profession has a different view than the ZiN. The medical profession considers the usefulness of Mammaprint proven in the MINDACT study and therefore includes it in the guidelines. The ZiN has higher requirements than the medical professions. According to interviewee 10, who was also involved in the assessment of Mammaprint, the ZiN and the medical specialists have another view on what the study design of clinical studies should be, which is also the case for other assessments by ZiN according to interviewee 10:

And, I have that all the time. You still see that in. Well, in all no matter what you're talking about then the clinical trials how they're designed, design of those trials those are almost never aligned with how actually ultimately the ZiN would like to see this and that's still true. That's very, very peculiar, that so the medical profession just has a different way of looking at it.

Interviewee 5 also indicates that the ZiN looks different to the design of clinical trials of diagnostic tests, which the interviewee related to the fact that they are used to assessing drugs and assesses diagnostiscs in the same way, while that might not fit the study designs for diagnostics:

ZiN is a somewhat complicated party in the sense that they are used to assessing drugs and you often do that with Overall Survival (OS) or Progression Free Survival (PFS) or response data, particularly PFS and OS and so when this diagnostic question came up, that's how they approached it and that's somewhat tricky because almost no studies have OS as an endpoint with a diagnostic biomarker, so that does make it very complex.

For the TAILORx study only the treatment of the moderate-risk patient group based on the Oncotype test is randomized [174]. The ZiN therefore only includes this subgroup. The genetic low-risk group is not given chemotherapy and the genetic high-risk group is given chemotherapy, they are thus not randomized. This as based on retrospective studies it would be unethical to randomize them according to

several parties that responded on the draft position reports and as mentioned by interviewees 3 and 12.

According to reactions on the draft position report in 2021 BVN/NFK and Exact Sciences call it unethical to randomize the low and the high-risk groups.

BVN/NFK: *We believe that with current knowledge about the prognostic value of these scores it is unethical to expect randomized trials for this*

Exact sciences:
Based on the prospective retrospective NSABPB20 [180] study have concluded that prior to the initiation of TAILORx, it had already been sufficiently demonstrated that patients with low-risk score do not benefit from chemotherapy while patients with high-risk score actually have a clinically relevant benefit from chemotherapy. Randomization of these groups in TAILORx could therefore no longer be justified and would not have been ethical.[93]

Nothing is yet known about the current decision round [2], and the ZiN also said it could not discuss this. However, three interviewees hinted at the current situation for Mammaprint as they did already received the draft position report. Once again, the ZiN seems to be bogging it down for Mammaprint on study-design grounds. The following was said by interviewee 12:

ZiN in their evaluation evaluated the lymph node-negative and the lymph node positive separately and said the confidence intervals are too big, but yes, if you cut the group in half and make it smaller, then the confidence interval automatically gets bigger. So then we got into another argument with the ZiN of yes, that clinical trial, it just has those patients that way, so you have to evaluate it based on that clinical trial, not your own way and that's just typical of the ZiN, that occasionally they do things their own way and they think that's the best way, but there's something to be said against that. So we are doing that now and I think we have convinced them now to say, no, you just have to look at that if it's one trial, then you have to look at that whole trial then you have to put that group so back together again and then the confidence intervals do satisfy so that that's going on now, so we'll see.

Few explanations emerged from the interviews as to why the ZiN has such a different view than the medical specialists. There seems to be much misunderstanding of the ZiN's approach. Parties indicate that they do not actually understand why the ZiN has such a different approach. A few interviewees suggest there is a knowledge gap between practice and policy making, which works both ways. It is also mentioned several times that the ZiN looks at things differently because the underlying motivation is different. The ZiN wants to keep care of accessible and affordable and does not want expensive, non-beneficial care in the package. Medical specialists want to provide the best care for the patient.

According to interviewee 6:

There is a huge knowledge gap between people in the field and policymakers. And also often with

[2] Shortly before the publication of this study, namely on October 30 2023, there was the announcement by the ZiN that Mammaprint is now included in basic health insurance. This position report could not be included in the analyses. However a quick check showed that the interviewees were indeed right. In the draft position report of 2023 the Mammaprint test got again a negative assessment, however after reactions of various actors on the draft position report, the ZiN reconsidered and did assessed Mammaprint positively in the final position report [161]. Mammaprint is now thus reimbursed from basic health insurance for a specific group of breast cancer patients.

people from insurance companies, we also see. These are people who understand insurance and know little about medicine, so there is a huge gap between them. Also from both sides. That's a barrier, I think. That should be much more shared.

According to interviewee 10:

The professional group of course wants the medicine to reach the patient and so does ZiN, but they also have a kind of perspective of, ultimately the citizen has to pay the premium and if that premium goes up because all the medicines come in that are too expensive and don't work well, then they also have a responsibility.

The ZiN (interviewee 17) says the following about this:

We are considering: is it added value if we are going to reimburse this from the basic insurance? Do we want to spend our premium money on this intervention from a solidarity perspective or could it be better spent elsewhere? We should only reimburse an extra test from basic insurance if taking this extra test leads to health benefits.

According to interviewee 13 it is important to have consensus about the study design before conducting studies like the MINDACT:

It is very important when you do this kind of research that you also try to coordinate in advance with the authorities who will have to assess it in what way they look at this kind of research

5.4.4. Priorities in Trade-off between Quality of Life and Survival Differ

While assessing the clinical utility of the tests, a trade-off must be made between the gain in quality of life on the one hand and the possibility of increased mortality on the other. This is because not administering chemotherapy provides a gain in quality of life, but also gives rise to the possibility that the cancer will still come back and someone will die from it, when this may not have happened if the patient had received chemotherapy. In this trade-off, medical specialists in the arena and the patient association prefer the quality of life, while the ZiN does not want to risk the possibility of more people dying [92, 93]. Until there is clear enough evidence for the ZiN that the probability of lower survival is very small and that there is explicit evidence of an increase in quality of life, they feel that clinical utility has not been demonstrated [92, 93]. Medical specialists and the patient association disagree on several points [92, 93]. First, they say there is sufficient evidence that it is safe to omit chemotherapy based on the Mammaprint and Oncotype [92, 93]. Second, quality of life and, in this case, the trade-off between quality of life and survival is in favor of the quality of life [92, 93]. Third, they find it odd that evidence is asked for that quality of life improves if a patient does not have to undergo chemotherapy, this is obvious according to the medical specialists in the arena [92, 93] and interviewees 3 and 12. The ZiN indicates that there must be very good evidence of an improvement in quality of life to accept the possibility of more deaths. They also recant on literature, where they say that in practice patients and medical specialists have the same acceptance limits as the ZiN:

Only if we can be sufficiently certain that this potentially significant loss in survival is compensated by a sufficient added value in terms of quality of life can we speak of clinical benefit. Patients indicated that the benefits of no chemotherapy are abundantly clear and they find it incomprehensible that we do not give a positive assessment of the positive effect on quality of life. [92]

... Not treating with adjuvant chemotherapy can lead to distant metastases that are no longer treatable and could possibly have been prevented with adjuvant chemotherapy. In this setting, as research also shows, we give more weight to survival than to quality of life. In the adjuvant setting, however, both medical specialists and patients will not readily accept a higher risk of death in exchange for better quality of life and less toxicity; it is known from the literature that most breast cancer patients find adjuvant chemotherapy acceptable in exchange for a limited increase in survival of ≤3% [92]

NFK/BVN further adds during the consultation meeting in 2018 that the advantages of not giving chemotherapy because of using the Mammaprint test is overlooked by the ZiN:

MammaPrint is designed to determine whether chemotherapy can be safely omitted. An important result, that 23% of patients can do without chemotherapy, is glossed over in the report. This means that out of every 100 patients, 23 are less exposed to the side effects of chemotherapy. These are well known and have both acute side effects but also long-term effects where cardiovascular problems and the so-called chemobrain come to mind. [92]

Measuring quality of life is something not included in the study design of the MINDACT [160, 169] or TAILORx [174]. In both cases quality of life was measured in a separately published study, but for a much smaller number of patients and for a much shorter period of time. However, HTA experts do seem to have wanted to put more focus on measuring the quality of life beforehand, but this was

almost not picked up by the study designer, the EORTC. This therefore seems to be a factor that not only played a role between ZiN and the medical specialists but also in the medical research world itself.

The following was said by interviewee 10:

and another sticking point at that time when the MINDACT study was set up, it was not yet common practice to assess the quality of life of patients during clinical trials. For the Health Technology Assessment (HTA), quality of life measures are necessary, for which permission was asked. This proposal was partly agreed upon for a subset of patients. It was only allowed to measure the quality of life once directly after the test results were available, and only for the Dutch participants in the trial.

During the assessment the ZiN really wanted to know quality of life of these patients and then including longer-term effects.

But yes, that was one of the shortcomings, because that wasn't standard practice then either to include quality of life in those kinds of clinical trials.

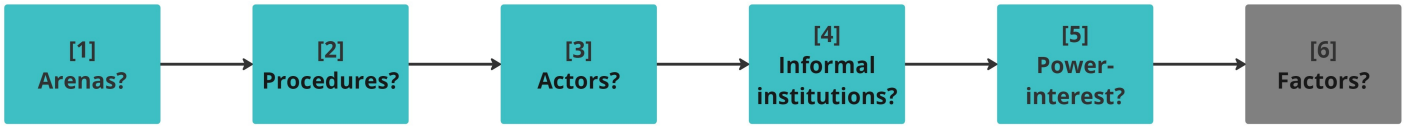

Figure 5.4: Institutional Actor Analysis

5.5. Power-Interest

To map the roles of actors, their level of interest in adoption and implementation, and their capacity to exert influence in the decision-making arena, a power-interest matrix can be highly informative. This matrix offers a clear visualization of these aspects. Therefore in this section the role of the actors in the arena will be specified and every identified actor will be mapped in a power-interest matrix (see figure 5.5). The interest of every actor is shown in table 5.3

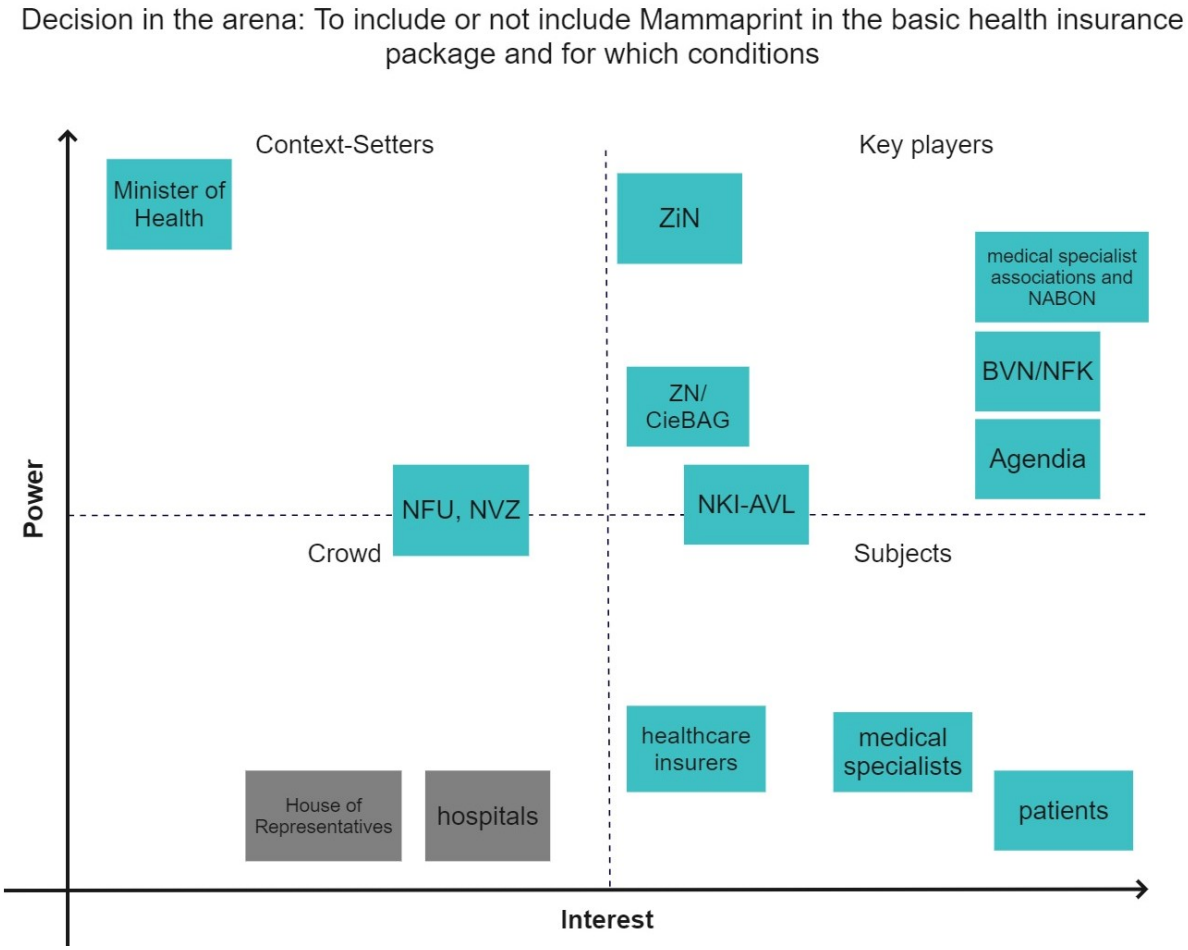

Figure 5.5: Power-Interest Matrix

Actor	Representing	Interest
ZiN	All health insured persons	Keep care affordable, accessible and of high quality
Medical Specialist Associations: NVMO, NVVP, NVCO, NVRO, NVvH, VKGN	Medical specialists in breast cancer care	Deliver high quality care to the patient
NABON	Medical specialist associations in breast cancer care	Deliver high quality care to the breast cancer patient
BVN/NFK	Breast Cancer Patients	Have high quality care (freely) available
ZN	Health Insurers	Keep care costs within budget
Agendia	Registration holder	Make a profit and improve breast cancer care
Hospital societies: NFU, NVZ	Hospitals	Deliver good care to patients, Keep costs of the hospital within budget, Make sure hospital is functioning in logistical and organizational terms
Minister of Health	Ministry of Health	Keep care affordable, accessible and of high quality
House of Representatives	Voting-Age Dutch population	Ensure fast adoption of promising innovation
NKI-AVL	Management, Researchers and Medical Specialists working at NKI-AVL	Innovate in cancer care and get innovation in cancer care quickly to patients, Keep their status as precious cancer center

Table 5.3: Identified Actors, the corresponding groups they represent and their interests

Based on reports of the ZiN [91], [92, 93], the website of the ZiN about the current decision round [94], a search on the website of the House of Representatives site, and the interviews with stakeholders an overview could be made from all actors in the decision arena. Interests of actors and informal rules within the arena could be mainly identified via interviews.

The ZiN is identified as a key player as it makes the final decision about reimbursement. In the first instance, the actors that have a say in the decision making-process are seen as other key players. Actors have a say in case they are able to give input to the ZiN. This means they either have a say during the drafting of the PICO and/or they received the draft position report and gave feedback on it in at least one of the decision rounds. All of these actors between the PICO and the draft position report do overlap except for the NFU and NVZ. The NFU however was only an observer during the consultation meeting in 2018 [92]. The NVZ declared the following during that same meeting:

Mr. Kemna explained that NVZ does not form an opinion on this, because it believes that a substantive response belongs primarily to the profession. First and foremost is the provision of good care to patients. In addition, there are organizational and financial aspects involved, and these turn out favorably in the case of Mammaprint.[92]

Based on this, NFU and NVZ are considered context-setters [81] as they have the power to influence the decision-making as they are invited around the table, but they do not use it as their interest is not high enough.

The following actors are thus considered key players:

- ZiN
- Medical Specialist Associations: NVMO, NVVP, NVCO, NVRO, NVvH, VKGN
- NABON
- BVN/NFK
- ZN
- Registration holders: Agendia, Exact Sciences

The following actors are thus considered context-setter:

- Hospital societies: NFU, NVZ

ZiN

Healthcare Institute Netherland (ZiN) decides upon reimbursement from basic health insurance for the Mammaprint and Oncotype tests. The ZiN is the gatekeeper in the arena, they decide who can enter round the table and who can not. The ZiN actually represents all health-insured persons in the Netherlands. The ZiN wants to ensure that the best possible care is provided, but the care remains affordable. The goal is quality, accessibility and affordability of the insured package [181].

Agendia

Agendia is the registration holder of the Mammaprint test. They have a financial interest. Agendia was founded in 2003 as a spin-off of the NKI-AVL. They got a lot of investments from private parties. In total 123 million dollars in 4 rounds ranging from December 2005 till July 2018 [182].

Exact Sciences

Exact Sciences is the registration holder of the Oncotype test. They have a financial interest. Exact Sciences is a molecular diagnostics company specializing in early-stage cancer detection. They are a multi-billion dollar company [183]. Upon the acquisition and merger of the company with Genomic Health in 2019, Exact Sciences became the registration holder of Oncotype.

Further, the following actors also play a role (see section 5.3), but more desk research and interviews had to be done to be able to determine their exact role in the arena.

- Minister of Health
- Medical Specialists
- Breast Cancer Patients
- Health Insurers
- NKI-AVL
- Hospitals
- House of Representatives

Minister of Health

The ZiN falls under the responsibility of the ministry of health as an independent administrative body. The ZiN has its own statuary mandate and can include Mammaprint and Oncotype without the signature of the minister of health. However, the minister has the power to overrule the ZiN. It can either exclude the test after approval by the ZiN and the minister can admit care to the basic package through ministerial regulation [184]. Interviewee 8 indicated that it is not acceptable for a minister to allow care that the ZiN has ruled does not meet the SWP. However, interviewees 8 and 17 mention it will occur and be accepted that a minister will exclude care that has been admitted because of costs that are too high. It is thus an informal rule that the minister of health might exclude care approved by ZiN, but will never include care rejected by ZiN. Interviewee 12 told that Agendia did try to arrange inclusion in basic health insurance via the minister after they received the healthcare innovation award in 2008 from his hands. However, the minister did not respond to that request according to interviewees 8 and 12. The minister of health is therefore seen as a context-setter as he has high power, but his interest in this specific case is low.

Medical Specialist Associations and Medical Specialists

The medical specialist associations represent the professionals that are a member of the association. The following associations are key players in the decision-making process: NVMO, NVVP, NVCO, NVRO, NVvH, VKGN. Every medical specialist is able to give input via his or her association. The medical specialists are the ones that use those tests in practice and therefore have an interest in the decision. They want to deliver the best care for their patients and according to the medical guidelines the gene expression tests may in certain cases be their CDSS [158]. As the medical specialist is not directly involved, but indirectly through the representatives of their associations, they are considered subject in the arena.

NABON

The NABON is a working group and not an association. This means that medical specialists who are also members of a medical specialist association can be members of this multidisciplinary working group. The NABON is a key player in the arena but does overlap with the medical specialist associations.

BVN/NFK and Breast cancer patients

The BVN is a member association of the overarching organization for cancer patients, the NFK. The BVN/NFK is a key player. The breast cancer patients are represented by BVN/NFK and can give input via them. For the patients reimbursement of the test is important first to be able to receive the best possible care as the tests can support in making treatment decisions. Another interest is financially as they otherwise have to pay for the test themselves or they might not be able to afford it. As the patients are not directly involved, but indirectly through the representatives of their association, they are subject in the arena.

ZN and Health Insurers

ZN is the trade association for all health insurers. ZN is a key player in the arena. Within ZN the responsibility is currently delegated to the Add-on Medicines Evaluation Committee (CieBAG) of ZN. CieBAG consists of 4 members from different health insurance companies. In practice, they are currently the ones receiving the draft position reports to give feedback on behalf of ZN. An interview revealed that in the current decision round, the CieBAG responded to the draft report in favor of Mammaprint. In 2009, when the dispute arose, all health insurers did assessments by themselves separately. However, in 2016 this changed and it was decided to do this jointly with all health insurers through ZN. This was changed to ensure that there was no difference in basic package per health insurer and because it was more efficient according to interviewee 18. However, as the ZiN has taken over the assessment on Mammaprint and Oncotype, it seems that the individual health insurers do not interfere in the arena as they leave this to the CieBAG responsible from the umbrella organisation ZN. Individual health insurers are therefore seen as crowd in this arena.

NKI-AVL

The Mammaprint test was invented at the NKI-AVL, but further developed into a spin-off called Agendia [185]. However, the NKI-AVL has done and continues to do much research on the Mammaprint. It is the hospital and research institute specialized in cancer. Both medical specialists as well as researchers work there in the field of breast cancer and other cancer areas. At the moment the NKI-AVL has a very small financial interest, namely 0.04% of the shares of Agendia in compensation for granting the rights to make the Mammaprint [185]. It is not likely that there is a major financial interest here at the NKI-AVL, as at least up till 2018 Agendia has not made any financial profit [185]. The exact interest of NKI-AVL is not explicitly mentioned, however, it seems that the NKI-AVL is very proud of the invention of the Mammaprint. At the time, it was groundbreaking research [75, 76] that led to the product. The NKI-AVL focusses on innovation in cancer care and are passionate about getting this innovation to patients quickly to deliver the best possible care. Many NKI-AVL researchers have been involved in the development and/or further research of the Mammaprint test. The ties with Agendia are very good. It became clear from the interviews that there is still great involvement with Agendia. For example, researchers and other NKI-AVL employees hold additional positions at Agendia. Some interviewees of the NKI-AVL talked in the 'we-form' when discussing how the decisions on reimbursement went, even though those persons did not hold a position at Agendia. It became clear that there is still much passion for the Mammaprint test within the NKI-AVL. It is difficult to say exactly what the interest of the NKI-AVL is. It seems that firstly the NKI-AVL believes in the added value of the Mammaprint and secondly they are very proud and want to strengthen and preserve their status as highly regarded cancer research centre and thirdly there is a chance that on an individual level there might be a (financial) conflict of interest, because some of the employees at the NKI-AVL have side positions at Agendia. Currently, the Mammaprint has already been denied inclusion in basic health insurance twice [91][92]. If this fails again in the current decision round, it means that a U.S. test is included in Dutch basic health insurance, while a Dutch test is not. This is while the Mammaprint is reimbursed in many other European countries [186] and the U.S.

The NKI-AVL is a difficult actor to place. This is because lots of medical specialists and researchers that work at the NKI-AVL are also part of the medical specialist associations and or the breast cancer medi-

cal guideline committee or are even shareholders or also working for Agendia. Despite the fact that the NKI-AVL is not listed by the ZiN, the gatekeeper in the arena, as an actor in the decision-making process, persons working at the NKI-AVL do participate and are sitting round the table with the ZiN albeit said through a different organization. In addition, of the individuals who attended the consultation meeting in 2018 [92], 3 of the 10 external individuals were employed at NKI-AVL. All 3 were also the main or co-author and/or study coordinator of the MINDACT study [160] and one of them was the founder and patent holder of Mammaprint. However, they officially participated in the consultation meeting from the position of NVMO/NABON, NVCO, and Agendia. Therefore the NKI-AVL is considered a key player as the NKI-AVL has more power than a subject while they can sit around the table in the arena.

