

## REVIEW ARTICLE OPEN



## Epidemiology

## The current status of risk-stratified breast screening

Ash Kieran Clift<sup>1,2</sup>, David Dodwell<sup>3</sup>, Simon Lord<sup>4</sup>, Stavros Petrou<sup>2</sup>, Sir Michael Brady<sup>4</sup>, Gary S. Collins<sup>5,6</sup> and Julia Hippisley-Cox<sup>2</sup>

© The Author(s) 2021

Apart from high-risk scenarios such as the presence of highly penetrant genetic mutations, breast screening typically comprises mammography or tomosynthesis strategies defined by age. However, age-based screening ignores the range of breast cancer risks that individual women may possess and is antithetical to the ambitions of personalised early detection. Whilst screening mammography reduces breast cancer mortality, this is at the risk of potentially significant harms including overdiagnosis with overtreatment, and psychological morbidity associated with false positives. In risk-stratified screening, individualised risk assessment may inform screening intensity/interval, starting age, imaging modality used, or even decisions not to screen. However, clear evidence for its benefits and harms needs to be established. In this scoping review, the authors summarise the established and emerging evidence regarding several critical dependencies for successful risk-stratified breast screening: risk prediction model performance, epidemiological studies, retrospective clinical evaluations, health economic evaluations and qualitative research on feasibility and acceptability. Family history, breast density or reproductive factors are not on their own suitable for precisely estimating risk and risk prediction models increasingly incorporate combinations of demographic, clinical, genetic and imaging-related parameters. Clinical evaluations of risk-stratified screening are currently limited. Epidemiological evidence is sparse, and randomised trials only began in recent years.

*British Journal of Cancer* (2022) 126:533–550; <https://doi.org/10.1038/s41416-021-01550-3>

## INTRODUCTION

Breast screening is widely implemented in many healthcare systems to reduce breast cancer mortality through the expedited diagnosis of smaller, asymptomatic breast cancers. For the majority of women, this uses mammography starting in middle age, although different regions differ in their screening recommendations and practices (Table 1). In rarer, very high-risk situations such as a known, high-penetrance genetic predisposition, earlier screening with magnetic resonance or ultrasound imaging is advocated [1, 2].

Whilst meta-analysis of randomised clinical trial data clearly demonstrates a reduction in the relative risk of breast cancer mortality due to screening, reduced breast cancer deaths may come at the expense of overdiagnosis (the identification and unnecessary treatment of clinically insignificant tumours), as well as the consequences of false-positive or false-negative results [3–8]. There is wide variability in the results seen in observational studies in relation to overdiagnosis estimations, depending on analytical or modelling approaches [9–12]. The summation of evidence assessed by the UK Independent Panel [3] was that breast screening does reduce mortality. It concluded that for every 10,000 women in the United Kingdom aged 50 years invited to screening for the next 20 years, 43 breast cancer deaths would be avoided and 681 tumours

(invasive or ductal carcinoma in situ [DCIS]) would be diagnosed, but 129 women would be overdiagnosed, i.e. three overdiagnosed cases per breast cancer death averted, although this calculated benefit–harm balance has been contested by some [13, 14].

Most countries use an age-based population-level breast screening strategy that reduces breast cancer mortality but does not account for the wide variation in individual women's cancer risks [15–18]. Identification of women at the highest risk of breast cancer may enable targeted intensification of early detection or preventive measures, and reduce the public health burden of this common malignancy. Motivated by such factors, 'risk-stratified' breast screening has emerged as a concept in which decisions to offer screening or the determination of screening frequency and modality (e.g. from mammography/tomosynthesis to magnetic resonance imaging [MRI]) are guided by accurate estimation of an individual woman's risk of breast cancer [15, 19–21]. Logically, risk-stratified screening would target intervention at those that stand to gain the most and reduce or stop screening in those that stand to gain little benefit, potentially also informed by the cost-effectiveness evidence.

However, the efficacy and feasibility of personalising screening strategies is uncertain and would require the meeting of several critical dependencies for implementation [21]. These include the need for rigorously developed and validated risk prediction

<sup>1</sup>Cancer Research UK Oxford Centre, Department of Oncology, University of Oxford, Oxford, UK. <sup>2</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK. <sup>3</sup>Nuffield Department of Population Health, University of Oxford, Oxford, UK. <sup>4</sup>Department of Oncology, University of Oxford, Oxford, UK. <sup>5</sup>Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology & Musculoskeletal Sciences, University of Oxford, Oxford, UK. <sup>6</sup>NIHR Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK. ✉email: [ashley.clift@phc.ox.ac.uk](mailto:ashley.clift@phc.ox.ac.uk)

Received: 22 January 2021 Revised: 25 August 2021 Accepted: 14 September 2021  
Published online: 26 October 2021

**Table 1.** Summary of national screening programme strategies or national body recommendations for screening women who are not at elevated risk of breast cancer (e.g. those without a known familial risk/genetic predisposition, or history of chest wall radiotherapy).

