Original Research Article

MDM Policy & Practice

MDM Policy & Practice
2025, Vol. 10(1) 1–15

DOI: 10.1177/23814683251328377
journals.sagepub.com/home/mpp
Sage

Bridging the Gap: A Mixed-Methods Study on Factors Influencing Breast Cancer Clinicians' Decisions to Use Clinical Prediction Models

Mary Ann E. Binuya, Sabine C. Linn, Annelies H. Boekhout, Marjanka K. Schmidt, and Ellen G. Engelhardt

Abstract

Background. Clinical prediction models provide tailored risk estimates that can help guide decisions in breast cancer care. Despite their potential, few models are widely used in clinical practice. We aimed to identify the factors influencing breast cancer clinicians' decisions to adopt prediction models and assess their relative importance. **Methods.** We conducted a mixed-methods study, beginning with semi-structured interviews, followed by a nationwide online survey. Thematic analysis was used to qualitatively summarize the interviews and identify key factors. For the survey, we used descriptive analysis to characterize the sample and Mann–Whitney *U* and Kruskal–Wallis tests to explore differences in score (0 = *not important* to 10 = *very important*) distributions. **Results.** Interviews ($N = 16$) identified eight key factors influencing model use. Practical/methodological factors included accessibility, cost, understandability, *objective* accuracy, actionability, and clinical relevance. Perceptual factors included acceptability, *subjective* accuracy, and risk communication. In the survey ($N = 146$; 137 model users), clinicians ranked online accessibility (median score = 9 [interquartile range = 8–10]) as most important. Cost was also highly rated, with preferences for freely available models (9 [8–10]) and those with reimbursable tests (8 [8–10]). Formal regulatory approval (7 [5–8]) and direct integration with electronic health records (6 [3–8]) were considered less critical. Subgroup analysis revealed differences in score distributions; for example, clinicians from general hospitals prioritized inclusion of new biomarkers more than those in academic settings. **Conclusions.** Breast cancer clinicians' decisions to initiate use of prediction models are influenced by practical and perceptual factors, extending beyond technical metrics such as discrimination and calibration. Addressing these factors more holistically through collaborative efforts between model developers, clinicians, and communication and implementation experts, for instance, by developing clinician-friendly online tools that prioritize usability and local adaptability, could increase model uptake.

Corresponding Author
Ellen G. Engelhardt, Division of Molecular Pathology/Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam, 1066 CX, The Netherlands; (e.engelhardt@nki.nl).

Highlights

- Accessibility, cost, and practical considerations, such as ease of use and clinical utility, were prioritized slightly more than technical validation metrics, such as discrimination and calibration, when deciding to start using a clinical prediction model.
- Most breast cancer clinicians valued models with clear inputs (e.g., variable definitions, cutoffs) and outputs; few were interested in the exact model specifications.
- Perceptual or subjective factors, including perceived accuracy and peer acceptability, also influenced model adoption but were secondary to practical considerations.
- Sociodemographic variables, such as clinical specialization and hospital setting, influenced the importance of factors for model use.

Keywords
breast cancer, clinical prediction models, clinician preferences, model implementation, personalized decision making, risk assessment

Date received: May 21, 2024; accepted: February 15, 2025

Division of Molecular Pathology, the Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands (MAEB, SCL, MKS, EGE); Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands (MAEB); Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands (MAEB, MKS); Department of Pathology, University Medical Center Utrecht, Utrecht, the Netherlands (SCL); Division of Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands (SCL); Division of Psychosocial Research and Epidemiology, the Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands (AHB, EGE).
The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: SCL reports grants paid to the institute from AstraZeneca, Eurocept Plaza, Roche, Genentech, Gilead Sciences, GSK, Novartis, and Agendia outside the submitted work; consulting fees from AstraZeneca paid to the institute; educational fees from Daiichi Sankyo paid to the institute; other financial support for attending meetings from Daiichi Sankyo; nonfinancial support from Genentech (drug), Roche (drug), Gilead Sciences (drug), Novartis (drug), Agendia (gene expression tests), and AstraZeneca (drug). SCL has a patent (PCT/EP2022/73958) pending on a method for assessing homologous recombination deficiency in ovarian cancer cells. The other authors declare no conflicts of interest. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study is part of the PREDICT-NL project (pathways to smart validation and clinical embedding of prediction tools for oncology and beyond) funded by Health˜Holland grant No. LSHM19121, received by MKS, EGE, and MAB. Research at the Netherlands Cancer Institute is supported by institutional grants of the Dutch Cancer Society; the Dutch Ministry of Health, Welfare and Sport; and individual donors.

Breast cancer remains a leading cause of mortality among women worldwide.[1] Adjuvant systemic treatments (e.g., chemotherapy, endocrine therapy) administered after tumor resection can reduce the risk of distant metastases, recurrence, and breast cancer–related mortality, although they may also lead to side effects, including fatigue, nausea, infertility, nerve damage, and osteoporosis, with rare cases of secondary cancers.[2] Clinical prediction models, such as PREDICT, provide individualized survival (or risk) estimates that can aid clinicians in evaluating the added benefits of these treatments.[3,4] The use of PREDICT and similar prognostic models can thus improve the identification of women most likely to benefit from additional interventions.[2] Beyond prognostication, models such as BOADICEA can assess an individual's predisposition to breast cancer and inform preventive interventions and screening strategies.[5] By enabling personalized risk assessments and the application of more tailored treatment strategies, clinical prediction models may help reduce overtreatment, limit low-value care, and improve quality of life.[1,2]

Despite the proliferation of clinical prediction models over the past two decades, a gap persists between model development and real-world implementation.[3,6–8] For instance, there are now more than 60 breast cancer prognostication models,[8] of which only a select few have been widely adopted in clinical practice.[9] Notably,