NFU, NVZ, and Hospitals

The NFU and NVZ are the societies for academic hospitals and hospitals respectively. The NFU and NVZ are context-setter in the arenas. Those organizations thus represent the academic and the non-academic hospitals. Given that the umbrella organizations are merely observers during the consultation meeting with the ZiN in 2018 or indicating that they are leaving this to the medical specialists [92], it is gained that the hospitals themselves also do not have a major stake in the decision for reimbursement. However, interviews did reveal that some hospitals in the Netherlands have an arrangement with Agendia, which still allows them to offer the Mammaprint to their patients without high costs. According to interviewee 12, this arrangement was only offered if a hospital requested it itself. Interviewee 1, a breast cancer surgeon, indicated that he/she finds the Mammaprint nicer to use than the Oncotype and that through this arrangement it is still possible to offer the Mammaprint to patients. Considering the attitude of the umbrella organizations and the fact that there is currently an arrangement, the interest of the hospitals does not seem very high and hospitals are therefore placed between crowd and context-setter in this arena. It should be mentioned that they could easily transfer to subject or even key player in case their interest becomes higher.

House of Representatives

In the House of Representatives, there have been several questions about Mammaprint over the years [175–178]. These questions were addressed to the minister of health. The House of Representatives itself has no power in the matter of reimbursement. They can ask questions to the minister but have no influence themselves as they can not overrule the ZiN and minister.

The following was said about the influence of the House of Representatives by interviewee 19:

...legislative amendments are voted on in the House of Representatives, but this is not about legislative amendments and this is about an assessment of how you have assessed the criterion of SWP. That is what the ZiN does, so the ZiN is an institute that hangs under the ministry of health and is its own kind of implementing body and the ministry of health does get questions about that and the ZiN in turn has to answer to the ministry of health and that accountability is then discussed in the House of Representatives. So also that's how you get a whole, well, It's pretty much a charade, because in the House of Representatives, they're not going to overrule the position of the ZiN.

Based on the fact that through all the years only 4 persons asked official questions in the House of

Representatives, their interest is rated as low. Therefore, they are placed in the crowd.

The power-interest matrix with all actors mapped can be seen in figure 5.5. In table 5.3 the actors, groups they represent and their interests are shown.

Concluding Remarks about the power-interest within the reimbursement arena

From this power interest analysis some interesting findings emerge. It becomes clear that the ZiN is the actor with the most power in this arena. Although the minister has the ultimate authority to intervene, there is a nuanced informality that makes direct interference very rare. Within the parliamentary context, the House of Representatives has limited influence in this decision-making process. Although they have the power to ask questions to the minister of health, their role remains primarily consultative. In the case of Oncotype, the ZiN adjusted its draft position, implying that the organizations that have formal power to influence do also informally have some power. The presence of Agendia or Exact Sciences, although around the decision-making table, casts a veil of ambiguity over their exact influence. It seems that their influence depends on the favor of the medical specialists. If the specialists are not supportive for Mammaprint, the registration holders' input could possibly not be of importance. Medical specialists and patients are subjects at the individual level, but can provide input through associations. The NKI-AVL has a rather unclear role and is a difficult actor to place and a difficult actor to delineate and determine who does act in their interest. This is because lots of medical specialists have dual roles, which implies the NKI-AVL could potentially have a big impact on the decision-making processes. In the case of Mammaprint, healthcare insurers are not leading the assessment procedures. The control was moved to the ZiN early on, and this control continues to rest with the ZiN in future decision-making rounds according to an informal rule. Health insurers join forces and orchestrate their involvement through the CieBAG Committee. This consolidated approach effectively delegates authority to the CieBAG and therefore individual health insurers are not directly involved in the decision-making process. Hospitals entrust the shaping of opinions to their medical specialists. It is unclear whether this would have been different had the logistical and financial complexity been greater.

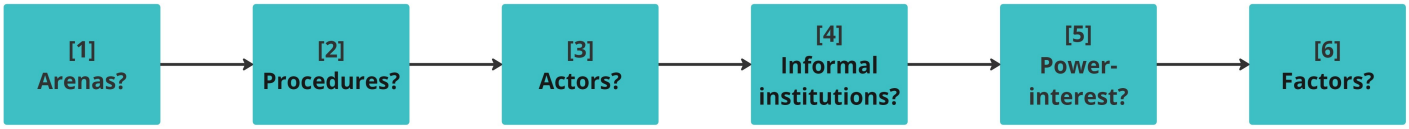

Figure 5.6: Institutional Actor Analysis

5.6. Factors Influencing Inclusion of Mammaprint in Basic Health Insurance

From all the steps in the IAA, factors influencing the adoption and implementation of Mammaprint could be identified. Firstly, by identifying critical decision arenas in chapter 4, it was discovered that the current bottleneck for Mammaprint lies in reimbursement by basic health insurance. When examining the procedures in step 2 of the analysis process, it was revealed that the process stalls during the SWP criterion. Through the identification of actors, as described in step 3, it becomes evident which actors are involved in the decision-making process and warrant further investigation of their underlying values and interests. This investigation takes place in step 4, where significant differences in opinions and motivations among actors in the reimbursement arena come to the forefront. Through this examination, it becomes clear that the challenge lies not solely in the SWP criterion itself, but rather in its interpretation as there exist multiple possible interpretations, and it is the perspective of the ZiN that contribute to the decision that Mammaprint does not comply with the SWP, while other actors in the arena see it differently. From the examination of informal institutions several insights arise. Firstly, Agendia and medical specialists in the reimbursement arena experience ambiguity regarding the requirements for inclusion in basic health insurance and suggest that diagnostic tests should be assessed differently than the way it is currently done by the ZiN. Secondly, medical specialists in the arena and Agendia assert that the burden of proof demanded by the ZiN is too high. Thirdly, the ZiN and medical specialists hold differing perspectives on the required study design to demonstrate the clinical utility of the tests. Lastly, there exists a trade-off between quality of life and survival in evaluating the Mammaprint test, while medical specialists in the arena emphasize quality of life, the ZiN seems to prioritize survival. Those factors are visualized in figure 5.7. The power-interest analysis reveals the ZiN as a critical actor in the arena, with ultimate authority. While formally the minister of health could intervene, the minister would not do so because of informal social norms. However, medical specialist associations and patient advocacy groups can influence the decision-making process. In figure 5.7 all key players are shown, although players are generalised as this will be different associations or organizations depending on the care context.

Factors identified during the case study that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• clinician involvement • availability of funding • availability of data	• CE approval • data sharing standards • legal Factors	• acceptance of AI in healthcare • respondence on market demand	• cost-effectiveness • installation and maintenance support	• clinical validity • clinician's profile • clinician's sense of added value • explainability • equitability • interoperability • ease of use • data security and privacy
			reimbursement by basic health insurance: • unclear requirements for diagnostic tests • high level of burden of proof • disagreement about study design • different priorities in trade-off between quality of life and survival	• inclusion in medical guidelines

Actors identified during the case study that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
• AI and cancer researchers • AMC leaders • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers	• AMC leaders • clinician and/or nurse • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• clinician and/or nurse • patients and care partners • AMC leaders • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators
			reimbursement by basic health insurance in NL: • ZiN • ZN • medical specialist associations • patient advocacy organisations • registration holder • (research institute)	

Figure 5.7: Factors and actors that influence the adoption and implementation of AI-CDSS

6

Factors Influencing Adoption and Implementation of AI-CDSS in Cancer Care

In the previous chapter, it is found what challenges Mammaprint is facing in the reimbursement arena. However, it's important to understand that this arena operates within a broader context. This chapter will address factors beyond the critical arena which potentially influence the processes within it. These contextual elements may provide additional insights into the challenges faced by Mammaprint in the reimbursement arena. This chapter thereby answers subquestion 4: 'What is the broader context that has an influence on those decision arenas?' These contextual factors could be identified through interviews and are partially supported by grey literature. The findings are structured using the phases framework.

6.1. Development

Developing an AI-CDSS within a hospital setting presents notable difficulties according to interviewees (4, 6, 7, 9, 12, 13, 18). Consequently, it is often considered launching a spin-off and involving investors in further product development. This results in commercialization. It does appear AI-CDSS development can also be done in collaboration with a philanthropic organization, allowing open-sourcing according to interviewee 5 [187].

Expertise of Product Development

There is a lack of expertise for comprehensive product development within a hospital as mentioned by interviewee 9:

first of all, as an organization we do not have the expertise to develop products, we do research but

product development is a profession in itself and we simply lack the knowledge for that

This encompasses regulatory compliance (as mentioned by interviewees 7, 9 and 18), scalability (as mentioned by interviewees 7 and 9), and marketing (mentioned by interviewee 9). The stringent regulations governing medical product development necessitate early-stage inclusion of regulatory expertise throughout the developmental process. Attempting to introduce it later could compromise previous steps. Scalability requires adeptness in translating lab-based tests seamlessly into existing diagnostic workflows.

Availability of Funding

It appears from interviews that hospitals themselves lack the financial resources required for comprehensive product development. The certification process and RCTs are considered very expensive by several interviewees (7, 9, 12, 13) and makes the financial burden very high. Although subsidies support scientific and applied research, funding typically ceases after a certain stage as mentioned by interviewee 9:

that funding only goes up to a certain stage of development and when it's really in the product development stage then there are really no subsidies to get and so then you have to go to the market for money

During the product development phase, subsidies seem to be scarce, prompting the need to engage investors from the market. Several interviewees (4, 6, 7) mention the availability of subsidies requires often the involvement of private partners to form public-private partnerships. Those subsidies are issued by Health Holland and will thus only be given in case a private partner also invests. For example interviewee 7 makes use of those subsidies:

I have some projects now and these are public-private partnerships and then you get some kind of grant from the government as a research institution ... and then also a company develops something

This initiative was according to interviewee 4 a policy plan:

It was actually the Minister of Economic affairs who wants to encourage public-private partnerships, it has been a policy plan and from there this call was set up

However, the availability of funding of investors differs depending on the interest rate according to interviewee 9. Low interest rates attract private equity investments, but the anticipation of rising rates may temper this trend. In times of very low interest rates, investors showed interest even at an early stage and actively contributed to the creation of companies as stated by interviewee 9:

so then they are also going to help build those businesses themselves, so I saw that a lot lately. I expect that's now declining because those interest rates are going up .... the money will slowly start to go out of the market a little bit that way. I already notice that it's more difficult to attract investors

According to interviewee 9, willingness to invest varies between diagnostic and pharmaceutical sectors. There is only a limited group that wants to invest in diagnostics. It is mainly therapeutics that

companies want to invest in.

A lot of investors are kind of pulling out of everything but therapeutics and this kind of CDSS which is ultimately diagnostics one way or the other and that's harder to get funded. At least, the traditional investors that I work with who have either never done that or are moving away from that because they can't get the return on investement there, the factor that they need to make their investment pay off that they can't get there ... that's what I do see a lot of in recent years ... it's getting harder and harder to get initiatives funded there ... it's only a very select few that will

However, there are not always investors needed as appeared during an interview with interviewee 5. An AI-CDSS named CUPPA, which stands for Cancer of Unknown Primary Prediction Algorithm, is developed in collaboration with a philanthropic-funded organization [187].

Involvement of Umbrella International Research Bodies

There are high requirements for the studies to demonstrate clinical utility. It emerged that umbrella international research bodies are involved in this, which in turn cooperate with national research bodies. At the European level for this is the EORTC, which is also the leader of the MINDACT trial [160]. At the national level, this is the Breast Cancer Research Group (BOOG), with which the EORTC collaborated for the Dutch patient population. In the U.S. the organisation that conducted the TAILORx was the Eastern Cooperative Oncology Group (ECOG)[174].

6.2. Regulation

European Regulation

The presence of regulatory limitations poses challenges in developing an AI-CDSS within a hospital and within a single institution according to several interviewees (4, 6, 7, 9, 12, 18). To get a CE mark within Europe, devices have to comply to the Medical Device Regulation (MDR) [188] or the In Vitro Device Regulation (IVDR) [189]. These rules have become stricter in recent years [190] as also stated by interviewee 7:

Those rules have only recently changed, that's a few years only the MDR, first you were allowed to do some kind of self-certification, then you were allowed to use it in your own hospital, but now it's very strict with the MDR.

An AI-CDSS must be trained on data from multiple institutions, not just a single hospital. Additionally, a company or institute developing AI-CDSS should appoint a regulatory officer in accordance with the European Medical Device Regulation (MDR). Therefore hospitals' organizational structures are not inherently suited for product development, as they often lack the prerequisites for certification. Interviewee 7 states the following about this:

if you want to apply for an MDR, then you also have to change the organization chart of your company for the fact to comply with the MDR, suppose you want to do that from a hospital, then you have to change the board of directors structure, because the MDR puts a requirement .... to the organiza-

tional chart of the company that develops things, you have to have quality assurance at a certain level, and a hospital has a board of directors for a hospital, but not for something that develops a medical device

Interviewee 18 also indicates that the IVDR makes it difficult for a hospital to develop its own molecular panels.

U.S. Regulation

In the realm of certification, the stories of Mammaprint and Oncotype shed light on a significant contrast in how they obtain approval. Mammaprint's path in the U.S. is notably rigorous, while Oncotype takes a different route. In the U.S., Agendia, the company behind Mammaprint, participated in an FDA pilot project. They secured approval for a complex assay that uses multiple markers with varying importance. This innovative approach exposed a lack of clear rules. On the other hand, the company behind Oncotype chose not to join the pilot project and found an alternative way, as also stated by interviewee 9:

They (Agendia) wanted to be the neatest kid in the class there, but that did lead to them actually being delayed in coming to market with their product, and Oncotype didn't want to do any of that, and they kind of beat Agendia in terms of who came to market first

Interestingly, FDA approval isn't mandatory for laboratory-developed tests and Oncotype is not FDA approved [140]. Mammaprint invested significant time in seeking FDA approval and as a result faced a decline in momentum and market presence. In 2013, Agendia's CEO stated that they initially expected FDA approval to boost their standing in the oncology market and increase sales [191], however, the reality was different while the process of FDA approval consumed a lot of time and resources [191]. This highlights that while certification is crucial, even basic certification can be sufficient for practical adoption.

6.3. Sell/Marketing

Type of Evidence

The approval process for AI algorithms under the MDR [188] or IVDR [189] framework may encompass retrospective assessments [188, 189]. However, the role of conducting RCTs remains somewhat ambiguous. Notably, in the cases of Mammaprint and Oncotype, conducting an RCT has been observed to significantly impact market acceptance, as highlighted by interviewees 9 and 12. Agendia, for instance, initially relied on retrospective evidence to support its product. However, to establish credibility with governmental bodies and insurance providers, the company recognized the imperative of undertaking a comprehensive prospective study [155]. Interviewee 9 and 12 also mention the need for an RCT for market acceptance. Interviwee 9 states the following:

For a lot of people, governments, insurance companies, that prospective study had to be there necessarily before they actually want to believe that it did what it was supposed to do.

Despite the RCT for Mammaprint, the MINDACT study [160], Agendia encountered reimbursement challenges for its product within the Netherlands, though the study did prove influential in facilitating market entry in other countries. Interviewee 12 mentions that it is reimbursed on the basis of this study in many countries:

It is thus ... reimbursed on the basis of clinical characteristics except in the Netherlands, ... (while it is reimbured) in France, in Italy, in Germany, in Belgium, in Czechoslovakia, in America, in Canada

Interviewee 9 also indicates that it did help in other countries:

it did help elsewhere though to sell this better.

Concerning the use of AI-CDSS in (cancer) screening, prospective validation studies are meant to be very important, as it was said multiple times by interviewee 11. Those studies are needed as evidence before being able to decide upon using AI-CDSS within national breast cancer screening programs. This is necessary because it is not clear from retrospective studies how the outcome of an AI-CDSS affects radiologist behavior as stated by interviewee 11:

Precisely with deep learning, you can end up in the situation that an AI indicates something that the radiologist does not see, and that naturally raises questions, so is the radiologist going to act on that, not knowing what the AI is looking at? Because the moment the AI starts seeing a lot of things, many of which do not turn out to be breast cancer on closer examination, yes, then you do something with the balance between the pros and cons of the program, and of course we would rather not have a lot of women being sent to the hospital with a suspicion of breast cancer who turn out not to have breast cancer which is why we also need prospective studies.

We're really not yet at implementation to say that now is the moment in time to implement. We know retrospectively that this AI is actually doing quite well. We should be doing prospective studies now to see what happens, and how it does affect those radiologists.

Marketing Region

In the realm of product marketing, a strategic orientation towards specific markets holds pivotal implications for the sustained success of an AI-CDSS as appeared during the case study interviews of Mammaprint. The intricate interplay between the characteristics of different markets, particularly evident in the European and U.S. contexts, introduces a dynamic that can either advance or impede progress. Remarkably, the U.S. market, as highlighted by interviewees, stands as a significantly expansive market. So indicates interviewee 12:

I would advise anyone starting this again to focus first on the U.S. market.

One of the reasons given by interviewee 12 is that the market is much bigger:

America first of all it is a large homogeneous healthcare market, 320 million people in one system, Europe is 27 countries that all have their own healthcare system.

As also mentioned by interviewee 1:

While the competitor (Oncotype) has very quickly ramped up in the U.S., which of course is a huge market. Which is quantitatively much faster and more.

Although alternative markets are occasionally discussed, there is a prevailing sentiment: none of the interviewees have substantial first-hand experience in this area. In addition, interviewee 7 mentions the enormous challenge of the elusive Chinese market, while acknowledging untapped opportunities in the Australian and New Zealand markets. Mammaprint has achieved notable success in the U.S. market, with an impressive 80% of their sales attributed to this region, as indicated by interviewee 12.

Lobby

The role of lobbying emerges as a potentially influential factor, yet its pathways and magnitude remain veiled in uncertainty. Among the interviewees (3, 10), a recurrent observation is the substantial financial resources at the disposal of Oncotype. This financial disparity raises questions about its implications and effects on market dynamics. So does interviewee 3 mention:

they are just so big that they just have a tremendous amount of money to be able to sponsor. ... for example, their marketing is much broader than Agendia's. Agendia is actually relatively small...Those (Exact Sciences), are just huge and everybody knows it and they just have the resources to, in America it's also just so you can pay doctors to use your test, or a hospital and go talk to a hospital and you make an agreement that they only buy your test and not the competitor's test.

Exact Sciences also tried lobbying in the Netherlands, which is illustrated in an investigative piece by 'Follow the Money' [192]. There are lobbyists, paid by Exact Sciences, for Oncotype in the Netherlands. These lobbyists engaged in casual meetings with medical specialists, advocating the superiority of the Oncotype test over Mammaprint. However, the impact and magnitude of these interactions remain enigmatic.

6.4. Purchase/Reimbursement

Healthcare System

The structure of healthcare systems seems to significantly influences the reimbursement landscape for Mammaprint. A stark contrast exists between the U.S. and Europe in this regard. In the U.S., a single healthcare system governs a vast market, allowing for direct negotiations with insurance providers, fostering a more direct interaction with key stakeholders. Conversely, Europe presents a fragmented market, with each country operating its own distinct system, with variances in product acceptance and reimbursement protocols, and diverse insurance companies as mentioned by interviewee 9:

A major obstacle for Mammaprint in the Netherlands was, and still is, the reimbursement, that reimbursement is I think also one of the reasons that Agendia moved to the U.S. ...., that is one very large market with only one care system and Europe is in this area a very fragmented market, each country

has its own systems, its own way of reimbursement, different insurance companies, so it's a very difficult market to penetrate. .... there the U.S. is a lot easier, so reimbursement has been a big issue.

The negotiation process in Europe often involves government entities, which can lead to slower decision-making according to interviewee 12:

in America you can deal with private health insurance companies, and they can just decide something independently. In Europe, you have to deal mainly with governments, and, that just goes much slower.

Reimbursement rules vary greatly between countries, exemplifying the intricate relationship between healthcare systems and financial considerations. Notably, the Netherlands upholds a system rooted in solidarity and equal access to healthcare services and seems to foster a more cautious approach to costly treatments and diagnostics. This contrasts with the U.S., where a more commercialized healthcare system seems to contribute to a higher level of acceptance for expensive medical services. The distinction between public and private healthcare systems further underscores the disparities in reimbursement strategies.

Lack of Clear Reimbursement Route for Basic Health Insurance

The process of reimbursement of AI-CDSS within the Dutch healthcare system lacks a straightforward route as interviews and grey literature revealed. This includes AI-CDSS making use of diagnostic procedures within cancer care, as well as the reimbursement of AI-CDSS employing image analysis methods.

Molecular Diagnostics

While certain AI-CDSS for diagnostics, such as Oncotype [93] are covered under the basic insurance package under certain conditions, the landscape for molecular diagnostics remains intricate. However, recently the field of molecular diagnostics has witnessed significant developments in delineating the possibilities for reimbursement. In 2021, the 'implementation process molecular diagnostics' [193] has been launched at the request of the Ministry of Health, Welfare and Sport [193]. Within this implementation process, the ZiN collaborates with the field to enhance the quality, accessibility, and affordability of molecular diagnostics. The outcomes are intended to contribute to tailored care for patients with metastatic cancer [193]. Under the auspices of ZN, the CieBAG committee [194, 195], along with its Molecular Diagnostics (MDx) working group [196], collaborates with healthcare professionals and patient associations to address the reimbursement of molecular diagnostics within this implementation process. Within molecular diagnostics there is a grouping of several different biomarkers. There is a clear reimbursement route for the group of biomarkers directly linked to a drug. This is for example for the EGFR mutation in long cancer patients which is coupled to a drug named Osimertinib [197]. Now, they are also looking within the 'implementation process molecular diagnostics' at making reimbursement possible for biomarkers that are not directly linked to a drug, but are linked to treatment decision-making. Interviewee 5 says the following about this:

The first group is biomarkers linked to a drug and the second group is biomarkers not necessarily linked to a drug, but which do support a treatment decision...the first group goes the classic way of the blue boxes as we call them, but then you have that second ... those are biomarkers that are a

little trickier, because the ZiN and the health insurers said that's not reimbursed, because that's not proven. So those practitioners said 'we can't work like this if we don't have these biomarkers', ... the consensus now is that the second group is also on that list and thus is also reimbursed care, so now we have two categories of reimbursed care, those blue boxes (group 1) linked to a drug, and these (group 2), we also call it prognostic factors, for example if you have a very poor prognosis you will start heavy therapy sooner, if you have a very good prognosis you might be able to delay for a very long time.