Country	Age group	Screening strategy
United Kingdom (NHS Breast Screening Programme) [161]	Women aged 50–70 Women aged 71 years and older	Invitation to mammography screening every 3 years Not invited—may self-refer
United States of America (United States Preventive Service Task Force) [162]	Women aged 40–49 years Women aged 50–74 years Women aged 75 years and older	Individual decision-making recommended Biennial mammography No recommendation: evidence insufficient to assess harms and benefits in this age group
Canada (Canadian Task Force on Preventive Health Care) [163]	Women aged 40–49 years Women aged 50–69 years Women aged 70–74 years	Not recommended; shared decision-making if desired Mammography every 2–3 years Mammography every 2–3 years
Netherlands (National Breast Cancer Screening Programme) [164]	Women aged 50–75 years	Invitation to mammography every 2 years
Australia (BreastScreen Australia) [165]	Women aged 40–49 years Women aged 50–74 years Women aged 74 years and older	Not invited, but may 'opt-in' Invitation to mammography every 2 years Not invited but may 'opt-in'
China (National Health Commission of the People's Republic of China) [166]	Women aged 20–39 years Women aged 40–69 years Women aged 70 years and older	Monthly breast self-examination, clinical breast examination 1–3 yearly Mammography every 1–2 years with ultrasound for women with dense breasts; monthly breast self-examination and annual clinical breast examination Monthly breast self-examination, annual clinical breast examination

models capable of stratifying women accurately, robust health economic evaluation, and clarity on the balance sheet of benefits and harms that would become the 'new norm' in clinical practice after prospective studies. Although some recent reviews have sought to summarise this field [21–23], they differ in the extent to which they cover each of these dependencies. In this narrative scoping review, the authors seek to provide a concise overview of the key topics within the risk-stratified screening literature and anticipate the potential effects of emerging evidence on clinical practice. First, the evidence for risk prediction modelling that may guide personalised screening is reviewed. Thereafter, the evidence from observational analyses of epidemiological or registry data is appraised prior to discussion of how ongoing prospective cohorts and trials have been designed in this area. Furthermore, health economic evidence and output from qualitative studies is synthesised. Throughout, there is an emphasis on evidence quality, its limitations and discussion of how unmet needs may be satisfied.

## METHODS

A scoping literature review was undertaken using Medline (PubMed) with the following search strategy: ("breast screening" OR "mammography") AND ("risk#adapted" OR "risk#stratified" OR "personalised" OR "personalized" OR "tailored" OR "risk#based"). Papers published in any language prior to 1 November 2020 were considered for inclusion in this review. We reviewed the reference lists of systematic reviews to identify key publications if not identified by the search strategy. We also searched clinicaltrials.gov on 1 November 2020 to identify ongoing interventional studies in this area (for search terms: "breast cancer", "screening" and "risk").

Reports retrieved were screened for inclusion based on title and abstract and, if relevant, were classified into five groups: papers

reporting risk prediction models, epidemiological analyses of risk-stratified screening or retrospective evaluations, prospective studies and trials of risk-stratified screening, health economic evaluations and qualitative research on feasibility or acceptability. Findings were synthesised narratively, informed by the narrative synthesis guidelines developed by the Cochrane Collaboration [24].

## Risk prediction models to guide personalised screening

Several risk prediction models for breast cancer incidence have been reported, which tend to incorporate 'classical' risk factors identified from epidemiological evidence, e.g. clinical, demographic or pharmacological exposures, but may also assimilate factors such as family history, genetic risk markers or polygenic risk scores and imaging-related parameters. Well-recognised 'risk factors' include breast density, first-degree family history of breast cancer, increased body mass index (BMI), nulliparity or young age at first birth, and such factors may be attributable for over 52% of risk [25]. The role of obesity relative to menopause, such as the apparent 'protective effect' of obesity on oestrogen receptor-positive cancers prior to menopause [26, 27], may be relevant but is not always incorporated into risk models [28].