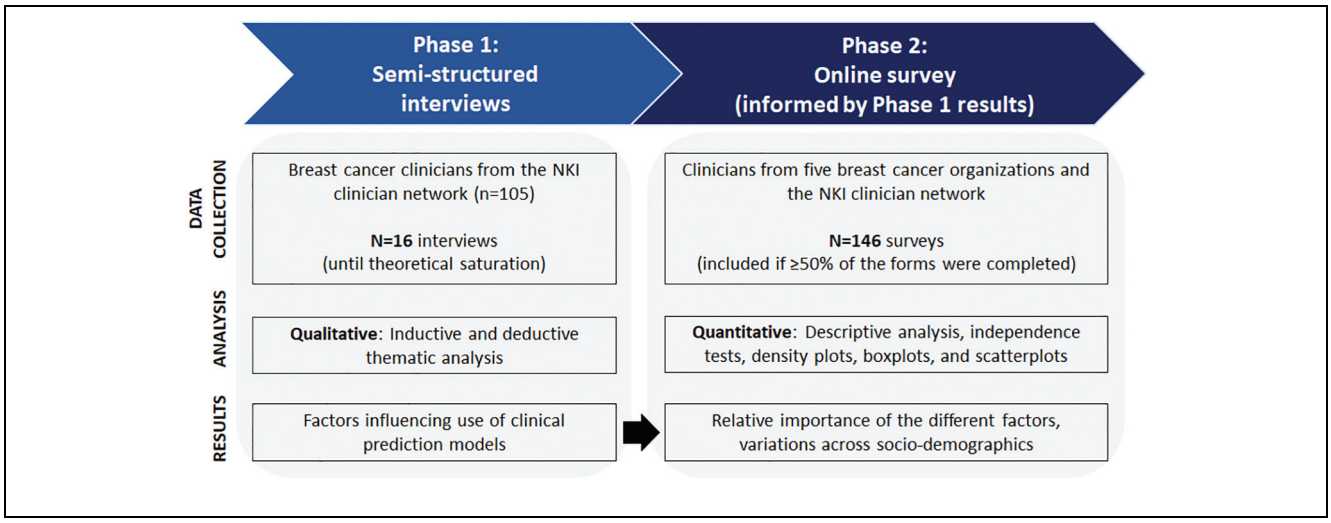

Figure 1 Flow diagram of the sequential exploratory mixed-methods approach applied in the study.
Phase 1 interview results directly informed the development of the survey questionnaire used for data collection in phase 2. The additional 5 organizations used to recruit clinicians in phase 2 were the Netherlands Cancer Institute's breast cancer working group, the Dutch Society for Radiation Oncology, the Dutch Society for Surgical Oncology, the Hereditary Breast and Ovarian Cancer Research Netherlands, and the Netherlands Comprehensive Cancer Organization. Due to limited access to external mailing lists used for survey distribution, calculating the response rate in phase 2 was not feasible.

high-performing machine learning–based models are not always extensively used; others lacking rigorous evaluation are sometimes implemented.[10,11] This variation in model uptake highlights the complexities of translating models from research into routine clinical practice. Understanding the barriers and facilitators to initial model adoption is essential to informing the design of models that can integrate more effectively into clinical practice and meet the needs of clinicians and patients.

Currently, most research on implementation barriers focuses on patient decision aids, which do not necessarily integrate data-driven algorithms.[12–17] In addition, literature on the requirements for implementing clinical prediction models often emphasizes performance metrics, such as discrimination and calibration, while overlooking broader factors that may influence their clinical use. Although some studies have assessed the impact of (widely) used breast cancer prediction models post implementation,[18,19] the factors driving their initial clinical adoption remain largely unexplored. In specialized fields such as breast cancer, where integration of validated prediction models can guide complex treatment decisions and improve patient outcomes, understanding what drives clinicians to adopt these tools is essential.

We aimed to 1) identify the key factors influencing breast cancer clinicians' decisions to initiate the use of clinical prediction models and 2) assess the relative importance of these factors. As a secondary exploratory aim, we examined whether the importance of these factors varied according to sociodemographic characteristics, including age, sex, clinical specialization, years of experience, and healthcare setting.

Methods

Study Design and Ethics Approval

We employed an exploratory sequential mixed-methods design (Figure 1). The study was approved by the Institutional Review Board at the Netherlands Cancer Institute (IRBd22-078). Verbal informed consent was obtained prior to the interviews, while survey participants provided informed consent on the first page of the questionnaire. No personal identifiers were retained during data collection or analysis.

Phase 1: Qualitative Semi-structured Interviews

Participants. We conducted semi-structured interviews to identify key factors influencing clinicians' decisions to start using clinical prediction models. Using purposive sampling, we recruited medical oncologists, radiation oncologists, surgical oncologists, clinical geneticists, and nurse specialists via the study team's breast cancer clinician network. All participants were familiar with, but not necessarily using, clinical prediction models. The

interview guide (Supplementary File A) was pilot tested with three clinicians prior to finalization.

Data collection. A trained epidemiologist with qualitative research experience (M.A.B.) conducted the interviews between March and November 2022, either in person, online, or by telephone, whichever was most convenient for the participant. The interviews had an average duration of 30 minutes. Data collection continued until we reached theoretical saturation (i.e., no new themes were emerging) and had sufficient diversity in clinical specialization, geographic location, and hospital type. All interviews were conducted in English, transcribed verbatim, and de-identified after transcription. Clinicians did not receive compensation for their participation.

Data analysis. We employed inductive and deductive thematic analysis to summarize the interview transcripts. Two researchers (M.A.B. and E.G.E.) independently double-coded all transcripts using ATLAS.ti versions 9 and, later, 23.[20] The first five interviews served as a training set for developing the initial codebook. Regular meetings were held to review the codes and resolve any discrepancies, with a third researcher (M.K.S.) providing input when needed. As coding progressed, initial codes were consolidated into broader themes representing key factors. Less prevalent themes were retained as subfactors or omitted based on relevance and data support. New themes were added to the codebook as they emerged. Previously coded interviews were reviewed to ensure that the new codes were captured, with recoding performed when necessary. Quotes from the interviews were used to illustrate the themes.

Phase 2: Quantitative Online Survey

Questionnaire development. An online survey was administered to a nationwide network of clinicians to assess the relative importance of factors identified in phase 1. The phase 1 codebook informed the design of the questionnaire. Clinicians were asked to provide sociodemographic information (age, biological sex, years of experience, clinical specialization, hospital type, and region) and rate the importance of each subfactor on a numeric scale from 0 (*not important*) to 10 (*very important*) (see Supplementary File B for the full questionnaire). To keep the survey concise, the selection of sociodemographic variables was limited to only key characteristics that we expected might influence initial model uptake. Clinicians who reported never using prediction models were asked about their reasons for non-use. The survey was pilot tested by 2 team members (M.K.S. and A.H.B.) and an experienced medical oncologist (S.C.L.). The anonymous survey was carried out using Castor EDC.[21]

Participants and data collection. The survey was conducted from May to October 2023 and recruited clinicians in the Netherlands involved in breast cancer care who self-identified as familiar with prediction models. Invitations, which included details about the study and a link to the survey, were sent via e-mail to the study team's clinician network, the Netherlands Cancer Institute's breast cancer working group, the Dutch Society for Radiation Oncology, the Dutch Society for Surgical Oncology, the Hereditary Breast and Ovarian Cancer Research Netherlands, and the Netherlands Comprehensive Cancer Organization.

Data analysis. Surveys were included in the analysis if they were at least 50% complete (i.e., the respondents had rated at least 1 cluster of factors). Descriptive statistics characterized the overall sample and subgroups of model users and non-users. Due to the imbalance between the groups, with only 9 non-users (6%) compared with 137 users (94%), regression analysis was unsuitable, as the limited number of non-users fell below the recommended events per variable threshold for reliable estimation. Although weighting adjustments were considered to address the imbalance, the limited data for non-users made this approach impractical and potentially biased. Consequently, subsequent analyses focused on the more robust group of model users. Nonparametric tests were used to compare score distributions across sociodemographic subgroups: the Mann–Whitney U test for pairwise comparisons between 2 groups and the Kruskal–Wallis test for comparisons across multiple groups. A 10-year cutoff was applied for years of experience, as this is a common threshold in clinical practice that distinguishes between early- and mid-career clinicians, potentially reflecting differences in familiarity with and use of clinical prediction models. A 50-year cutoff was used for age based on established epidemiological significance. To account for multiple testing and reduce the risk of type I errors, a strict significance threshold ($P \leq 0.002$, using Bonferroni correction) was used. Data distributions and subgroup comparisons were visualized using density plots and scatterplots. We