Mammaprint is also a form of molecular diagnostics that has an effect on treatment decision-making. If Mammaprint had come to market now and not gone immediately to ZiN for assessment, it would most likely follow the route to reimbursement that the CieBAG is currently devising according to interviewee 5.

The CUPPA algorithm [187, 198, 199], an AI-CDSS employing molecular diagnostics, stands out as a success story. Widely adopted by national molecular tumor boards, CUPPA facilitates the identification of potential treatment options for cancers with unknown primary sites. It used Whole Genome Sequencing (WGS) data as input, despite the substantial expense associated with WGS, it secured national reimbursement for patients with an unknown primary site within just 1.5 years according to interviewee 5:

That CUPPA has been reimbursed nationwide for 1.5 years and we started that 3 years ago.

A special type of billing code, Other Care Product code (OZP code) was introduced to ensure the possibility for reimbursement of CUPPA [200]. It's important to note that CUPPA possesses distinctive attributes: it was developed in collaboration with a philanthropically-funded organization and is therefore non-commercial and open-sourced on Github [199], it is used exclusively by medical specialists within national molecular tumor boards and viable alternatives for treatment decision-making were previously lacking according to interviewee 5. Furthermore, it remains unclear whether any charges are incurred for the use of the CUPPA algorithm, or that reimbursement payments exclusively cover the costs of WGS and the efforts of the medical specialists within a molecular tumor board.

Image Analysis Based AI-CDSS in hospital

The path towards reimbursing image analysis methods is considerably less defined than that of molecular diagnostics. Interviewee 6, who also develops image-based AI-CDSS, mentions the current ambiguity surrounding the feasibility of reimbursing such systems. Efforts are underway to establish 'guideline AI' [201], aiming to clarify aspects like reimbursement by insurers for the use of AI in healthcare [201]. This is still at an early stage and further developments in this area are needed.

The lack of a clear reimbursement route is also illustrated for another AI-CDSS based on image analysis methods in cancer care, namely Transpara [202]. Transpara is a spin-off from the Radboud UMC and has achieved CE and FDA approval [203].This AI-CDSS helps radiologists in the hospital with breast cancer screening and is used within several hospitals within the Netherlands [203]. Hospitals that implement Transpara bear the costs for its use [203]. There is an ongoing tension regarding the evidentiary requirements for reimbursement. In a research report on AI applications in healthcare commissioned by the ZiN and conducted by research firm Berenschot, the following is stated:

Financing of Transpara by health insurers has not yet been addressed at the current stage. Hospitals

currently pay for the application themselves. There is partial room in the budgets for such innovations, but the tension remains: when has a clear point been reached where sufficient evidence has has been gathered to qualify for health insurer funding, according to several actors. [203]

This issue of when there is enough evidence for reimbursement was also already highlighted in the previously mentioned 'AI Guideline' [201].

Image Analysis Based AI-CDSS in screening

Currently there are no AI-CDSS used in screening in the Netherlands, but developments are being done. In breast cancer screening, especially image analysis methods appear promising according to interviewee 11. The reimbursement for screening follows a completely different path. Costs are reimbursed from a budget for screening programs from the Ministry of Health and thus have nothing to do with health insurance. The Committee for Population Screening within the Health Council advises the Minister of Health about this. The pathway thus appears straightforward, but it has not occurred thus far.

Existence of Alternative Reimbursement Routes

The existence of alternative reimbursement routes besides reimbursement through basic health insurance could positively influence the adoption of AI-CDSS as learned from the case study on Mammaprint. These alternative reimbursement routes explain how it is possible that Mammaprint became well established in practice, while it is not in the basic health insurance package. First, during the RASTER study and the MINDACT, the Mammaprint was reimbursed in study context [160, 168]. Second, according to many interviewees (8, 10, 12, 17, 18) the Mammaprint test was reimbursed by many health insurers out of leniency. Although when again a negative assessment of the ZiN came out in 2018, many health insurers stopped this goodwill arrangement. Reimbursement out of leniency is also found by a more recent AI-CDSS, namely SkinVision, which is reimbursed out of leniency by one health insurer in the Netherlands [204][1]. Third, Mammaprint is currently in the supplementary insurance of a couple of health insurers according to interviewees 12 and 17. Fourth, even without supplemental insurance, you can still use a Mammaprint test at certain hospitals at no additional cost as it became clear from interviews that Agendia has arrangements with hospitals so that they can continue to use and offer the tests to patients without incurring high costs. How many hospitals this is about is unknown. Agendia itself confirmed the existence of this arrangement. The reason the arrangement with Agendia exists seems to be because of the fact that medical specialists were used to using the test and were eager to continue offering it. It is suggested about Agendia's rationale that they could also retain market share this way, however, this has not been mentioned by or asked of Agendia if this is indeed true. This does seem plausible, as it was also mentioned that many hospitals switched completely to Oncotype after the 2021 decision, as that test was reimbursed through basic insurance from then on. According to one of the interviews working at one of the hospitals that has an arrangement with Agendia, they would probably also have to switch if Mammaprint will not be reimbursed soon:

Because with us that is still financially reimbursed in a certain way. But the Mammaprint is not included in the package of reimbursed care from the health insurance companies and the Oncotype is.

[1] Skinvision is a home-based image analysis AI-CDSS designed to assess whether skin abnormalities warrant further examination by a medical specialist for potential cancer.

So ... if that doesn't change, the Mammaprint is going to disappear, at least for the Netherlands, then it's going to be Oncotype for us as well.

6.5. Use

Utilization of AI

At first glance, the utilization of AI may appear to be negligible, partly due to a lack of awareness among medical professionals regarding its presence. For instance, interviewees 4 and 6 indicate that Mammaprint is not associated with AI in the medical world. Medical specialists seem to be unaware that AI is employed in the classification of high and low-risk groups in tests such as Mammaprint and Oncotype. Interestingly, when asked what qualifies as AI, all interviewees (3, 4, 6) pointed to image analysis applications, often using advanced deep learning techniques. These applications are currently gaining attention in fields like radiology and pathology, marking image analysis as an emerging and innovative area according to interviewees 6, 7, 11, 12 and 13. For example, interviewee 3 mentions:

what we see as AI, for example, is that which is used a lot in radiology and pathology, so the image and computer-based learning

Interviewee 6 mentions:

I think we are now referring mostly to AI in the application for medical care to image analysis, so that's then in radiology and in pathology

the new revolution coming up after genetic classification or sequencing is digital image analysis using deep learning methods

It seems from interviewees 6, 7, 11, 12 and 13 that new areas where AI-CDSS are upcoming are thus now in image analysis methods in radiology and pathology.

Paradoxically, while Mammaprint is not associated with AI, medical practitioners viewed Mammaprint as black box at the beginning because of the use of gene expression profiles according to interviewees 7, 10, 12 and 19. This perspective can be attributed, in part, to the early stages of Mammaprint's development coinciding with the start of understanding the human genome. Back in the early 2000s, the first rough map of the human genome was unveiled. During this period, the application of genomics tests was groundbreaking, and the focus was primarily on the use of gene expression profiles rather than delving into the algorithms. For example, interviewee 10 says the following about this:

The Mammaprint was actually the first kind of DNA profile in the Netherlands at that time, which was really considered by a doctor as a kind of black box.

When asked about their perspective on AI's role in breast cancer care, interviewees (1, 2, 3, 5) generally exhibit enthusiasm. Interviewee 1 for example mentions:

I'm open to it, and I'm at the end of my career, but I think it's definitely worth exploding. I see the doctor's role more and more, and I'm not the only one, as a coach, as a translator of all those kinds of things that are becoming publicly available.

Interviewee 11 mentions the following about radiologists' attitudes toward the use of AI:

In the beginning, certainly at congresses there was very much such a debate of are we being replaced by the AI. I don't really see that anymore. Now it's much more about the question: 'Hey AI is very promising, how are we going to integrate it as good as possible into the process and how do we ensure that the radiologist and the AI work together as good as possible so that we can improve the process?' And that's kind of the quest where we are right now.

It's about, how can we best achieve that interaction? And if that means freeing up some radiologist time with that. Yes, that's absolutely fine, of course. There is a need everywhere, right? There is capacity problem everywhere. So it's only nice if the AI can make sure that the radiologist can focus on some other things as well.

Their enthusiasm, however, comes with a condition: they are not averse to "black box" algorithms, as long as their effectiveness is validated and proven. The AI-CDSS have to do what they claim they can do. However, this seems to be something more general and not specific to the use of AI.

Output type

It emerges that there is no clear preference among medical professionals for the type of output given by Mammaprint and Oncotype. In fact, these tests differ on that. The Mammaprint gives a scale with a cut-off point for high and low risk, so the score is somewhere on a scale. The Oncotype gives a scale with a low, medium and high risk group, so there is a score attached to this as well. Interviewee 3 indicates a high and low risk group is much clearer, however, this interviewee also indicates to be biased toward Mammaprint. Interviewees 1 and 2 indicate that they are used to the use and output of the Mammaprint test, but do not have a strong preference in the type of output. With both tests, you have some leeway and you are not bound by the result, but can still discuss that with the patient. Interviewee 1 says the following about this:

The cut-off point for both tests is not such a thing, as far as I'm concerned you'll be fine with both tests. It's binary on the one hand, red or green, on the other hand you still get the percentage where you're free to still have in joint decision making about 'okay it's green, but it's very close to that midline, maybe if you want that chemotherapy can still come into the picture', so with neither test you're necessarily stuck on a black or white decision making

Interviewee 4 and 6 also mention that a mere 'yes' or 'no' as output for an AI-CDSS is considered unpleasant. Interviewee 2 indicates to not prefer an output without underlying information:

then you're actually missing the underlying information, because then there's actually another entity

that is therefore deciding for you how information is interpreted or where you get the information from, so I would be a little more cautious about that.

From this it does thus seem to be important to include a score on a scale in addition to the risk group, as this allows the medical specialist to interpret it.

6.6. Concluding Remark

During the interviews with various stakeholders the broader context of the reimbursement arena of AI-CDSS in cancer care, and specifically Mammaprint, came to the light, which are shown in figure 6.1. It became evident that it is difficult to develop AI-CDSS within a hospital as funding and expertise of product development are missing. There are also regulatory requirements which pose impediments. In terms of funding, there is often dependence on investors and the government encourages public-private partnerships. Yet it appears that philanthropic-funded organizations can be an alternative during the product development phase, in which case the product does not need to be commercialized. Certification alone appears insufficient for marketing and reimbursement of the product, additional evidence is essential. This can be in the form of RCTs, for which the involvement of umbrella international research bodies is important. It became evident that the ease of adoption is highly contingent on the region or country, particularly evident in the sell/marketing and purchase phases, where differences between Europe with a primarily public and the U.S. with a primarily private healthcare market emerged in particular. It seems that when direct negotiations with healthcare insurers are feasible, the adoption of AI-CDSS is smoother compared to healthcare systems where government involvement is prevalent. There might be a role of lobbying, but this role remained vague. Furthermore, reimbursement pathways for AI-CDSS in the Netherlands vary significantly based on the underlying technique, namely molecular diagnostics or image analysis and based on the use in hospital or in screening. The reimbursement status of image analysis AI-CDSS remains unclear, with ongoing efforts to assess the feasibility of reimbursement. On the other hand, AI-CDSS utilizing molecular diagnostics, have been subject to reimbursement regulation development to clarify and improve reimbursement pathways via the 'implementation process molecular diagnostics' [195]. Besides being included in basic health insurance, alternative reimbursement routes do exist in the Netherlands, although often temporary in nature. The interviewees indicate Mammaprint is not seen as a form of AI within the medical field, but what the medical field does see as AI are image analysis techniques. These AI-CDSS using image analysis seem to be on the rise in radiology and pathology. According to interviewed medical specialists AI-CDSS should not provide a definitive, binary response; instead, it should be designed to allow a medical specialist to interpret the results.

Factors identified during the case study that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase	Use
• clinician involvement • availability of funding • availability of data	• CE approval • data sharing standards • legal Factors	• acceptance of AI in healthcare • respondence on market demand	• cost-effectiveness • installation and maintenance support	• clinical validity • clinician's profile • clinician's sense of added value • explainability • equitability • interoperability • ease of use • data security and privacy
			reimbursement by basic health insurance: • unclear requirements for diagnostic tests • high level of burden of proof • disagreement about study design • different priorities in trade-off between quality of life and survival	• inclusion in medical guidelines
• expertise of product development • availability of funding • involvement of umbrella international research bodies	• European Regulation • U.S. Regulation	• type of evidence • marketing region • lobby	• healthcare system • lack of clear reimbursement route for basic health insurance • existence of alternative reimbursement routes	• utilization of AI • output type

Actors identified during the case study that influence the implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase	Use
• AI and cancer researchers • AMC leaders • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• system developers	• AMC leaders • clinician and/or nurse • system developers • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators	• clinician and/or nurse • patients and care partners • AMC leaders • developers, regulators, political and financial stakeholders, ethicists, operational support, medical users, and maintenance operators
			reimbursement by basic health insurance in NL: • ZiN • ZN • medical specialist associations • patient advocacy organisations • registration holder • research institute	

Figure 6.1: Factors and actors that influence the adoption and implementation of AI-CDSS

7

Discussion

In this chapter first the findings from the literature and case study will be compared. Second, the factors identified during the case study on Mammaprint will be discussed. Third, the interpretations according to institutional theory and generalizability will be discussed. Fourth, there is a reflection on the methods used.

7.1. Relation Literature and Case Study

The literature (see chapter 3) primarily focused on AI-CDSS in a general context, whereas the case study specifically focused on AI-CDSS within breast cancer care (see chapter 4, 5 and 6), albeit with some insights extending to other forms of cancer. This juxtaposition offers an intriguing opportunity to discern whether factors identified in the broader realm of healthcare also apply to (breast) cancer care. Moreover, it prompts an exploration into any novel factors that may have emerged during the case study, absent from the initial literature review.

Development

During the literature review it was found that medical specialist involvement during the development process is an important factor [95–97, 101, 110, 111, 113]. It is very likely that this helped for Mammaprint as development took place within the NKI-AVL which also employs lots of medical professionals who are part of national medical specialist associations which helped diffusion within a network of medical specialists. The availability of funding Watson *et al.* [97] was found to be an important factor during the literature review and then especially the costs for clinical trials [115]. This is something which was also mentioned often during the case study interviews as an important reason why externalisation of product development is needed (see section 6.1). For the clinical trials, lots of data is needed [97, 102, 116] and should be acquired in the right way [97, 100, 101, 107, 115, 117]. This was implicitly validated by the fact that there was intensive involvement of international research bodies, who also managed the data acquisition process. In the context of the development phase, many factors outlined

in the literature thus resurface in the case study. A notable revelation is the imperative for externalizing product development (see section 6.1). Interviews indicate that this need stems not only from financial considerations, but also from the lack of specialized product development expertise within a hospital setting.

Regulation

In the regulation phase, uncertainties abound. While the literature suggests ambiguity in the standards required for CE approval [100, 101, 119], this did not arise in the case study interviews, potentially due to the absence of regulatory experts among the interviewees. Moreover, recent changes in Medical Device Regulation (MDR) and In Vitro Device Regulation (IVDR) in 2021 may have alleviated some of the uncertainties [188]. Nevertheless, European regulations pose challenges for in-hospital AI-CDSS development, necessitating collaboration across multiple institutions (see section 6.1). The case study also highlighted variations in FDA approval processes between Mammaprint and Oncotype, with uncertainty surrounding whether FDA approval was necessary (see section 6.1). Mammaprint underwent a rigorous pilot project for FDA approval, while Oncotype had a comparatively smoother path. Although the case study primarily focused on the purchase/reimbursement phase, discussions regarding data sharing standards and legal considerations were limited, suggesting they currently have minimal impact on Mammaprint, though this may have been different in the past.

Sell/Marketing

Regarding the slow acceptance of AI in healthcare, as highlighted in the literature [135–137], confirming or refuting this factor based solely on the case study findings is not possible. What does emerge is the ongoing proliferation of AI-driven solutions in healthcare, with medical specialists generally displaying receptiveness. However, it should be noted that factors influencing sector-wide acceptance may be varied, and this research specifically focuses on the adoption of Mammaprint, where AI did not play a role in reimbursement and in acceptance as most medical specialist were not aware of its use (see section 6.5). Similarly, concerning the respondence on market demand it is also very hard to either confirm nor deny the effect of this factor in the case study. This is complex as payers and users are distinct entities, making it hard to definitively attribute demand determination. What can be said is that Mammaprint is considered to provide added value in the Neterlands only for a small subset of breast cancer patients [158]. During the case study, novel factors emerged regarding the marketing of AI-CDSS (see section 6.3). For instance, in the case of Mammaprint, mere CE or FDA approval proved insufficient; additional evidence, in the form of an RCT, was required to instill trust, which was also mentioned for AI-CDSS applied in screening. Moreover, the geographical scope of marketing played a pivotal role. While the U.S. boasts a big unified market, the European landscape is markedly more fragmented, organized on a country-by-country basis. The role of lobbying in marketing, while mentioned in the case study, remained somewhat vague yet could potentially exert influence on marketing.

Purchase/Reimbursement

While the literature emphasizes cost-effectiveness as a crucial factor [101, 104, 129], it wasn't a primary consideration for Mammaprint's rejection in the Netherlands' reimbursement process. This has two sides, on the one hand Mammaprint is probably cost-effective [92]. On the other hand, the proce-

dure for reimbursement in basic health insurance only dives in the cost-effectiveness in the last phase, while Mammaprint is already rejected in an earlier phase because of clinical utility reasons [91, 92]. Installation and maintenance support was indicated by literature to be insufficient [97, 114]. During the case study not much is said about this, but it seemed that for Mammaprint the maintenance support is sufficient as hospitals associations indicated to be fully content with organizational and logistical concerns (see section 5.5). Noteworthy new factors emerged in the case study. There's notable uncertainty about the reimbursement pathway for basic health insurance for molecular diagnostic AI-CDSS like Mammaprint and also for in-hospital image analysis-based AI-CDSS (see section 6.4). Additionally, substantial disagreement was observed between medical specialists and the ZiN in the Netherlands regarding reimbursement. Medical specialists hold different perspectives on when clinical utility is adequately demonstrated and the risks they're willing to accept in the trade-off between quality of life and survival (see section 5.4). Furthermore, the healthcare system itself also seems to play a role. In the Netherlands, government agencies have the final say in decisions for reimbursement within basic health insurance as the system is rooted in solidarity. In contrast, the American system is more commercialized, allowing for direct negotiations with insurance providers (see section 6.4). In addition, it became clear that there are alternative pathways for (temporary) reimbursement in the Netherlands (see section 6.4).

Use

In the literature study most factors influencing the adoption and implementation of AI-CDSS appeared in the use phase (see section 3.5). This could be because of a bias in the search terms. It could however also be that this is actually the most studied phase. During the case study, the use phase was not the primary focus; hence, fewer factors were mapped in this phase. What stood out that medical specialists are open for the use of AI [108, 112] Gunasekeran *et al.* [105] Alsobhi *et al.* [104], something that could already be found in literature and could be validated during the case study (see section 6.5). Although in the case of diagnostic tests, medical specialists are not aware of the use of AI. Another thing that stood out is that medical specialist's have a sense of added value when it comes to the decision-making process for treatments [96, 101, 105, 107]. This was also found during the case study where medical specialists indicated that in the output of an AI-CDSS there should be leeway for the own interpretation of the medical specialist (see section 6.5). This also calls for systems that provide an explanation on which their recommendations are based.

7.2. Discussion of Factors identified in Reimbursement Arena

The inclusion of Mammaprint in basic health insurance encounters a bottleneck during the SWP criterion assessment, which evaluates its clinical utility. The SWP criterion is a formal institution mandated by law in the 'Besluit Zorgverzekering' in article 2.1, second paragraph [167]. However, institutions operate in a layered manner, and in this case, it is the ZiN that establishes the assessment framework for the SWP criterion [89], which holds independent legal significance. Additionally, the ZiN clarifies the nature, content, and scope of the insured services. The precise implementation of the SWP criterion is thus not mandated by law, but delegated to the sector. The SWP assessment framework [89, 170] provides guidance, but it also allows for considerable discretion, leading to multiple interpretations. Therefore, apart from creating the SWP assessment framework, the ZiN also possesses the authority to interpret it on a case-by-case basis. This grants the ZiN significant room to establish its own rules and bestows substantial decision-making power upon the ZiN in the arena.

7.2.1. No clear requirements for inclusion of diagnostic tests

The ZiN has lots of experience in assessing medications, but little experience in assessing diagnostic tests

Where the ZiN has extensive experience in evaluating medications, the same cannot be said for medical tests. Extramural medicines, those available at pharmacies with a doctor's prescription, are always assessed by the ZiN, which then advises the minister [162, 163]. Additionally, the ZiN evaluates medications leading to very high costs, falling under special consideration, in the 'lock of expensive drugs' where they also advise the minister [164]. However, in the realm of medical tests, the ZiN lacks the same level of experience. This probably because of two reasons. Firstly, assessments of intramural care, like diagnostic tests, that are not in the lock of expensive drugs are normally the responsibility of health insurers [163] and only come to the ZiN in exceptional cases. Secondly, the frequency of requests for medical tests is likely relatively low. The Mammaprint test was the first medical test assessed by the ZiN (see section 5.4.1). In the first decision round in 2010, the ZiN's draft position report [91] states that it assesses based on evidence-based medicine principles, but acknowledges limited experience with medical tests, as most experience has been gained with therapeutic interventions [91]. During this assessment, the ZiN also established a framework for evaluating medical tests, which is used within the SWP [91]. In 2011, further clarification on the SWP assessment framework is provided in the event of a medical test evaluation [170]. This is not a new way of assessment, but an explanation of how the SWP assessment framework can be applied to medical tests. In the SWP for medical tests [170], it becomes clear that, among other things, medical diagnostic tests in screening were examined as for screening the debate about when a test is beneficial was settled decades ago [170]. In this context, the criteria of Wilson *et al.* [205] from 1968 and the additional screening criteria set by the WHO in 2008 [206], which are also used by the RIVM for responsible population screening [207], are considered [170]. The starting point here is that tests are not applied when they do not have a positive effect on health or do more harm than good compared to the alternative of not testing or using another test, even if the test results themselves are valid [170]. However, there are two points that are not black and white. For instance, it is debatable when there can be talk of sufficient evidence, and in the case of Mammaprint, there is also a trade-off between potential positive and negative effects on health. Furthermore, it is notable that the medical specialists did include Mammaprint in the medical guidelines [156–158] while also using evidence-based medicine principles. This also shows that there

are different interpretations and that there is a tension in this regard.