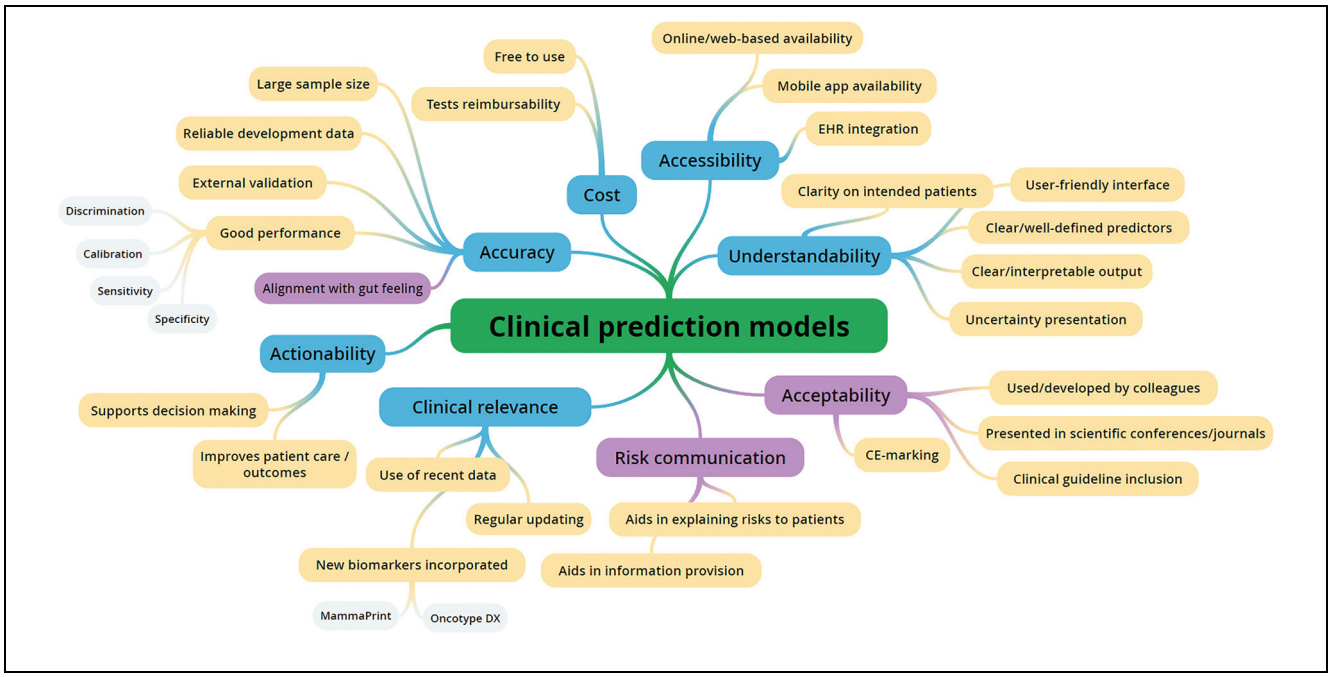

Figure 2 Factors influencing clinicians' decisions to use clinical prediction models.
The findings are based on interviews with 16 Dutch breast cancer clinicians. Key factors are represented in the first-level nodes, with the disaggregated subfactors shown in subsequent nodes. Blue nodes represent practical/methodological factors, whereas purple nodes indicate more perceptual factors.

performed all quantitative analyses using RStudio version 2023.06.0.[22]

Results

Phase 1: Qualitative Semi-structured Interviews

We interviewed 16 breast cancer clinicians: 6 medical oncologists, 3 radiation oncologists, 2 clinical geneticists, 2 surgical oncologists, 2 nurse specialists, and 1 radiologist. Ten clinicians worked primarily in academic medical centers, while others were from general teaching hospitals ($n = 3$) and general hospitals ($n = 3$). Most participants were from Western Netherlands ($n = 13$), 2 were from the South, and 1 was from the North. Most ($n = 13$) were female. The median age was 48 years (range, 31–62 years), and the median number of years of experience in breast cancer care, excluding training, was 14 years (range, 4–35 years).

Key Factors for Model Use

Eight factors (Figure 2) primarily influenced breast cancer clinicians' decisions to initiate the use of clinical prediction models. Practical/methodological factors included accessibility, cost, understandability, objective accuracy, actionability, and clinical relevance. Perceptual (personal or social) factors included subjective accuracy, acceptability, and risk communication. Below is a general summary of these factors. Illustrative quotes are available in Table 1.

- **Accessibility and cost:** Free, online accessible models were preferred. Opinions on integrating models directly into electronic health records (EHR) were mixed. Some saw it as a way to streamline risk assessment through automatic data extraction, while others raised concerns about data accuracy, potential patient misinterpretation, and redundancy with existing access methods.
- **Understandability:** Many clinicians emphasized the importance of clarity regarding the intended patient population, having tools with user-friendly interfaces, understanding the predictors that were used, and the model output. Most preferred models with graphical or tabular output presentations, while some favored outputs that conveyed uncertainty (e.g., via confidence intervals). Few mentioned that they needed to

Table 1 Illustrative Quotes from the Interviews

Factor	Quote
Accessibility	"Only via internet. I bookmark them mostly online. Some of them are on apps." (Clinician 16)
	"Through internet, Google." (Clinician 9)
	"It [EHR integration] would be nice to have, but not a must-have. I find it really easy to just open my browser and get it within seconds." (Clinician 4)
	"It [EHR integration] would be useful, but I wouldn't find it desirable. Because it would be kind of ground truth if you put them in the EHR." (Clinician 16)
	"I prefer to do it manually [rather than having a model integrated in EHR] because I can play with it." (Clinician 5)
	"Many users might not be cautious with interpretation of the numbers." (Clinician 7, referring to the availability of model results in EHR)
Cost	"All of these algorithms are mostly freely available." (Clinician 1)
	"It [PREDICT] is free to access." (Clinician 6)
	"[MammaPrint] is not completely free, but at least we have an arrangement with Agendia so that we can still use it. It's quite difficult for hospitals who have a lot of patients who might be eligible to do the test. I think (most) hospitals can't afford it because they have to pay it by themselves, and I think that makes it quite difficult (to use it)." (Clinician 4)
	"It [MammaPrint] is not covered anymore, so now we are even more restricted in using it." (Clinician 6, commenting on lack of reimbursement during our data collection)
Understandability	"I really use the quite commonly used prediction tools. If it's used by all oncologists, internists, or many doctors, then I'm not going to have a look into the statistics [algorithm] behind it." (Clinician 15)
	"The populations of which all those models are based are probably not completely representative for the patient at stake." (Clinician 16, expressing concern about using models when the intended patient population does not align with their own patients)
	"I think it [PREDICT] is very user-friendly." (Clinician 6)
	"I really like this [PREDICT's tabular output presentation] because it also gives you the confidence intervals and is numerical." (Clinician 1)
	"Sometimes you can look at the median survival and the survival benefit in terms of months or years. And that is sometimes very nice." (Clinician 7, referring to PREDICT's graphical presentation of survival over time)
	"There is a fraction of insecurity around those estimate, that would be normally sort of 95% confidence interval, for individual prediction, you would like to see something like that." (Clinician 3)
Accuracy	"In general I think you want to have a good sense about the accuracy of the prediction. Whether that's done with a C-index or an area under the curve or the difference between predicted and observed in subgroups, I think they are all sort of measuring the same aspect and you want to have feeling about that." (Clinician 4)
	"I really use the quite common-used prediction tools. So, PREDICT, it's used by many doctors, so I'm not going to have a look really into the statistics [algorithm] behind it. But when a tool is not validated for patients under twenty years, I'm not going to use it." (Clinician 15)
	"(I look at) the amount of data that is put into it [a model] and how the data was gathered. . . . For example, I think the modelling with cancer registry data, it should be taken with caution, because of the quality of the data." (Clinician 16)
	"I look at what kind of information they rely on and the validity of the model. I try to make sure that it's valid for my patients . . . [PREDICT] doesn't take into account comorbidities. For an average woman it's good, but for someone with more comorbidity it doesn't work. (So) we made a new model." (Clinician 7)
	"My intuition says that it [PREDICT] actually looks a little bit at the bright side of life. Sometimes, I don't feel that it does justice to the type of tumor. I have a hard time working with it. All the time it only shows you only very good results with patients." (Clinician 11)
Actionability	"[I use models] to actually be capable of tell the patients that their screening should go in this direction or maybe they should indeed visit a clinical geneticist rather than just stay here for regular control." (Clinician 3)
	"I already know that they don't have a very large advantage of their endocrine treatment. If they did experience side effects and they are in doubt whether to continue or to stop, then I use PREDICT to support the patients to stop. I show the numbers of two percent benefit in ten year survival, and then the decision is easily made." (Clinician 12)
	"This [PREDICT's graphical presentation of survival over time] is an incomprehensible curve for an old lady, because it only shows that she's going to die sooner or later, so I won't show it to them. But it helps me count back to the median overall survival gain. And that is sometimes a very useful number to decide on treatments." (Clinician 7)

(continued)

Table 1 (continued)

Factor	Quote
Clinical relevance	"What I really prefer is to have it developed on data as recent as possible and preferably Dutch-validated, or developed in a Dutch situation." (Clinician 4)
	"I think that with breast cancer, we are always a bit behind with data. Since breast cancer has such a long survival, you always need ten years follow-up at least. That makes all the data you're looking at now, data from patients that were treated at least ten years ago. And we know that the recurrences are a lot lower for patients that we are treating now. So we'll always take those numbers with a grain of salt." (Clinician 9)
	"I explained the boost dose [model] in radiotherapy, that's really (using) very old data, so you get a completely overestimate of the risk of recurrence . . . I think that it's really important to update existing models, which have been proven to be good, but they need to be updated with current new insights." (Clinician 2)
	"There are variables missing. I would like to have added the results of gene expression tools, so MammaPrint or Oncotype DX." (Clinician 12)
Risk communication	"You can use the models to give the patients an impression of what is the consequence of not having the radiation or not having the extra radiation, to provide extra background information for the patient, to support your decision to give a certain type of treatment." (Clinician 2)
	"I often propose: would it help you if I show you some graphs and figures about your risk and the benefit of adjuvant treatment? Many of them say, that would be nice. Then I look through the PREDICT tool. " (Clinician 7)
	"It always depends on who's in front of you. I think the puppets [icon arrays] is most illustrative (for patients) . . . I tell them, we have to treat 19 patients in vain to save the life of the 20th patient in the group with this therapy, and we don't know whether you'll be one of the 19. . . . For some people this is really a big risk if they would forgo this therapy, while for others, (they think) it's very likely that they will be one of the 19, so they can easily forgo it. . . . I think interpreting risk is so difficult for people." (Clinician 1)
	"It's really quite easy to explain it [PREDICT survival estimates] to patients. I always find the table to be the most enlightening, because the curves and the chart are a little less clear for patients, I think." (Clinician 6)
	"I have to be honest that I mostly use the table presentation. That makes the most sense to me, to make it clear to a patient. Just the numbers instead of diagrams." (Clinician 13)
Acceptability	"The way I start doing the things I do, is by talking to my colleagues and attending congresses. And if it's used in studies and I see the validation of the studies and it's getting published, I think, wow, it does have a role." (Clinician 12)
	"If you start working here, one of the first things that you learn about is the PREDICT website, because that's the one that we used most frequently." (Clinician 6)
	"A statistician who used to work here, he developed the algorithm for it." (Clinician 2)
	"It's important to put them [models] in the clinical guideline so that everybody uses them." (Clinician 7)
	"We're trying to follow the national (clinical) guidelines, but the thing is not all studies come in, and you have to wait for five years to have a guideline revised again. So, often, you already catch up, by time it's authorized one version, you already know there has been development over the last two years." (Clinician 14)
	"The good thing is that it was a clinical validated so it has an EU stamp [CE marking]. I think two years ago now. That's the reason why we can now use it in clinic as well." (Clinician 14)
	"I can't use it because it doesn't have CE marking for now, but we will start using it as soon as it's possible." (Clinician 7)

understand the algorithm or statistics underlying the model.

- **Accuracy:** A clear distinction was made between *objective* accuracy (measurable and quantifiable correctness, supported by evidence, e.g., from validation studies) and *subjective* accuracy (personal perception based on clinicians' expertise and intuition or "gut feeling"). Clinicians valued models developed from large, reliable datasets. While many mentioned technical metrics used in model validation, specifically discrimination and calibration, they had no strong preference for specific statistical metrics nor did they define thresholds for adequate model performance.

- **Actionability:** Clinicians emphasized the need for models to deliver clear, actionable information that guides treatment decisions and enhances patient compliance, ultimately leading to improved patient outcomes.
- **Clinical relevance:** Clinicians favored models aligned with contemporary clinical practice, especially those developed using recent data, incorporating novel insights (e.g., new genetic markers), and regularly updated to reflect advancements in the field.
- **Acceptability:** Social factors, such as peer use, conference presentations, and publications in high-impact journals, also influenced model use. Opinions on the necessity of clinical guideline recommendations varied: some considered them less critical due to being slightly behind current scientific publications; others emphasized their importance for model acceptance, especially in non-academic hospitals. Few clinicians mentioned CE marking as a requirement for model use.
- **Risk communication:** Clinicians preferred models that helped communicate risks and explain information (e.g., treatment benefits) to patients. They noted that simple icon arrays and tables were most effective when conveying risk information to patients.

Phase 2: Quantitative Online Survey

We included data from 146 clinicians in the analysis. Of these, 137 (94%) reported using clinical prediction models at least occasionally (Supplementary File C), while 9 (6%) indicated they never used such models. Table 2 summarizes the demographic and professional characteristics overall and by subgroup.

Model Users

Among the 137 clinicians who reported using models, 7 (5%) seldom, 39 (28%) sometimes, 69 (50%) regularly, and 22 (16%) always used models. The median age was 49 years (range, 28–64 years), and most (n = 106, 77%) were women. Most (n = 100, 73%) had more than 10 years of experience in breast cancer care. The sample comprised medical oncologists (n = 49, 36%), surgical oncologists (n = 24, 18%), radiation oncologists (n = 32, 23%), nurse specialists (n = 26, 19%), clinical geneticists (n = 4, 3%) and radiologists (n = 2, 1%). Clinicians worked across general hospitals (n = 31, 23%), general teaching hospitals (n = 49, 36%), academic medical centers (n = 47, 34%), and radiotherapy centers (n = 10, 7%). Table 3 lists the models with which the respondents were familiar. The PREDICT model was the most recognized (n = 127, 93%), followed by the Memorial Sloan Kettering Cancer Center lymph node metastases nomograms (n = 55, 40%) and the INFLUENCE model (n = 53, 39%).