Other actors in the reimbursement arena assert that the evaluation method is unsuitable for medical diagnostic tests, such as Mammaprint (see section 5.4). According to them, the ZiN assesses medical diagnostic tests as if they were medications, while they should be approached in a different manner. It appears logical that the ZiN would apply the familiar method that worked well for medications. Additionally, according to interviewees (see section 5.4) there is uncertainty surrounding the evaluation process, which undergoes changes throughout the decision rounds. This can be explained by the newness of the field, with the evaluation process still evolving. Furthermore, the ZiN may exercise additional caution in permitting these types of new diagnostic tests, as once they are approved, many more may follow, potentially leading to an unmanageable situation. In this sense, being an innovation and the first of its kind may work against Mammaprint.

Little changes in assessment procedures of diagnostic tests throughout the years, but recent improvements

Moreover, there appears to be a low sense of urgency at the ZiN. While a module on medical tests was published in 2011 [170], no further updates have been made over the years. Yet recently, since the past two years, this does seem to have turned around. On behalf of the Ministry of Health, the ZiN launched an implementation project for molecular diagnostics [193] with the aim of improving the quality, accessibility and affordability of molecular diagnostics with the field (see also section 6.4). The results should contribute to appropriate care for patients with metastatic cancer [193]. According to one interviewee (interviewee 19), a new module for the SWP dealing with molecular diagnostics will also follow from this. So there does seem to be a search now for an appropriate way of assessing molecular diagnostics in cancer care. Several actors are involved in this molecular diagnostics implementation process [193], many of whom are also involved in the reimbursement arena of Mammaprint. The committee CieBAG [194] from ZN is since recently responsible for molecular diagnostics [195]. From interview input it seems that CieBAG takes the lead in cooperation with ZiN on this, in close cooperation with medical professional organizations [196], among others. According to interviewees (5, 18 and 19) it seems to be promising and to actually improve reimbursement of medical molecular diagnostic tests considerably. This is 20 years after the invention of Mammaprint. So in terms of innovation, the legislation is years behind. It also remains to be seen whether this will change anything for Mammaprint as Mammaprint is currently in an assessment procedure [94].

7.2.2. (High) Level of burden of proof needed

There is a political, a commercial and a medical point of view about the required level

Regarding the level of burden of proof considered necessary for Mammaprint, there are large differences between the ZiN, the medical specialists in the reimbursement arena and the commercial actor. Where this is notable, there are possible explanations. For example, the ZiN wants the best evidence before allowing something into the basic package. The ZiN works on the principle that everyone helps pay for healthcare and therefore it must be proven useful before it is reimbursed. As for the medical field, they would like to be able to use it quickly in practice and as long as reimbursement is not properly regulated, this is inhibited. It is also indicated that 10-year RCTs are very difficult studies and that external factors are impossible to keep equal during such a long study duration (see section 5.4). The ZiN sees this different, because they see the 10-year RCT studies as less reliable because over the

years the use of chemotherapies changes. From the commercial point of view, the time frame at which evidence is needed exceeds the business case time frame. Furthermore, there are very high costs associated with RCTs that run for 10 years. All these different interests, objectives and experience leads to different point of views about the level of proof needed.

Insights on future requirements regarding the level of burden of proof

It further emerges from interviewee 5 that within the current trajectory 'implementation process molecular diagnostics' [193] (see also section 6.4), there is a high likelihood that phase 3 studies [1] will not be necessary for the inclusion of biomarkers in reimbursed diagnostics. In theory reimbursement was already possible as in the SWP it is also stated that phase 3 studies are not always required for inclusion in the insured package [89, 170]. However for in practice, a consensus now seems to be emerging within this trajectory 'implementation process molecular diagnostics' about when this is and is not necessary between the ZiN, medical specialists and health insurers. On the other hand, it becomes clear from interviewee 11 that the use of AI-CDSS in screening does require prospective studies before AI-CDSS will be implemented in national screening programs.

7.2.3. Disagreement on study design between medical professionals and ZiN

There is a big disagreement about the study design needed to prove clinical utility between the ZiN and the medical specialists within the reimbursement arena of Mammaprint. This is quite striking that there is such a big difference. The medical specialists consider the results of the MINDACT study [160] to be good enough to prove clinical utility and use the test in practice, while the ZiN argues that the study design in itself would not be able to answer this question. The study design gets stuck on several points, for example, the ZiN divides the study population into subgroups and, according to the ZiN, all patients should have received randomized treatments. As for the subgroups, some ambiguity remains. Indeed, medical specialists do indicate in medical guidelines that the Mammaprint test is only suitable for a specific group of patients [158]. Thus, they also ultimately select a subgroup from the entire study population for whom it would be useful. However, according to them, the entire study population should be considered as one group when it comes to measuring outcome measures. Furthermore, the medical specialists in the arena indicate that not all patients can be randomized because, according to medics, it is not ethically acceptable. There seems to be a knowledge gap here that works both ways. For example, the policy making field does not seem to have a good idea of how studies work in practice, so they may overlook what is and is not ethically acceptable. In addition, the medical field seems to overlook how the policy making field views clinical utility. Furthermore, the ZiN deems it necessary to have more quality of life data in order to weigh between the trade-off of quality of life and survival. However, medical specialists consider it evident that Mammaprint improves the quality of life for patients and find it strange that evidence is being requested for this. However, interviewee 13 pointed out that attempts were made within the field to gather more quality of life data for MINDACT (see section 5.4.3). However, at that time, it was not yet common practice to measure this for such studies, resulting in minimal data on quality of life being collected.

[1] Phase 3 refers to a large prospective and often randomized study that compares the safety and effectiveness of the new treatment against the current standard treatment, often using randomization.

7.2.4. Priorities in trade-off between quality of life and survival differ

There exists a trade-off between quality of life and survival in evaluating the Mammaprint test. There is a potentially large group that by using the Mammaprint does not have to contend with unpleasant side effects of chemotherapy. However, there is also a potentially very small group that will forgo chemotherapy by using the Mammaprint test, while the patient will still develop metastases and may even die from them. The question then is do you prioritize the improvement in quality of life of a large patient group over the potential death of a patient because of the Mammaprint? And then how heavily does quality of life versus survival weigh? This is where ethical choices must be made.

Survival is often considered as a primary objective in the medical world, as it serves as a tangible metric to demonstrate predictive capacity. If a test doesn't show a survival benefit, it may be deemed medically less useful. Additionally quality of life is more difficult to quantify and is often seen as a secondary objective. What makes it additionally challenging is the very limited availability of quality of life data for Mammaprint.

It became clear that medical specialists and the ZiN express different preferences in this regard (see section 5.4.4). These different preferences could be preceded by different ethical values, which are neither right nor wrong, but reflect ethical preferences. For example, the underlying ethical principle of choosing the greatest good for all, which is thus done at the moment when the improvement of quality of life for a large group is seen as more important than the potential death of a very small group, is a utilitarian principle. Whereas choosing that the risk of death should be very low is patient-centered and based on a deontological principle where death is seen as something to be avoided at all costs if necessary at the expense of a large patient group.

7.2.5. Potential Strategic Behavior

In the disagreements between the medical and political fields, there may also be an element of strategic behavior at play. Between the key players there are different objectives at play concerning the level of burden of proof and study design needed and the trade-off between quality of life and survival. This while there is a lack of clear criteria, which makes the discussion quite blurry, leaning towards a political conflict rather than a straightforward evidence-based one. It's possible that the arguments presented by key players may not be as robust as they are portrayed. There may be underlying motives and strategic behaviors at play, where stakeholders mobilize seemingly safe arguments to support their actual, potentially politically motivated, positions. This tactic of invoking higher values to justify actions is a common strategic behavior [208], often obscuring the true, more pragmatic motives. For example, in this case, such a higher value could be the survival or quality of life and using this to hide potential more pragmatic, possibly financial, motives. It is however not possible to proof strategic behavior [208].

7.3. Interpretations according to Institutional Theory

In the case of Mammaprint adoption and implementation in the Netherlands gets stuck during reimbursement by basic health insurance (see chapter 4). The SWP criterion appears as the point in the assessment for reimbursement where difficulties arise (see chapter 5). The SWP assessment framework [89, 170] is about proving clinical utility, however, this appears to be open to interpretation as there are no generally accepted rules and norms to decide upon clinical utility. Here the medical specialists participating in the reimbursement arena of Mammaprint disagree with the interpretation of the policy analysts in the arena. The medical specialists in the arena indicate that requirements for inclusion of diagnostic tests, like Mammaprint, are not clear. Also, there is no consensus about the level of burden of proof and the study design needed to proof clinical utility and there is no consensus about the trade-off between survival and quality of life between the medical specialists and policy analysts within the arena. When looking to the reimbursement pathways of other AI-CDSS (see section 6.4), a lack of a clear reimbursement route for basic health insurance also appears for other AI-CDSS. For AI-CDSS making use of molecular diagnostics improvements to facilitate the implementation, are currently made within the trajectory that is called 'implementation process molecular diagnostics' [193] (see section 6.4). For AI-CDSS making use of image analysis methods, there is still huge uncertainty about the route for reimbursement in basic health insurance. However, a start has been made to draw up 'guideline AI' [201] (see section 6.4) to clarify the placement of AI in healthcare including clarifying when AI is eligible for reimbursement. From all this it can be concluded that for AI-CDSS deployed in the hospital, there are no clear rules and norms that are generally accepted according to which policy decisions about the inclusion of AI-CDSS in basic health insurance should be conducted. This can be called an 'institutional void' as defined by Hajer [209], who describes it as the following:

there are no clear rules and norms according to which politics is to be conducted and policy measures are to be agreed upon. To be more precise, there are no generally accepted rules and norms according to which policy making and politics is to be conducted. Hajer [209]

7.4. Generalizability of Institutional Findings

When delving into the literature, institutional voids emerge as a well-recognized phenomenon in the field of economics and policy analysis. Institutional voids can manifest in various forms, including the absence or unreliability of market information, an uncertain regulatory environment, inefficient legal systems, and cumbersome bureaucratic processes [210–212]. They can also extend into the domain of infrastructural support and impact the prevailing cultural context [213–215].

Institutional voids have the potential to influence various aspects of the business environment, with both advantageous and disadvantageous consequences. The absence of market-supporting institutions, for instance, can significantly increase transaction costs for entering businesses [216]. In the presence of these voids, the business environment may become chaotic, expensive, and unpredictable, creating challenges for organizations in strategizing for profitability due to the increased complexity in planning, managing transaction costs, and establishing stable, long-term strategies [211]. However, in-

stitutional voids may also offer opportunities. In contexts where institutions are still evolving, innovative businesses can contribute to the development of new and effective institutions [209].

It is noteworthy that institutional voids are often prominent in emerging markets [210, 211]. These markets typically lack well-established institutions, which creates an environment full of institutional voids. Conversely, mature markets boast well-developed and highly regulated institutions. In some sectors, such as aviation and space exploration, institutions seem to be intentionally overregulated to ensure safety. The extensive regulations are meant to mitigate safety risks and set clear standards. Similarly, the healthcare sector, particularly in the approval of therapeutics, features comprehensive procedures and requirements to guarantee safety. However, institutional voids can still emerge within mature economies due to a 'changing world' [209]. In such cases, it is according to Hajer [209] not the absence of institutions, but rather the changing context of policy making and politics, that leads to challenges to established norms and practices and thus to institutional voids.

In the case of Mammaprint, as it sought to enter the market as a diagnostic test using gene expression, it encountered institutional voids. It did not fit the established regulatory procedures, leading to the creation of a special regulatory pathway in the U.S., where it became the first diagnostic test to receive FDA approval [74]. Similar regulatory approval was also obtained in Europe. Subsequently, Mammaprint faced another institutional void, this time related to reimbursement. This is a matter governed by political institutions. Nevertheless, it is not entirely surprising that Mammaprint encountered institutional voids. The Mammaprint test was developed around 2003 [74], during a time when genetics was a relatively new field, and the complete human genome had only recently been mapped. This is an example of a context of policy making that is expansive [209]. Hajer [209] in a paper of 2003, the same time as the invention of Mammaprint, also explicitly mentions genetics and biotechnology as spheres that are confronted with institutional voids at that time and gives as example the institutional void around property rights of genetic information [209].

It can thus be concluded that paradoxically, even within a highly regulated sector like healthcare, institutional voids can emerge, not because there is a lack of institutions – healthcare seems replete with them – but because innovative solutions often challenge and disrupt established institutional frameworks [209], creating a mismatch between innovation and existing institutions.

7.5. Reflection on Methods

Literature Review

Regarding the literature study in chapter 3, there are some limitations. It is focused on the use of AI-CDSS in the entire healthcare sector, whereas the main research question specifically pertains to cancer care. This broad focus does not provide certainty about the extent to which the factors identified in the literature study also apply to cancer care. This approach was necessary because literature focusing exclusively on AI-CDSS in cancer care does not cover the factors that impact adoption and implementation (see section 1.3).

Case Study

Single Case Study

The institutional actor analysis was applied in its entirety only to Mammaprint, focusing on a single case study. Yin [90] emphasizes a preference for a multiple case study, highlighting the limitations of a single case study in terms of generalizability and addressing alternative explanations. Despite these challenges, this research extends beyond Mammaprint by exploring implementation trajectories of other AI-CDSS in cancer care in section 6.4, drawing on interview insights and grey literature. However, it is important to note that the ability to address alternative explanations for the primary research question, "Why does the adoption and implementation of AI-CDSS in cancer care fail?" is constrained by the single case study approach. It is apparent that, for Mammaprint in the Netherlands, the current impediment lies in the reimbursement process for basic health insurance. Nevertheless, it is plausible that various other factors contribute to challenges in adoption and implementation for different AI-CDSS as various other potential limiting factors are identified in chapters 3 and 6.

Mammaprint is an Ambiguous Case

The Mammaprint case is difficult case in terms of potential conflict of interests (see section 5.5). The roles of representatives from medical professional organizations are somewhat ambiguous. This is due to the fact that several representatives of medical specialist associations have individual affiliations with NKI-AVL and/or Agendia. In the professional realm, they advocate on behalf of a single actor group within the reimbursement arena. As a result, it is not possible to say with certainty which interests they are representing. Are they advocating for the actor they formally represent at that moment, or are informal interests also at play?

Interviews

Snowball Sampling

The group of interviewees, while they were highly relevant, may also introduce potential biases within the group. Snowball sampling, in particular, may lead to a bias in participant selection due to the interconnectedness of individuals within the same professional network, possibly resulting in a homogeneity of perspectives. This potential bias could significantly impact the overall perspective and opinions gathered. For instance, it's noteworthy that the interviewed medical professionals exclusively utilized the Mammaprint test in their practice, rather than the Oncotype test. Also, all expressed a positive view of the Mammaprint test, with the exception of concerns about its cost. It remains uncertain whether this

shared enthusiasm is reflective of the broader sentiment among medical professionals in the breast cancer sector or if it's specific to the participants in this study. However, given the inclusion of these tests in clinical guidelines, a generally positive consensus regarding their usefulness is anticipated.

Interview Analysis

A total of 18 interviews with 19 interviewees, each lasting between 0.5 to 1.5 hours, were conducted. A semi-structured interview method was employed to accommodate the diverse backgrounds and roles of the interviewees. This approach required customized preparation for each interview, while providing unique perspectives on the case study. While this personalized approach proved invaluable due to the rich input gathered, it also necessitated an intensive and time-consuming analysis process. The interviews were analyzed by grouping participants based on their relevance to the arenas and phases, re-listening to recordings, extracting key insights and quotes, and categorizing them. While this method was effective, a more structured approach, such as thematic analysis through transcript analysis tools as atlas.ti, may offer better traceability of themes and may reduce bias introduced by the researcher.

(Counteract) Interviewer Bias

When conducting and analysing interviews, there is a risk of introducing bias. In case of conducting interviews the way of asking questions and the way of interacting with the interviewee can impose a bias in the outcomes. To mitigate potential interviewer bias, a concerted effort was made to maintain objectivity throughout the interview and analysis process. This involved employing strategies to encourage interviewees to speak and freely express their viewpoints by asking open-ended questions and avoiding leading cues that could steer their responses in a particular direction. The inclusion or exclusion of Mammaprint in basic health insurance emerged as a particularly sensitive and politically charged issue during the interviews. Careful attention was devoted to comprehending the underlying reasons for such stances without allowing personal opinions to overshadow the analysis. It was noted that extensive documentation of the assessment process of Mammaprint, including all final position reports [91–93], helped in a nuanced understanding. In the context of the guideline arena and the contextual factors discussed in chapter 6, a more neutral stance from interviewees was observed. Nevertheless, similar measures were applied to safeguard against potential biases. This ensured that the data collected and analyzed remained as objective and unbiased as possible, contributing to the integrity of the research findings.

Phases Framework

Usefulness

The phases framework played a crucial role in organizing and categorizing the research findings. It provided a structured approach to synthesize the literature and pinpoint the specific stage at which Mammaprint faced challenges. The framework proved especially effective in contextualizing insights from interviewees and in making comparisons across different participants. It's worth noting that these phases are interconnected and can mutually influence one another. Additionally, it became apparent that while the chronological order of the phases is generally accurate, some phases may overlap or extend beyond others, allowing for concurrent progress. For example, in the case of Mammaprint, it was already in the use phase, even though it had not yet achieved reimbursement in basic health insurance in the Netherlands. This was made possible through temporary reimbursement.

Limitations

The framework has its limitations. It doesn't differentiate between organizational, interpersonal, and intrapersonal factors, which if for example done in the study by [140]. Furthermore, it doesn't provide an in-depth exploration of different organizational levels. Research by Greenhalgh *et al.* [217] and Robert *et al.* [218] suggests that it is useful to examine the (non-)adoption of healthcare technologies at various levels, including the micro-level (individual adoption), meso-level (organizational assimilation), and macro-level (policy and regulatory environment) [217, 218]. This multi-level approach could offer a more comprehensive understanding of the complexities surrounding technology adoption in healthcare organizations.

Institutional Actor Analysis

Usefulness

The application of an institutional actor analysis provides a framework to initially delineate the procedures and identify the stakeholders involved in the reimbursement of Mammaprint. While the representation of the reimbursement arena may not accurately reflect reality, it serves as a pragmatic tool to dissect the dynamics within the arena. In this way a more nuanced understanding of actor interactions, power dynamics, and the multifaceted pathways that shape the adoption and implementation of AI-CDSS such as Mammaprint is gained. The framework has been instrumental in identifying critical points where adoption and implementation encounter impediments.

Using an institutional actor lens for investigating the adoption and implementation of AI-CDSS in cancer care is unique to the best of my knowledge.

Limitations

There is the possibility that the framework leads to an oversimplification of the decision-making process by primarily following formal procedures and identifying actors through this method. This may result in overlooking the influence of informal networks, relationships or coalitions [81]. Non-formal actors, especially, may be inadvertently marginalized. It is also possible that an actor perceived as a single entity actually also represents other actors informally. This could especially be the case for medical specialists as those are affiliated with various organizations, such as hospitals, research institutes, guideline committees, and medical specialist associations. This inherent interconnectedness between the actors in the arena adds a layer of intricacy, challenging the concept of distinct groupings and underscoring the need to avoid overly simplistic categorizations.

Furthermore, it is crucial to recognize that the insights are temporary in nature, as decision-making processes are inherently dynamic [81]. The analysis is a snapshot and as the landscape of AI-CDSS is rapidly changing, it may not fully reflect the current dynamics and influence of actors.

8

Conclusion

This chapter will first address the research questions and their corresponding answers. Subsequently, the academic contribution will be discussed, potential directions for future research will be suggested and policy advice will be proposed.

8.1. Research Question

The main research question for this thesis was as follows:

Why does the adoption and implementation of AI-CDSS in cancer care fail?

This main research question has been addressed via a case study employing an institutional actor analysis (see section 2.2.2). This overarching research question was broken down into four subquestions. A condensed overview of the responses to these subquestions is provided below. For a comprehensive analysis, please consult the relevant chapters, which are mentioned below for each subquestion.

1. What factors are identified in previous studies that influence the adoption and implementation of AI-CDSS generally?

For an overview of all the identified factors, see chapter 3. For a comparison between the factors identified through the literature and the factors identified through the case study, see the discussion section 7.1. From this subquestion it was found that in the literature there is a significant focus on factors that play a role in the use phase, while earlier stages like sell/marketing and purchase/reimbursement receive comparatively less attention, indicating a potential research gap. Also, the studies primarily center on the view of medical specialists, while a diverse array of stakeholders play pivotal roles in the adoption and implementation of AI-CDSS. However, only a limited number of papers have explored the viewpoints of multiple stakeholders. Therefore there is a need for a more comprehensive multi-actor perspective. In light of this, employing an institutional actor analysis framework (see section 2.2.2) will

provide a structured approach to understanding the challenges in adopting and implementing AI-CDSS.

2. Which decision arena(s) are critical for the adoption and implementation of a specific AI-CDSS in cancer care?

As the case study is about Mammaprint, this question will be answered about Mammaprint. The full answer is discussed in chapter 4. At first glance, it appears that both the arena in which inclusion of Mammaprint in guidelines and the arena in which inclusion of Mammaprint in basic health insurance is decided are critical to adoption. However, it appears that breast cancer guideline arena is currently not considered a pivotal factor in the adoption and implementation of Mammaprint. Although it played a critical role in earlier rounds, the absence of substantial challenges, supported by interviewees' accounts and limited documentation, suggests that further analysis of this arena is not meaningful. In contrast, inclusion in basic health insurance is a consistently critical decision arena for Mammaprint's adoption and implementation in the Netherlands [1] , with significant influencing factors identified through case study interviews and documentation of decision-making rounds [91–93].

3. What factors influence the adoption and implementation of this specific AI-CDSS in those critical decision arena(s)?

The full answer on this question is discussed in chapter 5. It is found that the medical specialists and their societies in the reimbursement arena were in favor of the use of Mammaprint and committed to get Mammaprint in basic health insurance, however the assessment of the ZiN, the actor with the highest power within the arena, was negative about the inclusion in basic health insurance both in 2010 [91] and in 2018 [92] [2]. The rejections by the ZiN were on the basis of the clinical utility of Mammaprint which is in the formal institution, namely the SWP criterion. However, it has been found that the challenge lies not only in the SWP criterion itself, but rather in its interpretation. These findings are described in section 5.4 and discussed in section 7.2. In short, it appears that the medical specialists in the reimbursement arena find the requirements for inclusion of diagnostic tests, like Mammaprint, unclear and that there is a disagreement in the level of evidence needed and in the study design needed to prove clinical utility. Also there seems to be a clash in values when it comes to the trade-off between quality of life and survival.

[1] Shortly before the publication of this study, namely on October 30, 2023, there was the announcement that Mammaprint is now included in basich health insurance [161]. So at the time of the analyses, the arena was critical because Mammaprint was not included in basic health insurance. Currently, after 13 years since the first assessment by the ZiN, this is no longer the case and so the arena is no longer critical. However, this arena has been critical for a very long time.