Relative Importance of the Different Factors for Model Use

Figure 3 shows the density distributions of scores (see Supplementary File D for descriptive statistics). The highest-rated factors were accessibility (availability as online tool: median = 9 [interquartile range {IQR} = 8–10]) and cost (free to use: 9 [8–10]; tests included reimbursable: 8 [8–10]). Other factors with strong ratings included actionability (good clinical utility: 9 [8–10]) and understandability (clarity in intended patients: 9 [8–9]; clearly defined predictors: 8 [8–9]; user-friendly interface: 8 [8–9]; clear output for clinicians: 8 [8–9]; clear output for patients: 8 [7–9]). Broad methodological components of accuracy were highly rated (reliable development data: 9 [8–10]; large development sample: 9 [8–9]; external validation: 8 [7–9]), along with performance metrics (good discrimination: 8 [7–8]; good calibration: 8 [7–8], good sensitivity: 8 [8–8]; good specificity: 8 [8–9]). Subjective accuracy (model estimates align with gut feeling: 7 [5–8]) received slightly lower scores among accuracy-related factors. Acceptability (used by colleagues: 8 [8–9]; included in clinical guidelines: 8 [7–8]) and risk communication (aids in information provision: 8 [8–9]; aids in risk communication: 8 [8–9]) were also highly rated. The lowest-rated factors were integration into EHRs (6 [3–8]), inclusion of new biomarkers (6 [5–8])—specifically Oncotype DX (6 [4–8]) and MammaPrint (6 [4–7])—and mobile application accessibility (5 [2–7]).

Assessment of Importance Scores across Sociodemographic Characteristics of Model Users

Table 4 summarizes significant differences by sociodemographic characteristics (complete test results in Supplementary Files E–J). Male clinicians rated clear output for clinicians and user-friendly interface as slightly more important compared with female clinicians. Clinicians with more than 10 years of experience placed higher importance on models with clearly defined predictors and app accessibility. Differences were also observed across clinical specializations, such as regarding the importance of CE marking, reimbursement for included tests, and clarity of patient-facing outputs. In addition,

Table 2 Characteristics of the Survey Participants

		Model Use[a]	
	Total	**No**	**Yes**
	(N = 146)	**(n = 9)**	**(n = 137)**
Age, years			
Mean (s)	49.3 (8.85)	51.9 (10.7)	49.2 (8.74)
Median [Min, Max]	49.0 [28.0, 64.0]	54.0 [34.0, 64.0]	49.0 [28.0, 64.0]
Sex			
Male	34 (23.3%)	3 (33.3%)	31 (22.6%)
Female	112 (76.7%)	6 (66.7%)	106 (77.4%)
Years of experience with breast cancer care			
1–2	4 (2.7%)	0 (0%)	4 (2.9%)
3–5	13 (8.9%)	1 (11.1%)	12 (8.8%)
6–10	22 (15.1%)	1 (11.1%)	21 (15.3%)
>10	107 (73.3%)	7 (77.8%)	100 (73.0%)
Clinical specialization			
Surgical oncologist (in training)	24 (16.4%)	0 (0%)	24 (17.5%)
Medical oncologist (in training)	50 (34.2%)	1 (11.1%)	49 (35.8%)
Radiation oncologist (in training)	32 (21.9%)	0 (0%)	32 (23.4%)
Radiologist (in training)	6 (4.1%)	4 (44.4%)	2 (1.5%)
Clinical geneticist (in training)	4 (2.7%)	0 (0%)	4 (2.9%)
Nurse specialist/nurse (in training)	30 (20.5%)	4 (44.4%)	26 (19.0%)
Hospital type			
General hospital	37 (25.3%)	6 (66.7%)	31 (22.6%)
General teaching hospital	52 (35.6%)	3 (33.3%)	49 (35.8%)
Academic medical center	47 (32.2%)	0 (0%)	47 (34.3%)
Radiotherapy center	10 (6.8%)	0 (0%)	10 (7.3%)
Region in the Netherlands			
North	17 (11.6%)	4 (44.4%)	13 (9.5%)
East	25 (17.1%)	3 (33.3%)	22 (16.1%)
West	88 (60.3%)	2 (22.2%)	86 (62.8%)
South	16 (11.0%)	0 (0%)	16 (11.7%)

[a]Model use was defined as "yes" if clinicians reported using clinical prediction models occasionally, regularly, or always and "no" if they reported never using them.

clinicians from general hospitals rated test reimbursement and inclusion of new biomarkers higher than those from academic medical centers or radiotherapy centers.

Non-users

Among the 9 who did not use clinical prediction models, 6 (67%) were affiliated with general hospitals. PREDICT was the most recognized model (Table 3). Radiologists found prediction models less relevant to their work, with some considering them redundant given the comprehensive clinical guidelines they follow. One medical oncologist cited frequent discrepancies between patients' interpretations and their own, which they felt made models inefficient for routine use.

Other Factors for Model Use

Ninety-five (65%) participants reported no additional criteria beyond those covered in the survey. The remaining participants mentioned factors consistent with those included in the survey. No new factors were thus identified.

Discussion

We employed a mixed-methods approach to identify factors influencing breast cancer clinicians' decisions to initiate use of clinical prediction models and to assess the relative importance of these factors. Our study identified eight key factors, broadly grouped into practical/methodological (accessibility, cost, understandability, objective

Table 3 Breast Cancer Prediction Model Familiarity among Survey Participants

	Model Use, *n* (%)	
	No	Yes
	(*n* = 9 [6])	(*n* = 137 [94])
PREDICT Breast Cancer Prognostication Tool	9 (100)	127 (93)
MSKCC Lymph Node Metastasis Nomograms[a]	2 (22)	55 (40)
INFLUENCE Model	1 (11)	53 (39)
IBTR! (Ipsilateral Breast Tumor Recurrence) Model	0	36 (26)
IBR (Ipsilateral Breast Relapse) Model	0	33 (24)
Model to predict the outcome of Oncotype DX	2 (22)	31 (23)
BOADICEA (CanRisk) Model	1 (11)	20 (15)
PORTRET Tool (Prognosis for Women 70 +)	0	11 (8)
CPS + EG (MD Anderson Recurrence Risk Model)[b]	0	10 (7)
Gail Model	1 (11)	7 (5)
MSKCC DCIS Recurrence Nomogram[c]	0	7 (5)
IHC4 Model (Postmenopausal Recurrence Risk)[d]	0	6 (4)
IBIS/Tyrer-Cuzick model	0	5 (4)
Fibrosis Prediction Model (Boost v. No Boost)	0	4 (3)
BCSC Model (Breast Cancer Screening Consortium)	1 (11)	2 (1)
Other models[e]	1 (11)	11 (8)

[a]Memorial Sloan Kettering Cancer Center (MSKCC) Breast Cancer Nomogram for Sentinel Lymph Node and/or Additional Nodal Metastasis.
[b]MD Anderson Model for Predicting the Risk of Recurrence after Neoadjuvant Therapy.
[c]Memorial Sloan Kettering Cancer Center Breast Cancer Nomogram for Recurrence of Ductal Carcinoma In Situ (DCIS).
[d]IHC4 Model for Predicting the Risk of Recurrence in Postmenopausal Women after Hormone Therapy.
[e]Including but not limited to other MD Anderson Breast Cancer Clinical Calculators, the CTS5 Model, the Age Gap Decision Tool, the Breast Reconstruction Risk Assessment (BRA) Score, Mayo Clinic's Prognostic Calculator for Triple Negative Breast Cancer, and the Stanford Decision Tool for Women with BRCA Mutations.

accuracy, actionability, clinical relevance) and perceptual factors (subjective accuracy, risk communication, acceptability). Sociodemographic characteristics showed potential influence on the perceived importance of these factors.

Our findings indicate that practical factors, particularly accessibility and cost, were the most influential when deciding to use clinical prediction models. Freely accessible online models were strongly preferred, highlighting the importance of digital availability. Nonreimbursable tests were identified as a barrier to model use. While clinicians valued incorporating new markers, such as gene signatures,[23] they emphasized the need to balance their use with out-of-pocket costs, particularly when the added predictive benefit is uncertain.[24–26] Notably, accessibility and cost were rated slightly higher than validation metrics, such as discrimination and calibration.[27] The stronger emphasis on practical considerations over technical metrics may stem from clinicians viewing model validation mainly as the responsibility of other stakeholders (e.g., statisticians and leading medical experts), while their primary concern is the effective integration of already validated tools and assessing their utility within clinical workflows.[28] Based on these findings, targeted measures could increase model adoption, such as ensuring health insurance coverage for expensive tests.[29–31] In addition, incentivizing developers or funding infrastructure to provide free online access to models could further facilitate their routine use in breast cancer care.[12,32]

Another key criterion prioritized by clinicians was understandability. Our findings indicate that most clinicians valued model explainability—specifically, understanding model inputs (e.g., included variables/tests) and outputs (e.g., results presentation)—more than algorithmic transparency.[33–37] While algorithmic transparency is often promoted to reduce biases and improve health equity,[38] our results suggest that only a few clinicians prioritized knowledge of the specific algorithms underlying models. Those who did were typically academically inclined, recognizing that a global model may perform differently in diverse local settings.[38] For most clinicians, understanding the predictors, definitions, cutoffs, and derivation population was sufficient.[39] To increase model adoption, developers should focus on creating intuitive, user-friendly online tools based on their models.[40]

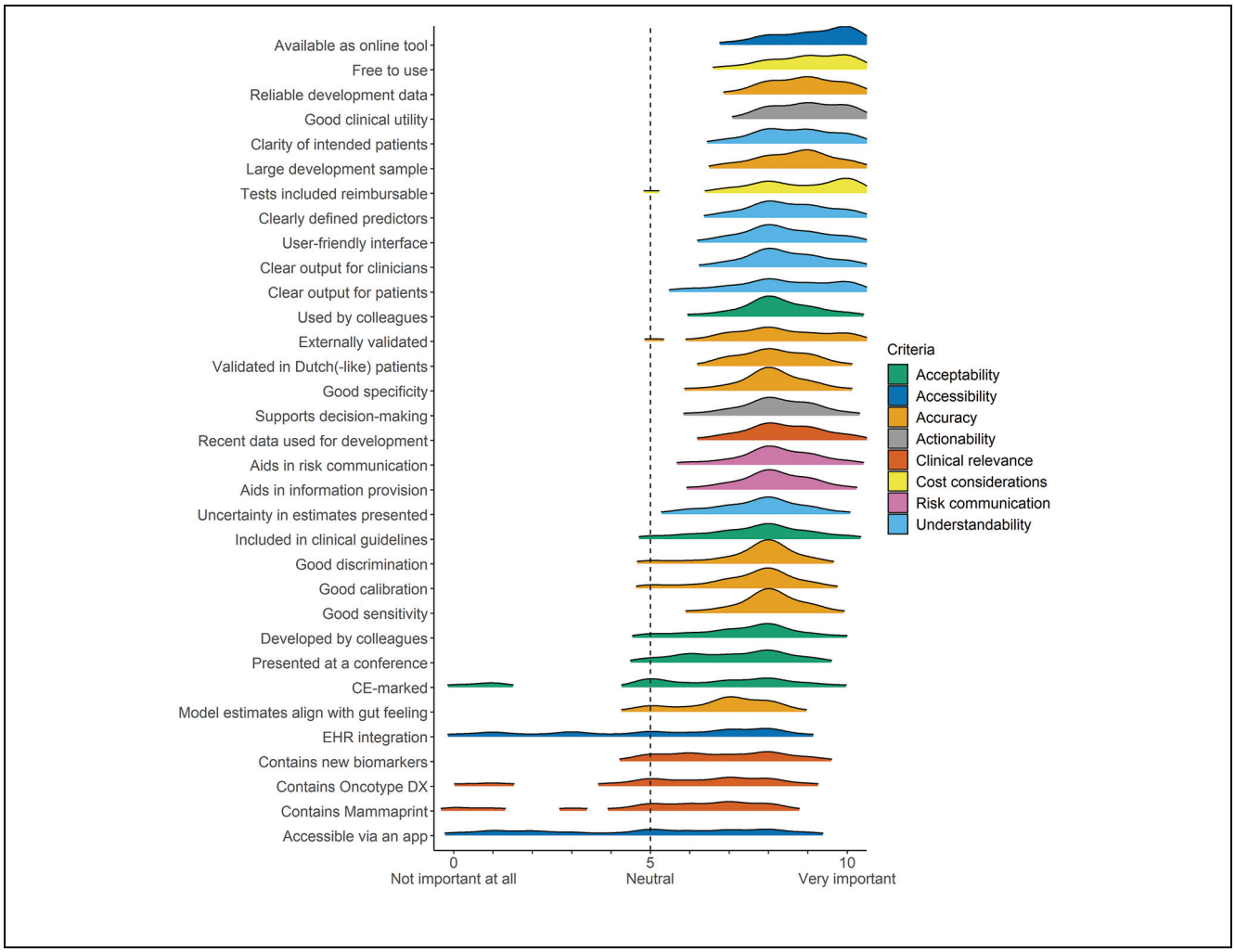

Figure 3 Density plots of the importance scores assigned by breast cancer clinicians who used clinical prediction models (n = 137).
EHR, electronic Health Records; CE, Conformité Européenne (CE).

Detailed information on model specifications or algorithm may be added as supplementary material for those interested.[41]

Trustworthiness, as highlighted in previous research,[37,42,43] is closely linked to model adoption. In our study, we found that trustworthiness was not a standalone factor but was embedded within others, such as understandability (e.g., trust from knowing a model's inputs and understanding its outputs), accuracy (e.g., trust from evidence of validity and alignment with clinical intuition), and peer acceptability (e.g., confidence from endorsements by trusted colleagues). To foster model trust, several strategies could be employed, including conducting validation studies to assess performance across diverse settings, presenting models at relevant clinical conferences, obtaining endorsements from medical societies, and developing (peer-led) training programs to improve model familiarity. Incorporating clinician feedback to adapt models locally for specific patients can also further build trust and address their specific needs.[44] Such a solution might be more straightforward and cost-effective than continuously deriving new models from scratch.[45,46]

In the European Union, prediction models intended to support clinical decision making, such as those used to directly guide diagnosis, prognosis, or treatment selection, are classified as medical devices under the Medical Device Regulation or the In Vitro Diagnostic Regulation if they rely on data from in vitro tests.[47,48] This