[2] Shortly before the publication of this study, namely on Oct. 30, 2023, there was the ruling that Mammaprint is now included in basic health insurance [161]. It was not possible to include the position report in the analyses of this study. However, some of the points of discussion are briefly mentioned in sections 5.2.4 and 5.4 because interviewees gave some away about that assessment while the final decision was not yet made.

4. What is the broader context that has an influence on those decision arena(s)?

The full answer on this question is discussed in chapter 6. In short, it became evident that it is difficult to develop AI-CDSS within a hospital as funding and expertise of product development are missing. It became clear that adoption and implementation is influenced by the type of healthcare system, where differences between Europe with a primarily public and the U.S. with a primarily private healthcare market emerged in particular. It also comes to the front that AI-CDSS cannot be seen as one term when it comes to reimbursement. In the Netherlands, the reimbursement paths are different for molecular diagnostic AI-CDSS, image analysis AI-CDSS used in the hospital and image analysis AI-CDSS used in screening. Furthermore, AI-CDSS using image analysis seem to be on the rise in radiology and pathology.

The culmination of all the answers to the sub-questions is brought together and discussed in sections 7.3 and 7.4. It is concluded that the adoption and implementation of AI-CDSS can get stuck because of the absence of supportive structures, regulations, and frameworks for reimbursement by basic health insurance, which is called an institutional void. There exists thus an institutional void for reimbursement from basic health insurance for AI-CDSS in the Netherlands.

8.2. Academic Contribution

The academic contribution works four ways. First, the study discovered evidence that supports previous research findings, indicating that the factors influencing adoption and implementation in healthcare also hold true for the context of cancer care. Second, new factors were identified that influence the adoption and implementation of AI-CDSS in cancer care. This was achieved by interviewing a unique and broad group of actors within the field. Third, it is displayed how an institutional actor framework could be successfully applied to an adoption and implementation problem in healthcare. The customization of the framework to this problem could be used in further research. Fourth, institutional voids are identified in reimbursement pathways which hamper the adoption and implementation of AI-CDSS in cancer care.

8.3. Future Research

There are multiple directions for future research which followed from this research. Initially, it is useful to examine multiple cases and be able to compare them. In this study, only the Mammaprint (and Oncotype) case was taken as the topic of interest. This limited the generalizability of the findings [90]. It would be interesting to explore more molecular diagnostic AI-CDSS for more types of cancer. For example, it would be interesting to investigate more molecular diagnostic AI-CDSS that also got stuck in the reimbursement phase to see if they ran into the same problems as Mammaprint. It would also be very interesting to investigate success factors of the adoption and implementation and therefore take an AI-CDSS as case study whose adoption and implementation has been successful. However, it is very likely that there are very few molecular diagnostic AI-CDSS for which this is applicable, but one AI-CDSS did emerge during the interviews that has been successfully adopted and implemented in a short amount of time, namely the CUPPA [187, 198, 199]. For AI-CDSS based on image analysis techniques, it is questionable whether any AI-CDSS have gone through all phases of adoption and implementation at all in the Netherlands. It would be very interesting to explore this further.

Given the perception, as revealed in this research, that the ZiN excels in evaluating drugs rather than medical diagnostic tests, it would be valuable to investigate how the ZiN handles assessments of other forms of care, such as imaging techniques like MRI and CT scans. This could provide insights into whether other innovations in healthcare face similar challenges during the reimbursement phase.

Further, it would be interesting to investigate AI-CDSS adoption and implementation beyond the borders of the Netherlands as it was mentioned in interviews that market entry and reimbursement is a lot easier in the U.S., because in Europe it requires cooperation with governments, while in the U.S. there is cooperation with private parties. Mammaprint is also mostly successful in the U.S.. In addition, it was mentioned in interviews that Mammaprint is reimbursed in many European countries, but not in the Netherlands even while it is invented here. It will be interesting to explore how the adoption and implementation of Mammaprint went in different countries.

Furthermore, it may be worthwhile to look at the funding models of developing AI-CDSS, because the research shows that in the Netherlands public-private partnerships are mainly encouraged in case research organizations want to develop AI-CDSS, leading to commercialization, while open-sourcing would also be possible, therefore it is interesting to see what the advantages and disadvantages are for future-proofing cancer care between commercializing and open-sourcing the systems, as was done with CUPPA.

Concerning the frameworks used, the IAA and phases framework, those proved valuable in this research. However testing those frameworks in other contexts would be interesting. Possibly the IAA framework could also be considered in cases where it is clear in advance that institutional voids exist to explore the exact formal and informal institutions in which they appear, but further research is needed to prove this. Additionally further developments of the phases framework when used in combination with the IAA framework could be valuable in terms of micro, meso and macro levels as the research showed that institutions are layered and also as other research about adoption of technology in healthcare do discern between those levels [217, 218].

8.4. Policy Advice

Given the research question, ”Why does the adoption and implementation of AI-CDSS in cancer care fail?”, the key policy consideration becomes, ”How can we prevent the adoption and implementation of AI-CDSS in cancer care from failing?” Since the study is focused on the Dutch context, these policy recommendations will specifically target the Dutch healthcare system.

Fill the institutional void through consensus seeking

The case study on Mammaprint has highlighted the existence of an institutional void in the institutions for reimbursement from basic health insurance which specifically pertains in the SWP criterion. This institutional void seems to apply for other types of AI-CDSS as well (see section 6.4 and section 7.3 for the discussion thereof). There is thus ambiguity regarding the reimbursement pathway for inclusion in basic health insurance. It is imperative to bring clarity to this aspect in order to fill the existing voids. In the case of Mammaprint, there are clear indications of disagreements between the policy analysts and the medical specialists (see section 5.4 for the identification and section 7.2 for the discussion). The substantial disparities in their perspectives on fundamental matters are striking. Establishing a consensus about these issues is crucial, which can be achieved through a transparent discussion involving stakeholders from various sectors. During these discussions, the opportunity can be provided to question values and formal and informal institutions, and to provide transparency about objectives.

This primarily revolves around the identified points of disagreement as discussed in chapter 5. Firstly, there is a need for consensus on the level of burden of proof required. This includes determining when retrospective or prospective studies are appropriate, as well as the follow-up duration. Addressing the level of burden of proof presents a dual challenge. On one hand, there has been a significant delay in inclusion in basic health insurance in the case of Mammaprint as a very high level was required. On the other hand, stringent requirements are understandable to ensure quality and safety. A middle ground in this could be to allow care temporarily which also allows time and cost coverage for evidence gathering for clinical utility [219]. In this way the care is accessible, while still being able to reject the care under consideration after more evidence has been collected. Presently, there are provisions in the Netherlands allowing for conditional approval [220, 221] or subsidies for promising care [222] to facilitate extensive data collection. It is recommended to continue and customize these arrangements further for AI-CDSS. Secondly, consensus is needed on the study design necessary and feasible to establish clinical utility. This includes discussing ethical considerations for randomization and measurements related to quality of life. Defining relevant outcome measures (PICO) is also crucial.Thirdly, consensus must be sought regarding the trade-off between quality of life and survival. Where does this trade-off lie, what are the underlying motivations, and what is deemed acceptable?

Pursue current initiatives to clarify the reimbursement pathway of molecular diagnostics and the use of AI

Furthermore, ongoing initiatives to clarify reimbursement in basic insurance for both molecular diagnostics, which is called 'implementation process molecular diagnostics' [193], and the use of AI, which is called 'guideline AI' [201], (both are mentioned in section 6.4) should be sustained and prioritized. This latter one is especially important because there are currently significant advancements in the field of radiology and pathology with respect to AI-CDSS, while there is considerable uncertainty regarding the path to reimbursement from basic health insurance and whether it is possible at all (see section 6.4). Filling these institutional voids is crucial to avoid the risk of substantial delays and the risk of always being too late [223].

Determine a vision for the role of AI-CDSS within the current framework for adequate cancer care

In addition, it is vital to approach AI-CDSS adoption and implementation in cancer care with foresight. Without a clear vision, progress in cancer care may be impeded. The ZiN's ongoing efforts in the framework for adequate cancer care [5, 6] could be complemented by a strategic plan for integrating AI-CDSS within this framework. Given the notable disparities in viewpoints between medical specialists and the ZiN identified in this research, collaborating closely with a diverse group of stakeholders is advisable. This vision can then be considered in formulating policies governing reimbursement within basic insurance for AI-CDSS in cancer care. Neglecting to establish a vision and strategy may inadvertently hinder, rather than enhance, the sustainability of cancer care.

References

1. MinisterievanFinanciën. *BMH 2 naar een toekomstbestendig zorgstelsel* Apr. 2020. https://www.rijksoverheid.nl/documenten/rapporten/2020/04/20/bmh-2-naar-een-toekomstbestendig-zorgstelsel.
2. *Zorguitgaven als percentage van het bruto binnenlands product* https://www.cbs.nl/nl-nl/nieuws/2022/27/zorguitgaven-stegen-in-2021-met-7-6-procent/zorguitgaven-als-percentage-van-het-bruto-binnenlands-product. (Accessed on 11/08/2023).
3. Nederland, Z. *Passende zorg | Werkagenda | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/passende-zorg. (Accessed on 08/24/2023).
4. *Samenvatting Integraal Zorgakkoord | Publicatie | Rijksoverheid.nl* https://www.rijksoverheid.nl/documenten/publicaties/2022/09/16/samenvatting-integraal-zorgakkoord. (Accessed on 08/24/2023). Sept. 2022.
5. Nederland, Z. *Signalement Passende zorg voor mensen met kanker: netwerk- en expertzorg | Rapport | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/rapport/2022/09/26/signalement-passende-zorg-kanker-netwerk--en-expertzorg. (Accessed on 08/24/2023). Sept. 2022.
6. Nederland, Z. *Passende zorg voor mensen met kanker | Werkagenda | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/kanker/passende-zorg-voor-mensen-met-kanker. (Accessed on 08/24/2023).
7. *Kanker: Zorguitgaven* July 2022. https://www.vzinfo.nl/kanker/zorguitgaven.
8. Kiemeney, L. *et al.* De kans op kanker voor Nederlanders. *Nederlands Tijdschrift voor Geneeskunde* **152,** 2233 (2008).
9. Nederland, Z. *Gezondheidszorg | Regering | Rijksoverheid.nl* https://www.rijksoverheid.nl/regering/coalitieakkoord-omzien-naar-elkaar-vooruitkijken-naar-de-toekomst/6.-gezondheid/gezondheidszorg. (Accessed on 08/24/2023).
10. Sutton, R. T. *et al.* An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ digital medicine* **3,** 17 (2020).
11. Cai, C. J. *et al. Human-centered tools for coping with imperfect algorithms during medical decision-making* in *Proceedings of the 2019 chi conference on human factors in computing systems* (2019), 1–14.
12. Moja, L. *et al.* Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. *American journal of public health* **104,** e12–e22 (2014).
13. Mohammadpour, S. *et al.* The Effects of Implementing Clinical Decision Support Systems on the Diagnosis, Treatment and Management of Cancers: A Systematic Review. *Journal of Payavard Salamat* **16.** cited By 0, 113–122. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141661235&partnerID=40&md5=a2734c6172a1774a8714575202b79d40 (2022).

14. Rahimi, R., Moghaddasi, H., Rafsanjani, K., Bahoush, G. & Kazemi, A. Effects of chemotherapy prescription clinical decision-support systems on the chemotherapy process: A systematic review. *International Journal of Medical Informatics* **122.** cited By 14, 20–26. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85057825037&doi=10.1016%5C%2fj.ijmedinf.2018.11.004&partnerID=40&md5=da532ef3ced60481a137bdb4599e0fda (2019).

15. Pan, H. *et al.* Concordance assessment of Watson for Oncology in breast cancer chemotherapy: First China experience. *Translational Cancer Research* **8.** cited By 4, 389–401. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065912442&doi=10.21037%5C%2ftcr.2019.01.34&partnerID=40&md5=5e05d202ccb8ac5c456e6cf6d31feb29 (2019).

16. Gu, J. *et al.* Deep learning radiomics of ultrasonography for comprehensively predicting tumor and axillary lymph node status after neoadjuvant chemotherapy in breast cancer patients: A multicenter study. *Cancer* **129.** cited By 1, 356–366. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85142279311&doi=10.1002%5C%2fcncr.34540&partnerID=40&md5=ae2e9e6ac0ae005641b9331ee0b2431e (2023).

17. He, F. *et al.* A Scoring System for Predicting Neoadjuvant Chemotherapy Response in Primary High-Grade Bone Sarcomas: A Multicenter Study. *Orthopaedic Surgery* **14.** cited By 0, 2499–2509. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137093203&doi=10.1111%5C%2fos.13469&partnerID=40&md5=cbda07370896605aa2777b76951bf97c (2022).

18. Filitto, G. *et al.* Automated Prediction of the Response to Neoadjuvant Chemoradiotherapy in Patients Affected by Rectal Cancer. *Cancers* **14.** cited By 4. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85129133809&doi=10.3390%5C%2fcancers14092231&partnerID=40&md5=848b748a9b78790f8262d85e73c26c21 (2022).

19. Byra, M., Dobruch-Sobczak, K., Klimonda, Z., Piotrzkowska-Wroblewska, H. & Litniewski, J. Early Prediction of Response to Neoadjuvant Chemotherapy in Breast Cancer Sonography Using Siamese Convolutional Neural Networks. *IEEE Journal of Biomedical and Health Informatics* **25.** cited By 20, 797–805. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85096874047&doi=10.1109%5C%2fJBHI.2020.3008040&partnerID=40&md5=7d429fb4036eb90b7dc638000d20a1fb (2021).

20. Song, D. *et al.* A Decision-Making Supporting Prediction Method for Breast Cancer Neoadjuvant Chemotherapy. *Frontiers in Oncology* **10.** cited By 3. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85099664088&doi=10.3389%5C%2ffonc.2020.592556&partnerID=40&md5=4cf44c9c59c104a23efd04b34757455c (2021).

21. Yang, L. *et al.* Prediction model of the response to neoadjuvant chemotherapy in breast cancers by a Naive Bayes algorithm. *Computer Methods and Programs in Biomedicine* **192.** cited By 14. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083086614&doi=10.1016%5C%2fj.cmpb.2020.105458&partnerID=40&md5=86e2cab57720ff4343d37c394d2fea27 (2020).

22. Tahmassebi, A. *et al.* Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Investigative Radiology* **54.** cited By 141, 110–117. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85056881745&doi=10.1097%5C%2fRLI.0000000000000518&partnerID=40&md5=0a84aef022c421cda4c8c1334f603e71 (2019).

23. De Ronde, J., Bonder, M., Lips, E., Rodenhuis, S. & Wessels, L. Breast cancer subtype specific classifiers of response to neoadjuvant chemotherapy do not outperform classifiers trained on all subtypes. *PLoS ONE* **9.** cited By 6. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84897805862&doi=10.1371%5C%2fjournal.pone.0088551&partnerID=40&md5=e1c533b613e2d279da29728061d9aa32 (2014).

24. Meyer-Base, A. *et al.* AI-Enhanced Diagnosis of Challenging Lesions in Breast MRI: A Methodology and Application Primer. *Journal of Magnetic Resonance Imaging* **54.** cited By 16, 686–702. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85089990009&doi=10.1002%5C%2fjmri.27332&partnerID=40&md5=2a3d3ea766945a2b39c38c48ecc14325 (2021).

25. Kleppe, A. *et al.* A clinical decision support system optimising adjuvant chemotherapy for colorectal cancers by integrating deep learning and pathological staging markers: a development and validation study. *The Lancet Oncology* **23.** cited By 4, 1221–1232. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136691615&doi=10.1016%5C%2fS1470-2045%5C%2822%5C%2900391-6&partnerID=40&md5=2135ea0324c34e4467cb4ec3424ac0de (2022).

26. Eccher, C., Seyfang, A. & Ferro, A. Implementation and evaluation of an Asbru-based decision support system for adjuvant treatment in breast cancer. *Computer Methods and Programs in Biomedicine* **117.** cited By 10, 308–321. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84908040620&doi=10.1016%5C%2fj.cmpb.2014.06.021&partnerID=40&md5=77a3cf711086bae55a0b8b374de2709d (2014).

27. Steele, S. *et al.* Time-dependent estimates of recurrence and survival in colon cancer: Clinical decision support system tool development for adjuvant therapy and oncological outcome assessment. *American Surgeon* **80.** cited By 9, 441–453. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84904547323&partnerID=40&md5=e95196e6f7b5009611fef1ba8d512d3d (2014).

28. Miao, R. *et al.* Beyond the limitation of targeted therapy: Improve the application of targeted drugs combining genomic data with machine learning. *Pharmacological Research* **159.** cited By 4. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85086462711&doi=10.1016%5C%2fj.phrs.2020.104932&partnerID=40&md5=dc4363961081a35bb8321cdd377b9e84 (2020).

29. Zhou, Z. *et al.* Characterization of aging cancer-associated fibroblasts draws implications in prognosis and immunotherapy response in low-grade gliomas. *Frontiers in Genetics* **13.** cited By 2. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85138064673&doi=10.3389%5C%2ffgene.2022.897083&partnerID=40&md5=e25796dc9f6f3e447a3ac5dae6bc3ab5 (2022).

30. Hartkopf, A. *et al.* Translational Highlights in Breast and Ovarian Cancer 2019-Immunotherapy, DNA Repair, PI3K Inhibition and CDK4/6 Therapy [Translationale Highlights Mamma-und Ovarialkarzinom 2019-Immuntherapien, DNA-Reparatur, PI3K-Inhibition und CDK4/6-Therapien]. *Geburtshilfe und Frauenheilkunde* **79.** cited By 9, 1309–1319. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85077019602&doi=10.1055%5C%2fa-1039-4458&partnerID=40&md5=891bd8d438cef0878a0406ad089d30ef (2019).

31. Xin, L. *et al.* Label-free assessment of the drug resistance of epithelial ovarian cancer cells in a microfluidic holographic flow cytometer boosted through machine learning. en. *ACS Omega* **6,** 31046–31057 (Nov. 2021).

32. Li, X. *et al.* An artificial intelligence model to predict survival and chemotherapy benefits for gastric cancer patients after gastrectomy development and validation in international multicenter cohorts. *International Journal of Surgery* **105.** cited By 0. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137368894&doi=10.1016%5C%2fj.ijsu.2022.106889&partnerID=40&md5=8d0ecf5cc92ed7d285ed8e49e86dd1d3 (2022).

33. Guo, X.-W. *et al.* Tumor-infiltrating immune cells based TMEscore and related gene signature is associated with the survival of CRC patients and response to fluoropyrimidine-based chemotherapy. *Frontiers in Oncology* **12.** cited By 0. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85138275031&doi=10.3389%5C%2ffonc.2022.953321&partnerID=40&md5=44c7befe3593c94d13aad47c610e6ac8 (2022).

34. Moradmand, H., Aghamiri, S., Ghaderi, R. & Emami, H. The role of deep learning-based survival model in improving survival prediction of patients with glioblastoma. *Cancer Medicine* **10.** cited By 5, 7048–7059. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85113678103&doi=10.1002%5C%2fcam4.4230&partnerID=40&md5=fbbb3281533c0ec245ddea983bc03d3c (2021).

35. Coquet, J. *et al.* Assessment of a Clinical Trial-Derived Survival Model in Patients with Metastatic Castration-Resistant Prostate Cancer. *JAMA Network Open* **4.** cited By 4. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85100325060&doi=10.1001%5C%2fjamanetworkopen.2020.31730&partnerID=40&md5=1f1b1d95c5dc732799b2daae4052b831 (2021).

36. Richardson, M. *et al.* Long term survival outcomes of stage I mucinous ovarian cancer - A clinical calculator predictive of chemotherapy benefit. *Gynecologic Oncology* **159.** cited By 5, 118–128. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85089557956&doi=10.1016%5C%2fj.ygyno.2020.07.011&partnerID=40&md5=e1d9b0e081a05e0d7526c917660c2ee9 (2020).

37. Wu, C.-F. *et al.* Disease-free survival assessment by artificial neural networks for hepatocellular carcinoma patients after radiofrequency ablation. *Journal of the Formosan Medical Association* **116.** cited By 25, 765–773. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85009808404&doi=10.1016%5C%2fj.jfma.2016.12.006&partnerID=40&md5=82ee26b42c970e57c9f53c5887b082ed (2017).

38. Ng, T., Chew, L. & Yap, C. A clinical decision support tool to predict survival in cancer patients beyond 120 days after palliative chemotherapy. *Journal of Palliative Medicine* **15.** cited By 15, 863–869. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84864609657&doi=10.1089%5C%2fjpm.2011.0417&partnerID=40&md5=4191f0300bc756ad14d021a8e205aee5 (2012).

39. Fang, J. *et al.* Value of texture analysis based on dynamic contrast-enhanced magnetic resonance imaging in preoperative assessment of extramural venous invasion in rectal cancer. *Insights into Imaging* **13.** cited By 0. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85142253585&doi=10.1186%5C%2fs13244-022-01316-2&partnerID=40&md5=65d12a70ff9934a06cca9bf409206c61 (2022).

40. Li, C., Chen, L., Chou, C., Ngorsuraches, S. & Qian, J. Using Machine Learning Approaches to Predict Short-Term Risk of Cardiotoxicity Among Patients with Colorectal Cancer After Starting Fluoropyrimidine-Based Chemotherapy. *Cardiovascular Toxicology* **22.** cited By 1, 130–140. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85119267523&doi=10.1007%5C%2fs12012-021-09708-4&partnerID=40&md5=70ea72b8e847813a65ee922f6d90e03b (2022).

41. Zini, E., Lanzola, G. & Quaglini, S. *Detection and management of side effects in patients with head and neck cancer* in. cited By 1 (2017). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85027279760&doi=10.1109%5C%2fRTSI.2017.8065888&partnerID=40&md5=afbf3c4a6650f5a7186ad3a38a91ca3b.

42. Ferroni, P. *et al.* Risk Assessment for Venous Thromboembolism in Chemotherapy-Treated Ambulatory Cancer Patients. *Medical Decision Making* **37.** cited By 56, 234–242. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85009730028&doi=10.1177%5C%2f0272989X16662654&partnerID=40&md5=4ffbcc2b2f60e4e12fe27b97ff9945d4 (2017).

43. Venäläinen, M. *et al.* Improved risk prediction of chemotherapy-induced neutropenia—model development and validation with real-world data. *Cancer Medicine* **11.** cited By 0, 654–663. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120410864&doi=10.1002%5C%2fcam4.4465&partnerID=40&md5=4a07a2b78af9bc87b032675c0cbba702 (2022).