Table 4 Significant Subgroup Differences in Importance Scores

Subgroup	Factor	Result (Median [IQR])	*P* Value
Sex	Clear output for clinicians	Male: 9 [8–10] Female: 8 [8–9]	<0.001*
Sex	User-friendly interface	Male: 9 [8–10] Female: 8 [8–9]	0.002*
Years of clinical experience	Accessible via an app	>10 y: 6 [3–8] ≤10 y: 3 [1–5]	<0.001*
Years of clinical experience	Clearly defined predictors	>10 y: 9 [8–9] ≤10 y: 8 [8–8]	0.002*
Clinical specialization	CE marking	Nurse specialists: 8 [7–9] Medical oncologists: 6 [5–8] Surgical oncologists: 6 [1–8] Radiation oncologists: 5 [3–7]	<0.001
Clinical specialization	Clarity of patient-facing outputs	Nurse specialists: 9 [8–10] Surgical oncologists: 9 [8–10] Medical oncologists: 8 [8–9] Radiation oncologists: 7 [6–8]	<0.001
Clinical specialization	Reimbursement for tests	Radiation oncologists: 7 [5–8] Medical oncologists: 9 [8–10] Surgical oncologists: 8 [8–10] Nurse specialists: 8 [8–10]	<0.001
Clinical specialization	Inclusion in clinical guidelines	Medical oncologists: 8 [7–9] Nurse specialists: 8 [7–9] Surgical oncologists: 7 [6–8] Radiation oncologists: 7 [6–8]	0.001
Hospital type	Clarity of patient-facing outputs	General hospitals: 9 [8–10] General teaching hospitals: 8 [8–9] Academic medical centers: 8 [7–9] Radiotherapy centers: 7 [6–9]	<0.001
Hospital type	Inclusion of new biomarkers	General hospitals: 8 [7–9] General teaching hospitals: 6 [5–8] Academic medical centers: 6 [5–8] Radiotherapy centers: 5 [3–6]	<0.001
Hospital type	Reimbursement for tests	General hospitals: 10 [8–10] General teaching hospitals: 9 [8–10] Academic medical centers: 8 [7–8] Radiotherapy centers: 8 [7–9]	<0.001
Hospital type	Inclusion of Oncotype DX	General hospitals: 7 [6–8] General teaching hospitals: 6 [5–8] Academic medical centers: 5 [2–7] Radiotherapy centers: 5 [1–6]	0.002

Subgroup comparisons were conducted using the Mann–Whitney *U* test (*); otherwise, the Kruskal–Wallis test was applied. No significant differences were observed based on clinicians' age (dichotomized at 50 y) or region. Complete test results are provided in Supplementary Files E–J.

classification requires the software or tool based on the model to receive CE certification, which may involve a rigorous, lengthy, and costly process to ensure compliance with harmonized standards (e.g., ISO 13485).[49] Few breast cancer prediction models have managed to obtain this certification.[11] Our findings suggest that most breast cancer clinicians assigned relatively low importance to CE marking. This may be because many already rely on published evidence from validation studies, making CE marking an added formality rather than a crucial indicator of validity or clinical utility.

Our analysis revealed differences in factor prioritization across clinician subgroups. Nurse specialists assigned greater importance to regulatory assurances, such as CE marking, potentially reflecting their emphasis on adherence to protocols and standardized care. In contrast, medical, surgical, and radiation oncologists, who may have greater familiarity with model validation literature,

perceived CE marking as less critical. Clinicians in general hospitals prioritized the inclusion of new markers in models more than those in academic settings did. This difference could potentially be attributed to the fact that the clinical utility of many new markers (e.g., gene signatures, tumor-infiltrating lymphocytes) was still under evaluation at the time of data collection.[31] Clinicians in academic or research settings may have had greater involvement in or access to these evaluations, which could have made them more cautious about incorporating new markers. Understanding the distinct needs of specific clinician subgroups could help further optimize the presentation and features of prediction models.

A key strength of our study is that the identified factors were derived directly from clinicians, ensuring the findings represent the views of end-users of clinical prediction models. In addition, our large sample of model users adds robustness to our main findings. There are also some limitations to consider. The reliance on clinician networks for recruitment may have introduced some selection bias. Next, despite a broad recruitment strategy, certain specializations (e.g., clinical geneticists and radiologists) were underrepresented, which may limit the generalizability to these subgroups. Furthermore, the small number of non-users might have restricted our ability to fully understand the barriers for model use. Self-reported data may have introduced social desirability bias, although we attempted to minimize this by ensuring anonymity and neutral phrasing in survey questions. Finally, to keep the survey brief, only specific sociodemographic variables were examined, which may have limited the depth of our subgroup analysis findings. Future studies should aim to confirm our findings with larger and more diverse samples. It is also important to explore whether similar factors are relevant in other medical domains and countries. Evaluating the effectiveness of different implementation strategies could further inform best practices for model implementation.

Conclusion

Clinical prediction models have the potential to support personalized, evidence-based decision making in breast cancer care. To our knowledge, our study is the first to investigate the broad range of factors influencing breast cancer clinicians' decisions to start using prediction models. Our findings suggest that practical factors, such as ease of access and cost associated with model use, outweigh methodological rigor and technical metrics in determining model use. To bridge the gap between model development and clinical use, improved collaboration among model developers, clinicians, and communication and implementation experts is crucial. Clinician-friendly design that prioritizes usability, utility, and adaptability to local settings could increase model uptake. In addition, advocating for health policy initiatives, such as the reimbursement of expensive tests, providing financial support for translating promising models online, and funding trainings or conference presentations, can help increase familiarity with and confidence in using prediction models, ultimately improving their clinical impact. More research is needed to identify barriers encountered by non-users and to investigate other potential sociodemographic variations, which could allow for a more tailored approach to the implementation of clinical prediction models.

Acknowledgments

The authors would like to thank all of the clinicians who participated in the interviews and online survey as well as the clinicians who helped pilot test the interview guide and questionnaire. We also would like to thank Prof. Dr. Eveline Bleiker for providing feedback on the draft article.

Authors Contributions

EGE, MAB, MKS, and AHB conceived the study and wrote the protocol. All authors revised and approved the protocol. MAB and EGE collected and analyzed the data. All authors helped in interpretation of the results. MAB wrote the first draft of the manuscript. EGE, MKS, AHB, and SCL revised the manuscript. All authors approved the final version for publication.

ORCID iDs

Mary Ann E. Binuya https://orcid.org/0000-0001-8682-8690
Ellen G. Engelhardt https://orcid.org/0000-0003-0016-4972

Supplemental Material

Supplementary material for this article is available online at https://doi.org/10.1177/23814683251328377.

References

1. Harbeck N, Penault-Llorca F, Cortes J, et al. Breast cancer. *Nat Rev Dis Primers*. 2019;5(1):66.
2. Zhao A, Larbi M, Miller K, O'Neill S, Jayasekera J. A scoping review of interactive and personalized Web-based clinical tools to support treatment decision making in breast cancer. *Breast*. 2021;61:43–57.
3. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med*. 2013;10(2):e1001381.