44. Zhou, L., Zheng, W., Huang, S. & Yang, X. Integrated radiomics, dose-volume histogram criteria and clinical features for early prediction of saliva amount reduction after radiotherapy in nasopharyngeal cancer patients. *Discover Oncology* **13.** cited By 0. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85145422292&doi=10.1007%5C%2fs12672-022-00606-x&partnerID=40&md5=c9245e17a6755c35c5909e5771cf9d32 (2022).

45. Draguet, C. *et al.* Automated clinical decision support system with deep learning dose prediction and NTCP models to evaluate treatment complications in patients with esophageal cancer. *Radiotherapy and Oncology* **176.** cited By 3, 101–107. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85139273995&doi=10.1016%5C%2fj.radonc.2022.08.031&partnerID=40&md5=84e3dc61b300eb34ab133fedc0e18ca3 (2022).

46. Tan, B. *et al.* Personalised, Rational, Efficacy-Driven Cancer Drug Dosing via an Artificial Intelligence SystEm (PRECISE): A Protocol for the PRECISE CURATE.AI Pilot Clinical Trial. *Frontiers in Digital Health* **3.** cited By 8. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85133132301&doi=10.3389%5C%2ffdgth.2021.635524&partnerID=40&md5=93011bb44d4c2286a9747b9cb9e45282 (2021).

47. Liang, Y., Leung, K.-S. & Mok, T. A novel evolutionary drug scheduling model in cancer chemotherapy. *IEEE Transactions on Information Technology in Biomedicine* **10.** cited By 46, 237–245. https://www.scopus.com/inward/record.uri?eid=2-s2.0-33645727174&doi=10.1109%5C%2fTITB.2005.859888&partnerID=40&md5=0b8c185de831e12b05ebf9f079149bac (2006).

48. Turki, T. & Wang, J. Clinical intelligence: New machine learning techniques for predicting clinical drug response. *Computers in Biology and Medicine* **107.** cited By 19, 302–322. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85061373671&doi=10.1016%5C%2fj.compbiomed.2018.12.017&partnerID=40&md5=41e6225f730641f8abc4759fafcb3822 (2019).

49. Sun, D. *et al.* Computerized Decision Support for Bladder Cancer Treatment Response Assessment in CT Urography: Effect on Diagnostic Accuracy in Multi-Institution Multi-Specialty Study. *Tomography* **8.** cited By 1, 644–656. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85125961149&doi=10.3390%5C%2ftomography8020054&partnerID=40&md5=21090c246c133da209e784d5f250a110 (2022).

50. Davies, N. *et al.* Added value of magnetic resonance spectroscopy for diagnosing childhood cerebellar tumours. *NMR in Biomedicine* **35.** cited By 2. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85116923616&doi=10.1002%5C%2fnbm.4630&partnerID=40&md5=7de8f60d405d2c44b73e526741aa5e74 (2022).

51. Sun, D. *et al. Effect of computerized decision support on diagnostic accuracy and intra-observer variability in multi-institutional observer performance study for bladder cancer treatment response assessment in CT urography* in. **12033.** cited By 0 (2022). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85132844504&doi=10.1117%5C%2f12.2611179&partnerID=40&md5=f1c1236270fdb1997cae425590116789.

52. Kaur, R., Juneja, M. & Mandal, A. Machine learning based quantitative texture analysis of CT images for diagnosis of renal lesions. *Biomedical Signal Processing and Control* **64.** cited By 1. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85096596623&doi=10.1016%5C%2fj.bspc.2020.102311&partnerID=40&md5=cfd911e44c696f8a177b664fd36352a4 (2021).

53. Cha, K. H. *et al.* Diagnostic accuracy of CT for prediction of bladder cancer treatment response with and without computerized decision support. *Academic radiology* **26,** 1137–1145 (2019).

54. Pairawan, S. *et al.* Combined MEK/MDM2 inhibition demonstrates antitumor efficacy in TP53 wild-type thyroid and colorectal cancers with MAPK alterations. en. *Sci. Rep.* **12,** 1248 (Jan. 2022).

55. Clay, R. *et al.* Computer Aided Nodule Analysis and Risk Yield (CANARY) characterization of adenocarcinoma: Radiologic biopsy, risk stratification and future directions. *Translational Lung Cancer Research* **7.** cited By 5, 313–326. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85048692235&doi=10.21037%5C%2ftlcr.2018.05.11&partnerID=40&md5=d53ec6623a0095fe285c05a6e3d27c8e (2018).

56. Liang, G., Fan, W., Luo, H. & Zhu, X. The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomedicine and Pharmacotherapy* **128.** cited By 51. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85084861048&doi=10.1016%5C%2fj.biopha.2020.110255&partnerID=40&md5=773bb3131d10c2da3bce09c2bdff2f40 (2020).

57. Noor, A. & Assiri, A. A Rule-Based Inference Framework to Explore and Explain the Biological Related Mechanisms of Potential Drug-Drug Interactions. *Computational and Mathematical Methods in Medicine* **2022.** cited By 0. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136891789&doi=10.1155%5C%2f2022%5C%2f9093262&partnerID=40&md5=82d9cc3d6a4eca8a84836d548b03d53f (2022).

58. Regan, K., Raje, S., Saravanamuthu, C. & Payne, P. Conceptual Knowledge Discovery in Databases for Drug Combinations Predictions in Malignant Melanoma. *Studies in Health Technology and Informatics* **216.** cited By 3, 663–667. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84952013707&doi=10.3233%5C%2f978-1-61499-564-7-663&partnerID=40&md5=695f2673eb5974d57c42c09e0dee3eb2 (2015).

59. Bayrak, T., Çetin, Z., Saygılı, E. & Ogul, H. Identifying the tumor location-associated candidate genes in development of new drugs for colorectal cancer using machine-learning-based approach. *Medical and Biological Engineering and Computing* **60.** cited By 0, 2877–2897. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135822523&doi=10.1007%5C%2fs11517-022-02641-w&partnerID=40&md5=90333371f29d0c4cc3c23c667ff8be7e (2022).

60. Tibor Fekete, J. & Győrffy, B. A unified platform enabling biomarker ranking and validation for 1562 drugs using transcriptomic data of 1250 cancer cell lines. *Computational and Structural Biotechnology Journal* **20.** cited By 4, 2885–2894. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85132706750&doi=10.1016%5C%2fj.csbj.2022.06.007&partnerID=40&md5=e718fbdbab8e66a1819debc8e3e74fef (2022).

61. Hadjiiski, L. *et al. Multi-institutional observer performance study for bladder cancer treatment response assessment in CT urography with and without computerized decision support* in. **11597.** cited By 2 (2021). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103693645&doi=10.1117%5C%2f12.2582331&partnerID=40&md5=b0cc9370cc8a77d5fd1e79dfaa1431d7.

62. Hadjiiski, L. *et al.* Intraobserver variability in bladder cancer treatment response assessment with and without computerized decision support. *Tomography* **6.** cited By 11, 194–202. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85086620814&doi=10.18383%5C%2fj.tom.2020.00013&partnerID=40&md5=8ebcf12951cbc9d9b7f50dfae11b660e (2020).

63. Chapman-Sung, D. *et al. Convolutional neural network-based decision support system for bladder cancer staging in CT urography: Decision threshold estimation and validation* in. **11314.** cited By 0 (2020). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85085500223&doi=10.1117%5C%2f12.2551309&partnerID=40&md5=cfd749accbb2e424e49d118bf3fee60c.

64. Wu, E. *et al.* Deep learning approach for assessment of bladder cancer treatment response. *Tomography* **5.** cited By 25, 201–208. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85071521999&doi=10.18383%5C%2fj.tom.2018.00036&partnerID=40&md5=e27ca71df51436a554a4921e160cb438 (2019).

65. Gandikota, D. *et al. Bladder cancer staging in CT urography: Estimation and validation of decision thresholds for a radiomics-based decision support system* in. **10950.** cited By 0 (2019). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068146929&doi=10.1117%5C%2f12.2513566&partnerID=40&md5=a8acf1e866d33fd79301c53da2a2bd24.

66. Gordon, M. *et al. Bladder cancer treatment response assessment with radiomic, clinical, and radiologist semantic features* in. **10575.** cited By 1 (2018). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85046252747&doi=10.1117%5C%2f12.2294951&partnerID=40&md5=93e058c5ba4dc051a20e426f1621a466.

67. Panesar, S., D'Souza, R., Yeh, F.-C. & Fernandez-Miranda, J. Machine Learning Versus Logistic Regression Methods for 2-Year Mortality Prognostication in a Small, Heterogeneous Glioma Database. *World Neurosurgery: X* **2.** cited By 44. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85061588749&doi=10.1016%5C%2fj.wnsx.2019.100012&partnerID=40&md5=f3fa25a206cb04b15e18dc08233fbd08 (2019).

68. Zhang, L. *et al.* Real-time automatic prediction of treatment response to transcatheter arterial chemoembolization in patients with hepatocellular carcinoma using deep learning based on digital subtraction angiography videos. *Cancer Imaging* **22.** cited By 3. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130636607&doi=10.1186%5C%2fs40644-022-00457-3&partnerID=40&md5=dfd31958f1b86aa3d03002ec3df4582d (2022).

69. Park, J.-H., Baek, J.-H., Sym, S., Lee, K. & Lee, Y. A data-driven approach to a chemotherapy recommendation model based on deep learning for patients with colorectal cancer in Korea. *BMC Medical Informatics and Decision Making* **20.** cited By 7. https://www.scopus.com/

inward/record.uri?eid=2-s2.0-85091565942&doi=10.1186%5C%2fs12911-020-01265-0&partnerID=40&md5=c5a7471fe276f24fcde14ff74b1ccbd3 (2020).

70. Yang, C.-Y., Shiranthika, C., Wang, C.-Y., Chen, K.-W. & Sumathipala, S. Reinforcement learning strategies in cancer chemotherapy treatments: A review. *Computer Methods and Programs in Biomedicine* **229.** cited By 2. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85144339392&doi=10.1016%5C%2fj.cmpb.2022.107280&partnerID=40&md5=d39279f5a57e3478089ef3c233fa62d0 (2023).

71. Ebrahimi Zade, A., Shahabi Haghighi, S. & Soltani, M. Reinforcement learning for optimal scheduling of Glioblastoma treatment with Temozolomide. *Computer Methods and Programs in Biomedicine* **193.** cited By 9. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083326884&doi=10.1016%5C%2fj.cmpb.2020.105443&partnerID=40&md5=c70bb1c187953d84cd33804d49e4a5c6 (2020).

72. Tseng, H.-H. *et al.* Deep reinforcement learning for automated radiation adaptation in lung cancer: *Medical Physics* **44.** cited By 108, 6690–6705. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85037829216&doi=10.1002%5C%2fmp.12625&partnerID=40&md5=64aadbb97a2f33ad182430a659f43b97 (2017).

73. Zhao, Y., Kosorok, M. & Zeng, D. Reinforcement learning design for cancer clinical trials. *Statistics in Medicine* **28.** cited By 151, 3294–3315. https://www.scopus.com/inward/record.uri?eid=2-s2.0-70449449564&doi=10.1002%5C%2fsim.3720&partnerID=40&md5=6d68df31db7c4ca684184e8a88243c93 (2009).

74. Stork-Sloots, L. *MammaPrint: From Bench to Bedside* https://www.slideshare.net/webgoddesscathy/lisette-stork-sloots-agendia. (Accessed on 07/15/2023).

75. Van't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *nature* **415,** 530–536 (2002).

76. Van De Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *New England Journal of Medicine* **347,** 1999–2009 (2002).

77. Burstein, H. *et al.* Customizing local and systemic therapies for women with early breast cancer: the St. Gallen International Consensus Guidelines for treatment of early breast cancer 2021. *Annals of oncology* **32,** 1216–1235 (2021).

78. *Agendia Announces CE Mark for NGS-Based MammaPrint BluePrint Kit Enhancing Access to Personalized Treatment for Breast Cancer Patients in Europe - Agendia Inc.* https://agendia.com/agendia-announces-ce-mark-for-ngs-based-mammaprint-blueprint-kit/. (Accessed on 07/15/2023).

79. Beumer, I. *et al.* Equivalence of MammaPrint array types in clinical trials and diagnostics. *Breast cancer research and treatment* **156,** 279–287 (2016).

80. *Tested and approved | Breast Cancer Diagnosis Treatment* https://www.mammaprint.co.nz/tested-and-approved/. (Accessed on 07/15/2023).

81. Enserink, B. *et al. Policy Analysis of Multi-Actor Systems* (TU Delft Open, 2022).

82. *Our supervision of medical technology | Medical Technology | Health and Youth Care Inspectorate* https://english.igj.nl/medical-technology/our-supervision-of-medical-technology. (Accessed on 10/05/2023).

83. Bryson, J. M. What to do when stakeholders matter: stakeholder identification and analysis techniques. *Public management review* **6,** 21–53 (2004).

84. Eden, C. & Ackermann, F. *Making strategy: The journey of strategic management* (Sage, 1998).

85. Hermans, L. M. & Cunningham, S. W. *Actor and strategy models: Practical applications and step-wise approaches* (John Wiley & Sons, 2018).

86. North, D. C. Economic performance through time. *The American economic review* **84,** 359–368 (1994).

87. Ostrom, E. *Understanding Institutional Diversity* (Princeton University Press, 2005).

88. Teisman, G. R. Models for research into decision-makingprocesses: on phases, streams and decision-making rounds. *Public administration* **78,** 937–956 (2000).

89. *Beoordeling stand van de wetenschap en praktijk 2023 | Publicatie | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/publicatie/2023/04/11/beoordeling-swp-2023. (Accessed on 07/16/2023).

90. Yin, R. K. *Case study research and applications* (Sage, 2018).

91. *Standpunt MammaPrint bij de behandeling van het mammacarcinoom | Standpunt | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/standpunten/2010/10/29/standpunt-mammaprint-bij-de-behandeling-van-het-mammacarcinoom. (Accessed on 07/16/2023). Oct. 2010.

92. *Standpunt MammaPrint® bij vrouwen met vroeg stadium borstkanker (herbeoordeling) | Standpunt | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/publicaties/standpunten/2018/09/27/standpunt-mammaprint-bij-vrouwen-met-vroeg-stadium-borstkanker-herbeoordeling. (Accessed on 07/16/2023). Sept. 2018.

93. *Standpunt Oncotype bij vroeg stadium borstkanker | Standpunt | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/publicaties/standpunten/2021/08/11/standpunt-oncotype-bij-vroeg-stadium-borstkanker. (Accessed on 07/16/2023). Aug. 2021.

94. *Standpunt Oncotype DX® en MammaPrint® bij vroeg stadium borstkanker | Werkagenda | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/kanker/standpunt-oncotype-dx-en-mammaprint-bij-borstkanker. (Accessed on 07/15/2023).

95. Sax, D. R., Sturmer, L. R., Mark, D. G., Rana, J. S. & Reed, M. E. Barriers and opportunities regarding implementation of a machine learning-based acute heart failure risk stratification tool in the emergency department. *Diagnostics* **12,** 2463 (2022).

96. Schwartz, J. M. *et al.* Factors influencing clinician trust in predictive clinical decision support systems for in-hospital deterioration: qualitative descriptive study. *JMIR Human Factors* **9,** e33960 (2022).

97. Watson, J. *et al.* Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers? *JAMIA open* **3,** 167–172 (2020).

98. Matthiesen, S. *et al.* Clinician preimplementation perspectives of a decision-support tool for the prediction of cardiac arrhythmia based on machine learning: near-live feasibility and qualitative study. *JMIR human factors* **8,** e26964 (2021).

99. Wang, D. *et al. "Brilliant AI doctor" in rural clinics: Challenges in AI-powered clinical decision support system deployment* in *Proceedings of the 2021 CHI conference on human factors in computing systems* (2021), 1–18.

100. Wijnhoven, F. Organizational learning for intelligence amplification adoption: Lessons from a clinical decision support system adoption project. *Information Systems Frontiers* **24,** 731–744 (2022).

101. Wijnhoven, F. *Challenges of adopting human-centered intelligent systems: An organizational learning approach* in *Human Centred Intelligent Systems: Proceedings of KES-HCIS 2020 Conference* (2021), 13–25.

102. Kux, B. R., Majeed, R. W., Ahlbrandt, J. & Röhrig, R. *Factors Influencing the Implementation and Distribution of Clinical Decision Support Systems (CDSS).* in *GMDS* (2017), 127–131.

103. Petitgand, C., Motulsky, A., Denis, J.-L. & Régis, C. in *Digital Personalized Health and Medicine* 1001–1005 (IOS Press, 2020).

104. Alsobhi, M. *et al.* Facilitators and Barriers of Artificial Intelligence Applications in Rehabilitation: A Mixed-Method Approach. *International Journal of Environmental Research and Public Health* **19,** 15919 (2022).

105. Gunasekeran, D. V. *et al.* Acceptance and perception of artificial intelligence usability in eye care (APPRAISE) for Ophthalmologists: a multinational perspective. *Frontiers in medicine* **9,** 875242 (2022).

106. Van Gils, A. M. *et al.* Assessing the views of professionals, patients, and care partners concerning the use of computer tools in memory clinics: international survey study. *JMIR formative research* **5,** e31053 (2021).

107. Cabral, B. P., Braga, L. A. M., Syed-Abdul, S. & Mota, F. B. Future of Artificial Intelligence Applications in Cancer Care: A Global Cross-Sectional Survey of Researchers. *Current Oncology* **30,** 3432–3446 (2023).

108. Choudhury, A. Factors influencing clinicians' willingness to use an AI-based clinical decision support system. *Frontiers in Digital Health* **4,** 920662 (2022).

109. Choudhury, A., Asan, O. & Medow, J. E. Effect of risk, expectancy, and trust on clinicians' intent to use an artificial intelligence system–Blood Utilization Calculator. *Applied Ergonomics* **101,** 103708 (2022).

110. Kashyap, S., Morse, K. E., Patel, B. & Shah, N. H. A survey of extant organizational and computational setups for deploying predictive models in health systems. *Journal of the American Medical Informatics Association* **28,** 2445–2450 (2021).

111. Van de Burgt, B. W. *et al.* Combining text mining with clinical decision support in clinical practice: a scoping review. *Journal of the American Medical Informatics Association* **30,** 588–603 (2023).

112. Knop, M., Weber, S., Mueller, M. & Niehaves, B. Human factors and technological characteristics influencing the interaction of medical professionals with artificial intelligence–enabled clinical decision support systems: literature review. *JMIR Human Factors* **9,** e28639 (2022).

113. Schwartz, J. M., Moy, A. J., Rossetti, S. C., Elhadad, N. & Cato, K. D. Clinician involvement in research on machine learning–based predictive clinical decision support for the hospital setting: A scoping review. *Journal of the American Medical Informatics Association* **28,** 653–663 (2021).

114. O'Sullivan, D., Fraccaro, P., Carson, E. & Weller, P. Decision time for clinical decision support systems. *Clinical medicine* **14,** 338 (2014).

115. Varghese, J. Artificial intelligence in medicine: chances and challenges for wide clinical adoption. *Visceral medicine* **36,** 443–449 (2020).

116. Amann, J., Blasimme, A., Vayena, E., Frey, D. & Madai, V. I. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC medical informatics and decision making* **20,** 1–9 (2020).

117. Tran, K. A. *et al.* Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Medicine* **13,** 1–17 (2021).

118. Baxter, S. L. & Lee, A. Y. Gaps in standards for integrating artificial intelligence technologies into ophthalmic practice. *Current Opinion in Ophthalmology* **32,** 431–438 (2021).

119. Harwich, E. & Laycock, K. Thinking on its own: AI in the NHS. *Reform Research Trust* (2018).

120. Chua, I. S. *et al.* Artificial intelligence in oncology: Path to implementation. *Cancer Medicine* **10,** 4138–4149 (2021).

121. Varghese, J. Artificial intelligence in medicine: chances and challenges for wide clinical adoption. *Visceral medicine* **36,** 443–449 (2020).

122. Weidele, D. K. I. *et al. AutoAIViz: opening the blackbox of automated artificial intelligence with conditional parallel coordinates* in *Proceedings of the 25th International Conference on Intelligent User Interfaces* (2020), 308–312.

123. Tozzi, A. E. & Cinelli, G. Informed consent and artificial intelligence applied to RCT and COVID-19. *BioLaw Journal-Rivista di BioDiritto,* 97–108 (2021).

124. Pierce, R. L., Van Biesen, W., Van Cauwenberge, D., Decruyenaere, J. & Sterckx, S. Explainability in medicine in an era of AI-based clinical decision support systems. *Frontiers in Genetics* **13,** 903600 (2022).

125. Heinrichs, B. & Eickhoff, S. B. Your evidence? Machine learning algorithms for medical diagnosis and prediction. *Human brain mapping* **41,** 1435–1444 (2020).

126. Academy, B. & the Royal Society. *Data management and use: governance in the 21st century* 2017.

127. Amann, J., Blasimme, A., Vayena, E., Frey, D. & Madai, V. I. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC medical informatics and decision making* **20,** 1–9 (2020).

128. Denecke, K. & Baudoin, C. R. A review of artificial intelligence and robotics in transformed health ecosystems. *Frontiers in medicine* **9,** 795957 (2022).

129. Ng, W. Y. *et al.* Updates in deep learning research in ophthalmology. *Clinical Science* **135,** 2357–2376 (2021).

130. Goh, W. P., Tao, X., Zhang, J. & Yong, J. Decision support systems for adoption in dental clinics: a survey. *Knowledge-Based Systems* **104,** 195–206 (2016).

131. Shah, T., Rabhi, F. & Ray, P. Investigating an ontology-based approach for Big Data analysis of inter-dependent medical and oral health conditions. *Cluster Computing* **18,** 351–367 (2015).

132. Krumholz, H. M. Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. *Health Affairs* **33,** 1163–1170 (2014).

133. Wang, Y., Kung, L. & Byrd, T. A. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological forecasting and social change* **126,** 3–13 (2018).

134. Abujaber, A. A., Nashwan, A. J. & Fadlalla, A. Harnessing machine learning to support evidence-based medicine: A pragmatic reconciliation framework. *Intelligence-Based Medicine* **6,** 100048 (2022).

135. Abujaber, A. A., Nashwan, A. J. & Fadlalla, A. Enabling the adoption of machine learning in clinical decision support: A total interpretive structural modeling approach. *Informatics in Medicine Unlocked* **33,** 101090 (2022).

136. Stone, P. *et al.* Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence. *arXiv preprint arXiv:2211.06318* (2022).

137. Cam, A., Chui, M. & Hall, B. Global AI Survey: AI proves its worth, but few scale impact (2019).

138. Pereira, L., Rijo, R., Silva, C. & Martinho, R. Text mining applied to electronic medical records: a literature review. *International Journal of E-Health and Medical Communications (IJEHMC)* **6,** 1–18 (2015).

139. Von Wedel, P. & Hagist, C. Economic value of data and analytics for health care providers: hermeneutic systematic literature review. *Journal of Medical Internet Research* **22,** e23315 (2020).

140. Roberts, M. C. *et al.* ONCOLOGISTS'BARRIERS AND FACILITATORS FOR ONCOTYPE DX USE: QUALITATIVE STUDY. *International journal of technology assessment in health care* **32,** 355–361 (2016).

141. Walter, Z. & Lopez, M. S. Physician acceptance of information technologies: Role of perceived threat to professional autonomy. *Decision Support Systems* **46,** 206–215 (2008).

142. Varonen, H., Kortteisto, T., Kaila, M. & Group, E. S. What may help or hinder the implementation of computerized decision support systems (CDSSs): a focus group study with physicians. *Family practice* **25,** 162–167 (2008).

143. Tucci, V., Saary, J. & Doyle, T. E. Factors influencing trust in medical artificial intelligence for healthcare professionals: A narrative review. *Journal of Medical Artificial Intelligence* **5** (2022).

144. Romero-Brufau, S. *et al.* A lesson in implementation: a pre-post study of providers' experience with artificial intelligence-based clinical decision support. *International journal of medical informatics* **137,** 104072 (2020).

145. Shortliffe, E. H. & Sepúlveda, M. J. Clinical decision support in the era of artificial intelligence. *Jama* **320,** 2199–2200 (2018).

146. Alabi, R. O. *et al.* Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. *Artificial intelligence in medicine* **115,** 102060 (2021).

147. Bur, A. M. *et al.* Machine learning to predict occult nodal metastasis in early oral squamous cell carcinoma. *Oral oncology* **92,** 20–25 (2019).

148. Altmann, A., Toloşi, L., Sander, O. & Lengauer, T. Permutation importance: a corrected feature importance measure. *Bioinformatics* **26,** 1340–1347 (2010).

149. Zhou, N. *et al.* Concordance Study Between IBM Watson for Oncology and Clinical Practice for Patients with Cancer in China. *Oncologist* **24.** cited By 44, 812–819. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053281103&doi=10.1634%5C%2ftheoncologist.2018-0255&partnerID=40&md5=dcf3ae5782fe599415954665d3cae14f (2019).

150. Bates, D. W., Auerbach, A., Schulam, P., Wright, A. & Saria, S. Reporting and implementing interventions involving machine learning and artificial intelligence. *Annals of internal medicine* **172,** S137–S144 (2020).

151. Ciecierski-Holmes, T., Singh, R., Axt, M., Brenner, S. & Barteit, S. Artificial intelligence for strengthening healthcare systems in low-and middle-income countries: a systematic scoping review. *npj Digital Medicine* **5,** 162 (2022).

152. Fan, X. *et al.* Utilization of self-diagnosis health chatbots in real-world settings: case study. *Journal of medical Internet research* **23,** e19928 (2021).

153. Almaghrabi, N. S. & Bugis, B. A. Patient confidentiality of electronic health records: A recent review of the Saudi literature. *Dr. Sulaiman Al Habib Medical Journal* **4,** 126–135 (2022).

154. Gianfrancesco, M. A., Tamang, S., Yazdany, J. & Schmajuk, G. Potential biases in machine learning algorithms using electronic health record data. *JAMA internal medicine* **178,** 1544–1547 (2018).

155. *Improved breast cancer care by innovative diagnostic test MammaPrint | MammaPrint | Project | Fact sheet | H2020 | CORDIS | European Commission* https://cordis.europa.eu/project/id/672570. (Accessed on 07/17/2023).

156. *Richtlijn MAMMACARCINOOM* https://www.nabon.nl/wp-content/uploads/2022/11/Richtlijn-Mammacarcinoom-2008.pdf. (Accessed on 07/16/2023). Sept. 2008.

157. *Mammacarcinoom Richtlijn* https://www.nabon.nl/wp-content/uploads/2022/11/Richtlijn-mammacarcinoom-2012.pdf. (Accessed on 07/16/2023). 2012.

158. *Borstkanker - Algemeen - Richtlijn - Richtlijnendatabase* https://richtlijnendatabase.nl/richtlijn/borstkanker/algemeen.html. (Accessed on 07/16/2023).

159. *Richtlijncommissie - NABON* https://www.nabon.nl/richtlijncommissie/#:~:text=De%20richtlijncommissie%20werkt%20aan%20modulaire,stellen%20de%20gereviseerde%20teksten%20op.. (Accessed on 07/16/2023).

160. Cardoso, F. *et al.* 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *New England Journal of Medicine* **375,** 717–729 (2016).

161. *Standpunt - MammaPrint en Oncotype DX vergoede zorg voor bepaalde groep vrouwen | Standpunt | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/standpunten/2023/10/24/standpunt---mammaprint-en-oncotype-dx-vergoede-zorg-voor-bepaalde-groep-vrouwen. (Accessed on 11/03/2023).

162. *Beoordeling van nieuwe zorg | Over ons | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/over-ons/werkwijzen-en-procedures/adviseren-over-en-verduidelijken-van-het-basispakket-aan-zorg/beoordeling-van-nieuwe-zorg#:~:textDe%20beoordeling%20verschilt%20per%20zorgvorm&text=De%20wijze%20waarop%20wij%20dit,de%20wetenschap%20en%20de%20praktijk'.. (Accessed on 07/15/2023).

163. Bijl, E. *Advies en standpunt: wat is het verschil?* | *Magazine* | *Zorginstituut Nederland* https://www.zorginstituutnederland.nl/over-ons/publicaties/magazine/2020/04/23/advies-en-standpunt-wat-is-het-verschil. (Accessed on 07/15/2023). Apr. 2020.

164. *Sluis voor dure geneesmiddelen* | *Over ons* | *Zorginstituut Nederland* https://www.zorginstituutnederland.nl/over-ons/programmas-en-samenwerkingsverbanden/horizonscan-geneesmiddelen/sluis-voor-dure-geneesmiddelen. (Accessed on 07/15/2023).

165. Staal, P. *et al. Betekenis en beoordeling criterium 'plegen te bieden'* | *Rapport* | *Zorginstituut Nederland* https://www.zorginstituutnederland.nl/over-ons/publicaties/rapport/2008/11/17/betekenis-en-beoordeling-criterium-plegen-te-bieden. (Accessed on 07/16/2023). Nov. 2008.

166. *Reglement Adviescommissie Pakket Zorginstituut Nederland* | *Besluit* | *Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland. (Accessed on 07/21/2023).

167. *wetten.nl - Regeling - Besluit zorgverzekering - BWBR0018492* https://wetten.overheid.nl/BWBR0018492/2023-01-01. (Accessed on 08/26/2023). Aug. 2023.

168. Bueno-de-Mesquita, J. M. *et al.* Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). *The lancet oncology* **8,** 1079–1087 (2007).

169. Cardoso, F., Piccart-Gebhart, M., Van't Veer, L., Rutgers, E., *et al.* The MINDACT trial: the first prospective clinical validation of a genomic tool. *Molecular oncology* **1,** 246–251 (2007).

170. *Medische tests (beoordeling stand van de wetenschap en praktijk)* | *Standpunt* | *Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/standpunten/2011/01/20/medische-tests-beoordeling-stand-van-de-wetenschap-en-praktijk#:~:text=Zorginstituut%20Nederland%20Publicaties-,Medische%20tests%20(beoordeling%20stand%20van%20de%20wetenschap%20en%20praktijk),wordt%20vergoed%20uit%20het%20basispakket. (Accessed on 07/17/2023). Jan. 2011.

171. Piccart, M. *et al.* 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *The Lancet Oncology* **22,** 476–488 (2021).

172. *Final assessment report on MammaPrint® - Added value of using the gene expression signature test MammaPrint® for adjuvant chemotherapy decision-making in early breast cancer - EUnetHTA* https://www.eunethta.eu/final-assessment-report-on-mammaprint-added-value-of-using-the-gene-expression-signature-test-mammaprint-for-adjuvant-chemotherapy-decision-making-in-early-breast-cancer/. (Accessed on 07/17/2023). Jan. 2018.

173. De Glas, N. A. *et al.* Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. *The Lancet Oncology* **15,** 722–729 (2014).

174. Sparano, J. A. *et al.* Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *New England Journal of Medicine* **379,** 111–121 (2018).

175. *Antwoord vragen Van Gerven over de vergoeding van de mammaprint vanuit de zorgverzekering* | *Tweede Kamer der Staten-Generaal* https://www.tweedekamer.nl/kamerstukken/kamervragen/detail?id=2011Z16056&did=2011D42044. (Accessed on 07/17/2023). Aug. 2011.

176. *Het advies om de Mammaprint niet te vergoeden vanuit de basisverzekering | Tweede Kamer der Staten-Generaal* https://www.tweedekamer.nl/kamerstukken/kamervragen/detail?id=2018Z17350&did=2018D47125. (Accessed on 07/17/2023). Oct. 2018.

177. *Het bericht dat de MammaPrint niet vergoed wordt | Tweede Kamer der Staten-Generaal* https://www.tweedekamer.nl/kamerstukken/kamervragen/detail?id=2018Z17648&did=2018D47678. (Accessed on 07/17/2023). Oct. 2018.

178. *Het bericht 'MammaPrint niet in basispakket' | Tweede Kamer der Staten-Generaal* https://www.tweedekamer.nl/kamerstukken/kamervragen/detail?id=2018Z17864&did=2018D48143. (Accessed on 07/17/2023). Oct. 2018.

179. Drukker, C. A. *et al.* A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. *International journal of cancer* **133,** 929–936 (2013).

180. Paik, S. *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor–positive breast cancer. *Journal of clinical oncology* **24,** 3726–3734 (2006).

181. *Beoordeling dure specialistische geneesmiddelen | Over ons | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/over-ons/werkwijzen-en-procedures/adviseren-over-en-verduidelijken-van-het-basispakket-aan-zorg/beoordeling-van-geneesmiddelen/beoordeling-dure-specialistische-geneesmiddelen. (Accessed on 07/17/2023).

182. *Agendia - Company Profile - Tracxn* https://tracxn.com/d/companies/agendia/__WMxbQX7NzAZTLiJ0SBLUVdWE91g2WxkvJDH79m4KW7c#:~:text=Agendia%20has%20raised%20a%20total,lead%20by%20Athyrium%20Capital%20Management. (Accessed on 07/17/2023).

183. *Exact Sciences | EXAS Stock Price, Company Overview & News* https://www.forbes.com/companies/exact-sciences/?sh=2cc6d850131f. (Accessed on 07/17/2023).

184. *wetten.nl - Regeling - Zorgverzekeringswet - BWBR0018450* https://wetten.overheid.nl/BWBR0018450/2023-01-01. (Accessed on 07/17/2023).

185. *Mammaprint* https://www.avl.nl/nieuwsberichten/2018/mammaprint/. (Accessed on 07/17/2023). Oct. 2018.

186. Klinkhamer, J., De Lameillieure, K., Audeh, W. & van der Meijden, C. POSC212 A Most-Similar Case Analysis of Gene Expression Test Reimbursement Assessment in Europe. *Value in Health* **25,** S149 (2022).

187. Foundation, H. M. *CUPPA: Localisatie van de primaire tumor op basis van het DNA van de uitzaaiing - Hartwig Medical Foundation* https://www.hartwigmedicalfoundation.nl/cuppa-localisatie-van-de-primaire-tumor-op-basis-van-het-dna-van-de-uitzaaiing/. (Accessed on 09/29/2023). Sept. 2021.

188. *EUR-Lex - 32017R0745 - EN - EUR-Lex* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745. (Accessed on 10/22/2023).

189. *EUR-Lex - 32017R0746 - EN - EUR-Lex* https://eur-lex.europa.eu/eli/reg/2017/746/oj. (Accessed on 10/22/2023).

190. *New Regulations of Medical Devices* https://health.ec.europa.eu/medical-devices-sector/new-regulations_en. (Accessed on 10/22/2023).

191. Wolfson, B. J. *Agendia took gamble on FDA approval for breast-cancer test – Orange County Register* https://www.ocregister.com/2013/03/27/agendia-took-gamble-on-fda-approval-for-breast-cancer-test/. (Accessed on 08/24/2023). Mar. 2013.

192. Stevens, J. *Klokkenluider borstkankertest blijkt zelf lobbyist* https://www.ftm.nl/artikelen/waarschuwing-voor-onbetrouwbare-borstkankertest. (Accessed on 08/24/2023). July 2016.

193. *Uitvoeringstraject Moleculaire diagnostiek | Werkagenda | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/kanker/uitvoeringstraject-moleculaire-diagnostiek. (Accessed on 09/28/2023).

194. *Formulieren - ZN - Formulieren* https://www.znformulieren.nl/formulieren?folderid=6829604864&title=CieBAG. (Accessed on 10/22/2023).

195. *Moleculaire diagnostiek - Zorgverzekeraars Nederland* https://www.zn.nl/dossiers/moleculaire-diagnostiek/. (Accessed on 10/22/2023).

196. *Formulieren - ZN - Formulieren* https://www.znformulieren.nl/formulieren?folderid=8443822080&title=MDx. (Accessed on 10/11/2023).

197. *Pakketadvies osimertinib (Tagrisso®) bij niet-kleincellige longkanker (NSCLC) | Advies | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/adviezen/2021/11/15/pakketadvies-osimertinib-tagrisso-bij-niet-kleincellige-longkanker-nsclc. (Accessed on 10/22/2023).

198. Schipper, L. *et al.* Complete genomic characterization in patients with cancer of unknown primary origin in routine diagnostics. *ESMO open* **7,** 100611 (2022).

199. *hmftools/cuppa/README.md at master · hartwigmedical/hmftools · GitHub* https://github.com/hartwigmedical/hmftools/blob/master/cuppa/README.md. (Accessed on 09/29/2023). Mar. 2023.

200. *Prestatiebeschrijvingbeschikking - Whole Genome Sequencing, diagnose en behandeladvies bij Primaire Tumor Onbekend (PTO) - TB-REG 21658-01 - Nederlandse Zorgautoriteit* https://puc.overheid.nl/nza/doc/PUC_638877_22/1/. (Accessed on 10/23/2023).

201. *Leidraad kwaliteit AI in de zorg | Publicatie | Data voor gezondheid* https://www.datavoorgezondheid.nl/wegwijzer-ai-in-de-zorg/documenten/publicaties/2021/12/17/leidraad-kwaliteit-ai-in-de-zorg. (Accessed on 08/25/2023). Dec. 2021.

202. *ScreenPoint Medical raises $28 million for further development - Series C funding led by Insight Partners - Radboudumc* https://www.radboudumc.nl/en/news-items/2021/screenpoint-medical-raises-28-million-for-further-development. (Accessed on 10/23/2023).

203. *Onderzoeksrapport: Artificiële intelligentie en Passende zorg | Rapport | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/rapport/2022/09/29/onderzoeksrapport-artificiele-intelligentie-en-passende-zorg. (Accessed on 10/23/2023).

204. *Onderzoeksrapport: Zorginnovatieverhalen - een verkenning van 4 innovatievoorbeelden in de zorg | Rapport | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/publicaties/rapport/2021/06/25/kpmg-rapport-zorginnovatieverhalen. (Accessed on 10/23/2023).

205. Wilson, J. M. G., Jungner, G., Organization, W. H., *et al.* Principles and practice of screening for disease (1968).

206. Andermann, A., Blancquaert, I., Beauchamp, S. & Déry, V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bulletin of the World Health Organization* **86,** 317–319 (2008).

207. *Theorie en achtergrond | RIVM* https://www.rivm.nl/bevolkingsonderzoek-en-screening en/theorie-en-achtergrond. (Accessed on 10/25/2023).

208. Ten Heuvelhof, E. *Strategisch gedrag in netwerken: wat het is, hoe we het ontwikkelen, wat we ervan moeten vinden* (Boom uitgevers, 2016).

209. Hajer, M. Policy without polity? Policy analysis and the institutional void. *Policy sciences* **36,** 175–195 (2003).

210. Khanna, T. & Palepu, K. Why focused strategies. *Harvard business review* **75,** 41–51 (1997).

211. Khanna, T. & Palepu, K. G. *Winning in emerging markets: A road map for strategy and execution* (Harvard Business Press, 2010).

212. Liedong, T. A., Peprah, A. A., Amartey, A. O. & Rajwani, T. Institutional voids and firms' resource commitment in emerging markets: A review and future research agenda. *Journal of International Management* **26,** 100756 (2020).

213. Bruton, G. D., Ahlstrom, D. & Li, H.-.-L. Institutional theory and entrepreneurship: where are we now and where do we need to move in the future? *Entrepreneurship theory and practice* **34,** 421–440 (2010).

214. Spigel, B. & Harrison, R. Toward a process theory of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal* **12,** 151–168 (2018).

215. McAdam, M., Crowley, C. & Harrison, R. T. "To boldly go where no [man] has gone before"-institutional voids and the development of women's digital entrepreneurship. *Technological Forecasting and Social Change* **146,** 912–922 (2019).

216. Khanna, T. & Yafeh, Y. Business groups in emerging markets: Paragons or parasites? *ECGI-Finance Working Paper* (2005).

217. Greenhalgh, T. *et al.* Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of medical Internet research* **19,** e8775 (2017).

218. Robert, G., Greenhalgh, T., MacFarlane, F. & Peacock, R. Organisational factors influencing technology adoption and assimilation in the NHS: a systematic literature review. *Report for the National Institute for Health Research Service Delivery and Organisation programme* (2009).

219. IJzerman, M. J. *et al.* Towards routine implementation of liquid biopsies in cancer management: it is always too early, until suddenly it is too late. *Diagnostics* **11,** 103 (2021).

220. *Voorwaardelijk toegelaten zorg | Verzekerde zorg | Zorginstituut Nederland* https://https://www.zorginstituutnederland.nl/Verzekerde+zorg/voorwaardelijke-toegelaten-zorg. (Accessed on 09/29/2023).

221. *Voorwaardelijke toelating weesgeneesmiddelen, conditionals en exceptionals | Werkagenda | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/voorwaardelijke-toelating-weesgeneesmiddelen-conditionals-en-exceptionals. (Accessed on 09/29/2023).

222. *Veelbelovende zorg: subsidieregeling voor onderzoek naar potentieel veelbelovende zorg | Werkagenda | Zorginstituut Nederland* https://www.zorginstituutnederland.nl/werkagenda/veelbelovende-zorg. (Accessed on 09/29/2023).

223. IJzerman, M. J., Koffijberg, H., Fenwick, E. & Krahn, M. Emerging use of early health technology assessment in medical product development: a scoping review of the literature. *Pharmacoeconomics* **35,** 727–740 (2017).

224. Thakur, R., Hsu, S. H. & Fontenot, G. Innovation in healthcare: Issues and future trends. *Journal of Business Research* **65,** 562–569 (2012).

225. Rutgers, E. *et al.* CBO-richtlijn'Behandeling van het mammacarcinoom', Dutch Institute for Healthcare Improvement guideline," Treatment of breast cancer". *Nederlands Tijdschrift voor Geneeskunde* **146,** 2144–2151 (2002).

A

AI-CDSS Literature Search

The goal of the literature search is to identify the existing research in the field of AI-driven CDSS in cancer care and gain an understanding of the types of studies that have been conducted to date.

Different search terms were tried. For clinical decision support system (CDSS) there are some variants, such as clinical decision aid or clinical decision support program. However, it is noteworthy that not all the tools mentioned in the retrieved papers are computerized, and a substantial number of them are not employed for analytical purposes, but rather for databases or platforms. In case 'clinical' is not used in front of the other words, the searches are not healthcare specific, but come up with applications in lots of other industries. For AI a broad search is used that includes many different key terms for AI to be as complete as possible. So the search includes "simple" machine learning as well as deep learning, reinforcement learning and text mining. The search is then combined with terms about cancer. This because the focus area is specifically on cancer care. Using those three search terms resulted in too many papers (1539 on Scopus) of which many were about the prospective of AI and were also not cancer specific, but mentioning cancer besides other healthcare areas or medical conditions. Therefore the search terms treatment or chemotherapy were added to focus on cancer treatment decisions and to not get papers about the general prospective of AI in healthcare. This resulted in around 609 papers on Scopus, which is too much to manually analyze. The application area of the AI-driven CDSS was besides in chemotherapy treatment decisions also in other areas such as radiation dose decisions. To get a more workable number of papers the scope of the search was made smaller. Now the focus was only on chemotherapy and not on treatment in general. This resulted in a workable number of papers (131). The final search term combinations for Scopus and Pubmed are shown below.

The search was conducted using the following search terms on Scopus:
(TITLE-ABS-KEY ((clinical AND decision AND support AND system) OR cdss)
AND TITLE-ABS-KEY ((artificial AND intelligence) OR 'ai' OR (machine AND learning) OR (deep AND learning) OR (neural AND networks) OR (reinforcement AND learning) OR (neural AND language AND processing) OR 'nlp' OR (text AND mining))
AND TITLE-ABS-KEY (cancer OR oncolog*)
AND TITLE-ABS-KEY (chemo*)

And the following search terms on PubMed: (All fields)
((clinical decision support system) OR cdss)
AND ((artificial intelligence) OR ’ai’ OR (machine learning) OR (deep learning) OR (neural networks) OR (reinforcement learning) OR (neural language processing) OR ’nlp’ OR (text mining))
AND (cancer OR oncolog*)
AND (chemo*)

B

Explanation Phases Framework

The phases framework as described in the methodology (chapter 2) and shown in figure B.2, is based on the life cycle of medical technology as described by the Dutch ministry of health [82] as shown in figure B.1.