4. Wishart GC, Azzato EM, Greenberg DC, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res*. 2010;12(1):R1.
5. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med*. 2019;21(8):1708–18.
6. Damen JA, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016;353:i2416.
7. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ*. 2020;369:m1328.
8. Phung MT, Tin Tin S, Elwood JM. Prognostic models for breast cancer: a systematic review. *BMC Cancer*. 2019;19(1):230.
9. Dutch Breast Cancer Guideline (Landelijke richtlijn borstkanker). Available from: https://richtlijnendatabase.nl/richtlijn/borstkanker/startpagina_-_borstkanker.html [Accessed 26 September, 2024].
10. Hueting TA, van Maaren MC, Hendriks MP, Koffijberg H, Siesling S. The majority of 922 prediction models supporting breast cancer decision-making are at high risk of bias. *J Clin Epidemiol*. 2022;152:238–47.
11. van Maaren MC, Hueting TA, Völkel V, van Hezewijk M, Strobbe LJ, Siesling S. The use and misuse of risk prediction tools for clinical decision-making. *Breast*. 2023;69:428–30.
12. Abell B, Naicker S, Rodwell D, et al. Identifying barriers and facilitators to successful implementation of computerized clinical decision support systems in hospitals: a NASSS framework-informed scoping review. *Implement Sci*. 2023;18(1):32.
13. Meunier PY, Raynaud C, Guimaraes E, Gueyffier F, Letrilliart L. Barriers and facilitators to the use of clinical decision support systems in primary care: a mixed-methods systematic review. *Ann Fam Med*. 2023;21(1):57–69.
14. Harry ML, Truitt AR, Saman DM, et al. Barriers and facilitators to implementing cancer prevention clinical decision support in primary care: a qualitative study. *BMC Health Serv Res*. 2019;19(1):534.
15. O'Donnell S, Cranney A, Jacobsen MJ, Graham ID, O'Connor AM, Tugwell P. Understanding and overcoming the barriers of implementing patient decision aids in clinical practice. *J Eval Clin Pract*. 2006;12(2):174–81.
16. Watson DB, Thomson RG, Murtagh MJ. Professional centred shared decision making: patient decision aids in practice in primary care. *BMC Health Serv Res*. 2008;8:5.
17. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ*. 2006;333(7565):417.
18. Guler EN. Gene expression profiling in breast cancer and its effect on therapy selection in early-stage breast cancer. *Eur J Breast Health*. 2017;13(4):168–74.
19. Feldman M, Stanford R, Catcheside A, Stotter A. The use of a prognostic table to aid decision making on adjuvant therapy for women with early breast cancer. *Eur J Surg Oncol*. 2002;28(6):615–9.
20. ATLAS.ti. ATLAS.ti (version 9) [Qualitative data analysis software]. ATLAS.ti Scientific Software Development GmbH; 2023.
21. Castor EDC. Castor Electronic Data Capture. 2019. Available from: https://castoredc.com [Accessed 08 November, 2023].
22. R Core Team. *R: A Language and Environment for Statistical Computing (Version 2022.07.2)*. Vienna (Austria): R Development Core Team; 2023.
23. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415(6871):530–6.
24. Andre F, Ismaila N, Allison KH, et al. Biomarkers for adjuvant endocrine and chemotherapy in early-stage breast cancer: ASCO guideline update. *J Clin Oncol*. 2022;40(16):1816–37.
25. Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2019;30(10):1674.
26. NICE. Early and locally advanced breast cancer: diagnosis and management. July 2018 [updated January 2024]. Available from: https://www.nice.org.uk/guidance/ng101/chapter/recommendations#adjuvant-chemotherapy-for-invasive-breast-cancer [Accessed 09 February, 2024].
27. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128–38.
28. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagn Progn Res*. 2018;2:11.
29. Zorginstituut Nederland. Standpunt MammaPrint® bij vrouwen met vroeg stadium borstkanker (herbeoordeling). 2018. Available from: https://www.zorginstituutnederland.nl/publicaties/standpunten/2018/09/27/standpunt-mammaprint-bij-vrouwen-met-vroeg-stadium-borstkanker-herbeoordeling
30. Retèl VP, Byng D, Linn SC, et al. Cost-effectiveness analysis of the 70-gene signature compared with clinical assessment in breast cancer based on a randomised controlled trial. *Eur J Cancer*. 2020;137:193–203.
31. Zorginstituut Nederland. Standpunt - MammaPrint en Oncotype DX vergoede zorg voor bepaalde groep vrouwen. 2023. Available from: https://www.zorginstituutnederland.nl/publicaties/standpunten/2023/10/24/standpunt–mammaprint-en-oncotype-dx-vergoede-zorg-voor-bepaalde-groep-vrouwen [Accessed 16 February, 2024].
32. Giddings R, Joseph A, Callender T, et al. Factors influencing clinician and patient interaction with machine learning-based risk prediction models: a systematic review. *Lancet Digit Health*. 2024;6(2):e131–44.

33. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion*. 2020;58:82–115.
34. Moncada-Torres A, van Maaren MC, Hendriks MP, Siesling S, Geleijnse G. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Sci Rep*. 2021;11(1):6968.
35. El-Sappagh S, Alonso JM, Islam SMR, Sultan AM, Kwak KS. A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Sci Rep*. 2021;11(1):2660.
36. Du Y, Rafferty AR, McAuliffe FM, Wei L, Mooney C. An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus. *Sci Rep*. 2022;12(1):1170.
37. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. In: Finale DV, Jim F, Ken J, et al., eds. *Proceedings of the 4th Machine Learning for Healthcare Conference*. Proceedings of Machine Learning Research; 2019. p 359–80.
38. Paulus JK, Kent DM. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *NPJ Digit Med*. 2020;3:99.
39. Bergquist M, Rolandsson B, Gryska E, et al. Trust and stakeholder perspectives on the implementation of AI tools in clinical radiology. *Eur Radiol*. 2024;34(1):338–47.
40. Sandhu S, Lin AL, Brajer N, et al. Integrating a machine learning system into clinical workflows: qualitative study. *J Med Internet Res*. 2020;22(11):e22421.
41. Schwartz JM, George M, Rossetti SC, et al. Factors influencing clinician trust in predictive clinical decision support systems for in-hospital deterioration: qualitative descriptive study. *JMIR Hum Factors*. 2022;9(2):e33960.
42. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *J Biomed Inform*. 2021;113:103655.
43. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?" Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco (CA): Association for Computing Machinery; 2016. p 1135–44.
44. Jenkins DA, Martin GP, Sperrin M, et al. Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems? *Diagn Progn Res*. 2021;5(1):1.
45. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Med Res Methodol*. 2022;22(1):316.
46. van der Plas-Krijgsman WG, Giardiello D, Putter H, et al. Development and validation of the PORTRET tool to predict recurrence, overall survival, and other-cause mortality in older patients with breast cancer in the Netherlands: a population-based study. *Lancet Healthy Longev*. 2021; 2(11):e704–711.
47. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. 2017. Available from: https://eur-lex.europa.eu/eli/reg/2017/745/oj [Accessed 23 October, 2023].
48. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU. 2017. Available from: https://eur-lex.europa.eu/eli/reg/2017/746/oj [Accessed 23 October, 2023].
49. European Commission. Notified Bodies Survey on certifications and applications (MDR/IVDR). 2024. Available from: https://health.ec.europa.eu/document/download/59b9d90e-be42-4895-9f6f-bec35138bb0a_en?filename=md_nb_survey_certifications_applications_en.pdf [Accessed 10 March, 2024].