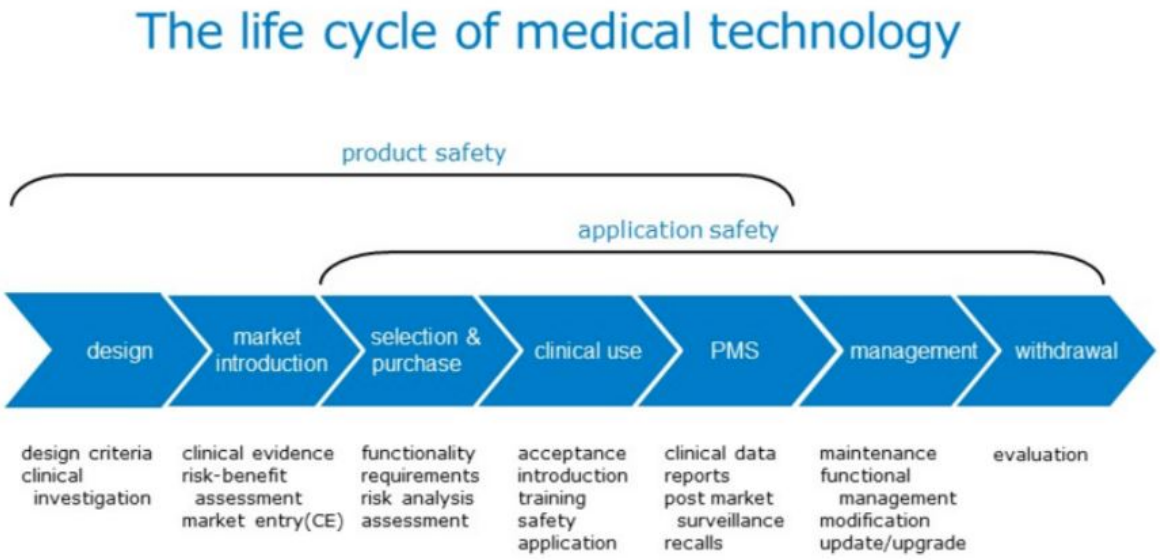

Figure B.1: The life cycle of medical technology [82]

Only the initial four phases of the framework are included in this research, as the latter three phases (post-market surveillance, management, and withdrawal) are deemed outside the scope. This is because the study's focus is on adoption and implementation, rather than the entirety of the product's life cycle. Also a new phase is introduced as the 'sell/marketing' phase is not within the diagram of the lifecycle of medical technology, but is deemed important for the adoption of innovation [224]. This sell/marketing phase is after market introduction and before selection and purchase. The phases framework employed in this research (see figure B.2) utilizes distinct terminology to precisely delineate the scope of each phase. Despite differing names, the content remains consistent, signifying the same underlying concept as the phases used in the life cycle diagram. The design phase corresponds to the

development phase, the market introduction phase aligns with the regulation phase, and a new phase, termed the 'sell/marketing' phase, has been introduced, then the selection & purchase phase equates to the purchase/reimbursement phase, and the clinical use phase denotes the use phase.

Phases in implementation and adoption of AI-CDSS

Development	Regulation	Sell/Marketing	Purchase/Reimbursement	Use
The first phase is the development phase, where research focuses on the development, design, and testing of AI-CDSS	The second phase is the regulation phase, where research discusses legal and regulatory frameworks, as well as ethical and legal considerations surrounding data privacy and security.	The third phase is the sell/marketing phase, where research examines how AI-CDSS are presented and promoted to potential buyers	The fourth phase is the purchase/reimbursement phase, where research explores the decision-making process involved in purchasing and implementing AI-CDSS, including reimbursement regulations by national health insurance institutes and health insurance companies	The fifth phase is the use phase, where research studies the actual implementation and use of AI-CDSS in a clinical setting. This phase is unique to the hospital setting, as it involves the day-to-day use of the system by clinicians and other healthcare professionals

Figure B.2: Phases in adoption and implementation of AI-CDSS

C

Overview of question topics, characterization of interviewees and interview analysis

C.1. Topics discussed during case study interviews

During the interviews there were three main groupings of topics discussed, which are: the medical guidelines, the reimbursement by basic health insurance and AI-CDSS. Within these three main topics, subtopics were present as discussed below. For all of these initially planned for subtopics, some questions that are asked are given as example to get an idea about the type of questions asked. Additionally, other topics emerged during the interviews from findings and subsequent follow-on questions, not initially planned for, but discussed in case it was relevant for the main and/or subquestions.

C.1.1. Medical Guidelines

General procedures of medical guideline drafting

- What are the procedures for drafting the medical guidelines?
- Who is involved in drafting the medical guidelines?

Procedures for drafting the breast cancer guidelines

- What parties are involved in drafting the breast cancer guidelines?
- What are the role of the parties involved in drafting the breast cancer guidelines?

Process of inclusion of Mammaprint in the breast cancer guidelines

- What were the reasons to include Mammaprint in the breast cancer guidelines in 2008?
- Why were there updates of the way Mammaprint is included in the breast cancer guidelines in 2012 and 2020?
- Were there points of discussion for taking up Mammaprint in the medical guidelines?

Role of guidelines in adoption and implementation of care

- Why are there medical guidelines?
- What is the role of medical guidelines in the adoption and implementation of new forms of care?

C.1.2. Reimbursement by basic health insurance

General procedures for assessments of care for inclusion in basic health insurance

- What are the procedures for assessing if care is eligible to get reimbursed by basic health insurance?
- Who is involved in assessing if care is eligible to get reimbursed by basic health insurance?

Assessments of Mammaprint for inclusion in basic health insurance

- How did the assessment for Mammaprint in 2010 (and 2018 and now, 2023) went?
- Which parties were involved in the assessment of Mammaprint?
- Were there points of discussion about the assessment of Mammaprint?

Assessments of Oncotype for inclusion in basic health insurance

- How did the assessment for Oncotype in 2021 (an now, 2023) went?
- What differences between Oncotype and Mammaprint caused that Oncotype is reimbursed and Mammaprint is not?

C.1.3. AI-CDSS

- Do you have any experience using AI-CDSS?
- Do you have any experience with the reimbursement of AI-CDSS?
- What are the qualities you would consider necessary for an AI-CDSS?

Mammaprint

- Do you have experience in using Mammaprint and/or Oncotype?
- When do you use Mammaprint and/or Oncotype?
- Do you have a preference between Mammaprint and Oncotype?

The development of an AI-CDSS

- How does the development of an AI-CDSS proceed?

Starting a(n AI-CDSS) spin-off

- Why was a spin-off created from Mammaprint?
- How does starting a spin-off work?

AI-CDSS and screening

- What is the current status of AI-CDSS in respect to screening?

C.2. Characterization of Interviewees

In table 2.1 in chapter 2 an overview of the interviewees is shown. To give a bit more background about the interviewees that discussed the reimbursement and/or guideline arena, additional information is given in table C.1. In this additional table it is shown if they were previously and/or currently involved. Also it is shown if they said general stuff about the procedures for reimbursement in basic health insurance and/or the uptake in medical guidelines as they have experience with this, not for Mammaprint, but for other sorts of care.

Table C.1: Characterization of the interviewees in the guideline and the reimbursement arena

Involvement status	Involvement	Interview ID
Characterization of Interviewees involved in the Guideline arena		
involved in the arena	Interviewees that are directly involved in the drafting of the breast cancer guidelines	1, 2, 14, 15
previously involved	Interviewees that were involved in the drafting of the breast cancer guidelines during all guidelines in which Mammaprint is included	1, 14
currently involved	Interviewees that are currently involved in the drafting of the breast cancer guidelines	1, 2, 15
general procedural validation	Interviewees that elaborated and/or validated the general guideline drafting procedures which are described in grey literature	1, 2, 4, 14, 15, 16
Characterization of Interviewees involved in the Reimbursement arena		
involved in the arena	Interviewees that are directly involved in the discussion about the reimbursement of Mammaprint in basic health insurance	3, 4, 8, 10, 12, 17, 18, 19
previously involved	Interviewees that were involved in one or more of the previous decision rounds about reimbursement of Mammaprint	4, 8, 10, 12, 17
currently involved	Interviewees that are involved in the current decision round about reimbursement of Mammaprint	3, 12, 17, 18, 19
general procedural validation	Interviewees that elaborated and/or validated (part of) the procedure for reimbursement of Mammaprint in basic health insurance which are described in grey literature	3, 4, 10, 12, 17, 19

C.3. Codes and Statements identified from the Interview Analysis

Codes were created during the interview analysis for statements that were of relevance for answering one or more of the subquestions. In table C.2 the codes and statements that came up during the interviews are shown. The codes could thus only be identified after the interviews were held from answers of the interviewees on open questions. For all the statements it is indicated which interviewees stated it. The statements are grouped according to the following themes:

- Guidelines
- Reimbursement
- No clear requirements for inclusion of diagnostic tests
- (High) Level of burden of proof needed
- Disagreement on study design between medical specialists and ZiN
- Priorities in trade-off between quality of life and survival differ
- Development
- Regulation
- Sell/Marketing
- Purchase/Reimbursement
- Use

Table C.2: The codes and statements that came up during the interviews coupled to the interviewees that stated it and grouped per theme

Code	Statement	Interview ID
Guidelines		
importance of guidelines	Interviewees underscore the importance of guideline inclusion for the acceptance of AI-CDSS, including Mammaprint, in clinical practice.	3, 6, 7, 9, 12, 14
previously no disagreement	Interviewees mentioned that they did not recall encountering significant challenges in previous decision rounds	1, 14
currently no disagreement	Interviewees indicated that there is likely no alteration to the section concerning gene expression profiles as there is no disagreement at this point on how Mammaprint is included in the guidelines.	2, 15

interplay guidelines and reimbursement	Interviewees mention there is usually consideration in terms of reimbursement or non-reimbursement and inclusion in guidelines. However, even while Mammaprint is not reimbursed by basic health insurance, it is in the guidelines as medical specialists deemed its utilization necessary in specific cases.	14, 15
Reimbursement		
importance of inclusion in basic insurance	Interviewees underscore the importance of inclusion in basic health insurance for the use of AI-CDSS, including Mammaprint, in clinical practice.	2, 3, 6, 9, 12, 18
effects of reimbursement of Oncotype	interviewees mention that many hospitals switched to Oncotype when Oncotype was reimbursed by basic health insurance and Mammaprint was not	2, 3
reimbursement and use in practice	interviewees mention that reimbursement problems cause Mammaprint to be used less in practice	3, 18
problems reimbursement and moving to U.S.	interviewee mentions the problems with reimbursement in the Netherlands was one of the reasons the registration holder moved to the U.S.	9
reimbursement and business case	interviewees mention the non-reimbursement of AI-CDSS make a business case difficult	6, 12
previous disagreement	Interviewees indicate there have been disagreements about inclusion in basic health insurance of Mammaprint	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19
current disagreement	Interviewees indicated there is once again in the current (2023) decision round disagreement over inclusion in basic health insurance	12, 19
alternative reimbursement routes for Mammaprint	Interviewees indicate that Mammaprint has been extensively utilized in practice for a prolonged period, even without basic insurance coverage as over the years, various alternative funding options have emerged and continue to exist	2, 3, 8, 10, 12, 17

arrangement with registration holder of Mammaprint	Interviewees mentioned some hospitals have arrangements about reimbursement with Agendia, the registration holder of Mammaprint, such that they can still use the Mammaprint test even though it is not reimbursed by basic health insurance	2, 3, 12
reimbursement out of leniency or in supplementary insurance	Interviewees mentioned that most health insurers reimbursed Mammaprint out of leniency and/or in supplementary insurance	8, 10, 12, 18, 17
usage of the RASTER study	Interviewees indicated that the CvZ did little with the results of the RASTER study, although the CvZ did pay for the study	10, 12
reassessments of care by the ZiN	Interviewees indicate that as soon as a drug or medical test is under review by the ZiN, it remains there. This is not officially established according to the interviewees, but is an informal rule	17, 19
No clear requirements for inclusion of diagnostic tests		
evidence diagnostic tests	Interviewees mention it is not clear what the ZiN wants to see as evidence of clinical utility for diagnostic tests in general	3, 5, 6, 10, 12
Subjectivity in procedure	Interviewees mention there is subjectivity in the assessment procedure	6, 19
Changing requirements	Interviewees mention requirements for Mammaprint change throughout the decision rounds	4, 12
expertise of the in ZiN drug assessments	Interviewees mention the ZiN is experienced in doing drug assessments	1, 3, 5, 12
expertise of the ZiN in diagnostics test assessments	Interviewees mention the ZiN is not experienced in doing assessments of diagnostic tests	1, 5, 12
assessment of drugs vs. diagnostics	Interviewees indicate that drugs and diagnostics should be assessed differently	1, 3, 5, 12
(High) Level of burden of proof needed		
follow-up duration	Interviewees mention the follow-up duration of 5 years was not long enough according to the ZiN	4, 12, 13

disagreement about follow-up	Interviewees explicitly said they disagree with the ZiN about the follow-up duration needed to prove clinical utility	4, 12
retrospective vs. Prospective	The interviewee mentions the requirement of a retrospective study for reimbursement in basic health insurance for a diagnostic test as the Mammaprint is unreasonably high	12
difficulties of RCTs	Interviewees mention the RCT studies (the MINDACT and the TAILORx) cannot be done again because of being very expensive, taking a very long time, being very difficult in practical and logistical terms and being unethical to randomize because there is already enough evidence for the tests	3, 12
Disagreement on study design between medical specialists and ZiN		
creation of subgroups	Interviewees mention that the patient group of the MINDACT study is split in subgroups by the ZiN which reduces the confidence intervals, while this should be seen as one patient group	3, 4, 12, 19
primary and secondary endpoints	Interviewees mention that the ZiN uses not the primary endpoints of the MINDACT trial, but the secondary when assessing the clinical utility	5, 12
randomization and ethics	Interviewees mention that not all patients groups can be randomized because of ethical considerations	3, 12
study design requirements	Interviewees mention it is more often that the study design of clinical trials does not align with what the ZiN requires when assessing	5, 10, 12, 13
Priorities in trade-off between quality of life and survival differ		
quality of life data	Interviewees mention the ZiN would like to have seen more quality of life data, while the MINDACT did not included this	3, 4, 10, 12
evidence of effects of chemotherapy	Interviewees mention it is odd that evidence is asked to prove that not undergoing chemotherapy will lead to improved quality of life as this is obvious	3, 12
Development		

AI-CDSS development in hospital setting	Interviewee mention there are difficulties developing an AI-CDSS within a hospital setting in the regulatory domain in the financial domain in the scalability domain	 7, 9, 18 4, 6, 7, 9, 12, 13 7
Regulation		
costs of certification process	Interviewees mention the certification process of AI-CDSS are very expensive	7, 9
costs of RCTs	Interviewees mention the execution of RCTs are very expensive	7, 9, 12, 13
public-private parnterships	Interviewees mention the availability of subsidies requires often the involvement of private partners to form public-private partnerships.	4, 6, 7
Sell/Marketing		
RCT and Market acceptance	Interviewees mention that conducting an RCT has been observed to significantly impact market acceptance	9, 12
Oncotype financial resources	Interviewees indicate there are substantial financial resources at the disposal of Oncotype.	3, 10
Purchase/Reimbursement		
Mammaprint reimbursed out of leniency	According to many interviewees the Mammaprint test was reimbursed by many health insurers out of leniency.	8, 10, 12, 17, 18
Use		
Mammaprint not associated with AI	Interviewees mention Mammaprint is generally in the medical world not associated with AI	4, 6
image analysis associated with AI	Interviewees indicate image analysis based techniques are primarily seen as AI	3, 4, 6
self-learning associated with AI	The interviewee indicates to associate self-learning processes with AI	13
AI-CDSS in radiology and pathology	Interviewees mention developments with AI are in radiology and pathology and/or mention examples of AI-CDSS developments in radiology or pathology	6, 7, 11, 12, 13
use of AI	Interviewees indicate the use of AI is not a problem as long as it is proven reliable	1, 2, 3, 5

gene expression was seen as black box	Interviewees indicate Mammaprint was seen at black box at the start, because of the use of gene expression profiles which were something very new	7, 10, 12, 19
output of Mammaprint	The interviewee indicates to be in favor of the output of the Mammaprint test compared to the Oncotype as a high or low risk group is easier interpretable than an intermediate risk group	3
output and room for interpretation	Interviewees indicate to not have a preference in output type of Mammaprint or Oncotype as they prefer room for interpretation which is possible for both output types as a score bar is also included	1, 2
binary output	Interviewee mentions a mere 'yes' or 'no' as output would be unpleasant	4, 6

D

Case study Medical Guideline Arena

During the identification of critical decision arenas, it was found that the inclusion of and the way how Mammaprint is included in the medical guidelines is currently not a critical decision arena. However it has been a critical decision arena during earlier decision-making rounds. In first instance it would have been interesting to identify factors that influence the uptake in the medical guidelines. However, after checking the grey literature and interviewing actors involved in the guideline drafting procedure, no difficulties about the gene expression module appeared. Probably also partly as interviewees did not remember the details and meeting minutes were not available. To give some insights in the drafting of medical guidelines, the arena of the breast cancer guideline drafting and the appearance of Mammaprint in the guidelines, this is taken up in this part of the appendix.

D.1. Dutch Medical Guidelines drafting procedure

No one is required by law to establish medical guidelines. However, in principle The law states that you must provide care that meets the to the quality standard, as established by the scientific association, which is why the existence of a guideline does matter. Medical guidelines are created by mandated members of medical scientific professional associations. There are also members of the patients' association on the committee. The knowledge institution of the Federation of Medical Specialists (FMS) does the process guidance for guideline drafting. Each guideline revision cycle starts with a bottleneck analysis. This is followed by a prioritization of the most important bottlenecks as according to the interviews not all can be addressed due to budget, people, and time constraints. Next, these pieces of the guideline are revised or developed. This drafting is done using the principles of evidence-based medicine. Once the guidelines are finished, the draft version is sent to the medical scientific association and the patients' association. There is then the opportunity to respond. After this, any feedback can be processed. Once the medical scientific associations are in agreement, the guidelines can be authorized and published on the guideline database of the FMS.

D.2. Genesis of the National Breast Cancer Guidelines

The first national breast cancer guidelines were developed in collaboration with the National Breast Cancer Consultation Netherlands (NABON) and Dutch Institute for Healthcare Improvement (CBO) in 2002 [225]. These were developed following evidence-based medicine [225]. In 2005 IKNL gets involved and in 2012 IKNL takes over from CBO. Then in 2017, FMS gets involved, where at first they only support financially, in the current round, the whole process management is in the hands of FMS. For process guidance during the process, the responsibilities changed from CBO to IKNL to FMS. The NABON has been involved from the beginning and still is. The NABON has no formal or legal status, but it is a collaborative effort among medical specialists in the breast cancer field with a number of areas of focus and the guidelines is one of them. As of 2015, there is officially a guidelines committee. When composing the committee, the geographical distribution of committee members, proportional representation from the various associations and bodies involved, as well as a spread in academic background were taken into account as much as possible [159].

D.3. Appearance of Mammaprint and Oncotype throughout the years

The appearance of Mammaprint and Oncotype in the Dutch national breast cancer guidelines throughout the years can be seen in table D.1. The first appearance of gene expression profiles in the guidelines was in 2008. It was then recommended that they be used only in study settings. In 2012 there was an update. Then it was advised for hormone-sensitive invasive ductal carcinoma to use gene expression profiles in case of doubt about giving adjuvant chemotherapy after having used the traditional prognostic factors. In 2020 there was again an update of the guidelines, here it is mentioned that the test may have added value and the group for whom it can be used is made more specific, and again there should be doubt about the indication of adjuvant chemotherapy. For further details on appearance in the guidelines, see table D.1. It became clear from interviewees in the breast cancer guideline arena that the guidelines are currently under review and a new version will be released soon (in the second half of 2023).

D.4. Drafting Breast Cancer Guidelines

The general process of drafting and revising medical guidelines is described earlier in section D.1. For breast cancer specific the currently involved actors are the FMS as process facilitator and the medical specialist associations as owner and drafters of the guidelines. Besides this the draft version of the guidelines are sent to ZN and ZiN for feedback. These actors could be identified via the most recent guidelines [158], the current guideline committee on the NABON website [159] and the case study interviews. During the interviews, the actors could be further validated and their specific roles could be identified. The guidelines themselves also discuss procedures and responsible persons [156] [157] [158]. According to the guidelines, gene expression profiles were one of the bottlenecks in all three decision rounds [156] [157] [158]. The interviewees in the breast cancer guideline arena (see figure 2.1) indicate that gene expression profiles were not identified as bottlenecks to tackle in the current round of decisions. Thus, it seems that in the upcoming new version, there will be no changes in the gene expression module compared to the 2020 guideline. It is unclear if the module about gene ex-

Publication Years	Appearance in the Guidelines	Advice
2008 [156]	Several gene expression profiles have been shown, for patients with negative axillary lymph nodes, to better distinguish subgroups with favorable or unfavorable prognosis and predict response to drug treatment. These profiles include the MammaPrint® 70-gene profile or ”Amsterdam signature, the 76-gene profile or ”Rotterdam signature” and the 21-gene profile or Oncotype DXTM panel. Large studies, on the order of several thousand patients, have been initiated in recent years to evaluate the prognostic and predictive value of the available gene profiles to adequately validate and to further test their usefulness in clinical decision-making. At this time it is therefore advised that gene expression profiles should preferably be determined as part of these studies.	The available commercial gene expression profiles are preferably applied in study settings
2012 [157]	Several gene expression profiles have been shown in retrospective studies to better distinguish subgroups with favorable or unfavorable prognosis in patients than traditional risk estimates.	In individual cases in hormone-sensitive invasive ductal carcinoma, validated gene expression profiles can be used if there is doubt about the indication for adjuvant chemotherapy based on traditional prognostic factors.
2020 [158]	Gene expression profiles such as MammaPrint and OncotypeDX have prognostic significance in addition to the known classical clinical and pathological factors. However, they cannot replace the known classical clinical and pathological factors. These gene expression profiles do not constitute a predictive test for the effect of chemotherapy. Based on currently available data, no preference can be given to MammaPrint or OncotypeDX when deciding on adjuvant chemotherapy.	Based on the MINDACT study, MammaPrint may have an added value in patients with pT1-2N0 and pT1N1 (maximum 1 positive axillary gland) ER+HER2- invasive carcinoma NST (formerly invasive ductal breast cancer), in whom adjuvant chemotherapy is considered based on clinical and pathological factors (see also adjuvant chemotherapy module). This includes patients with pT1-2N0 or pT1N1 disease with the following features: grade 1 3-5 cm pN0; grade 1 2-5 cm pN1; grade 2 2-5 cm pN0; grade 2 0-5 cm pN1; grade 3 1-2 cm pN0.

Table D.1: An overview of the appearance of Mammaprint in the guidelines throughout the years

pression profiles were on the initial list with bottlenecks. It was mentioned by interviewee 16 that in the current revision round about 160 bottlenecks were identified and that they picked up 14 of them. In the guidelines it is mentioned per bottleneck which group of persons is responsible for revision. These specialists write the guidelines on behalf of their medical specialty society and therefore on behalf of their constituency. According to interviewee 14, at the time of meetings, the medical specialists also have a letter with them stating that they have been delegated by their medical specialist association. On the NABON website it is stated which individuals are involved on behalf of which medical scientific societies for the current decision round [159]. However, as these guidelines have not been published

yet, there is no information available online about the current bottlenecks and the responsible persons per bottleneck.

As far as interviews revealed with medical specialists and the process facilitator in the breast cancer guideline arena (interviewees 1, 2, 14 and 15), everyone agreed with the way gene expression testing came into the guidelines. However, no minutes or reports are available of how the decisions in the various decision rounds exactly went. Since some decision rounds did take place many years ago (the first one around 15 years ago), it should be mentioned that the interviewees also indicated that they did not remember it all in detail or were not involved in all decision rounds.