Some models have also sought to incorporate markers of genetic risk, such as the inclusion of *BRCA* genotype in the Tyrer–Cuzick (also known as the 'IBIS') model [29], as well as the incorporation of a polygenic risk estimation incorporated into a later update thereof [30]. Some comprise predominantly genetic information, such as BOADICEA [31–33] and BRCAPRO [32, 34]. Genetic predisposition, through either highly penetrant mutations such as *BRCA1/2* [35, 36] or subtler single-nucleotide polymorphisms (SNPs) may indeed affect breast cancer risk; however, only 25–50% of the familial risk can be explained by known genetic variants [35, 37–40], and only 16% of the risk of nonfamilial breast cancer is accounted for by SNPs [36]. Furthermore, breast density

and textural features [41] may be relevant to breast cancer risk and have been explored as covariates in risk prediction models rooted in mammographic image analysis or as additions to covariate panels in updated versions of statistical models. Generally, assessment of breast density could be determined by visual assessment scales [42], or automated/algorithmic approaches [43–45], but the most common appears to be the four-category ‘Breast Imaging Reporting and Data System’ classification [46] (BI-RADS) [30, 47, 48]. Modelling approaches utilised include mathematical modelling [29], statistical regression [49, 50] or ‘machine learning’ techniques [51].

Whilst the implementation of risk-stratified screening to entire populations depends on the prospective outcome and economic evaluations, the requirement for accurate and robust multivariable risk prediction models is the *sine qua non* of any such approach [52]. All risk prediction models intended to be used to inform clinical decision-making should be transparently reported, and robustly evaluated in terms of various performance metrics [53]. Strong internal validation is recommended, and appropriate external validation using data sources that are independent of those used to generate the model may also be useful [54, 55]. Important considerations include the discrimination of models, i.e. how well they distinguish between women who do develop breast cancer versus those who do not [56]; calibration, i.e. the degree of agreement between the predicted risks and the observed risks [57]; and assessment of ‘net benefit’ using decision-curve analysis [58, 59]. In terms of discrimination, the ‘area under the curve’ (AUC) or identical ‘c-statistic’ may be used for binary outcomes, or the ‘c-index’ for survival data [60] both range between 0 and 1, with values of 1 corresponding to perfect prediction and 0.5 reflecting discrimination no better than a coin toss. Other metrics to consider include the proportion of variance explained by the model, such as the  $R^2$  [2, 61]. It is increasingly clear that average performance metrics or ‘overall’ assessments of model performance in populations are insufficient to truly assess clinical utility on deployment, as there may be differences in model performance between regions, ethnic groups or even age groups [62, 63].

Table 2 describes the development and validation results regarding key published risk prediction models. Our search strategy sought to identify original reports and secondary validation studies without restricting the latter to those performed by model developers. Interestingly, a recent systematic review by Louro et al. [64], which appraised the evidence for the Breast Cancer Risk Assessment Tool (BCRAT), Breast Cancer Surveillance Consortium (BCSC), Rosner and Colditz, IBIS and other models using the ISPOR-AMCP-NPC [65] questionnaire rather than PROBAST [66] found that it was challenging to recommend any model for the purposes of risk-stratified screening. Importantly, some risk prediction models were missed by their search strategy [50, 67], and models focussing on genetic determinants of risk were intentionally excluded [64]. The emergent pattern is that incorporation of multiple data forms, such as adding breast density or other mammographic features, or genetic information yields incremental gains in model performance [68], although these tend to be relatively small [30, 69]. Specifically, regarding incorporating breast density on discrimination, the increase in the AUC of published models ranges from 0.03 to 0.14 [69]. The effects on calibration are less clear and the effects on net benefit are not reported.

It is important to note that simply comparing AUC/c-indices of different model development studies does not constitute a meaningful comparison as a breast cancer risk is strongly influenced by age, and the AUC/c-index may be influenced by the heterogeneity of study population, the prediction horizons used and the source of the population used may differ across studies. For example, the AUC/c-index of a model developed in a cohort with a very broad age range is not directly transferable to a separate cohort of women with a narrower age range, such as women eligible for screening currently.

The O/E ratio is widely used to assess calibration, and simply compares the overall number of observed cases versus the number predicted by a model for a given population, and as a standalone metric is insufficient, as over-prediction in sub-groups can be compensated by under-prediction in others and vice versa [57]. A more comprehensive analysis of alignment between predicted versus observed risks for individual study participants could include the use of a calibration plot displaying (mis)alignment across levels of risk [70]. Some papers report as one aspect of their analyses the hazard ratio of pre-selected highest risk groups (such as top tenth) to middle-risk groups (such as middle 80%) [47]. This alone may provide an incomplete assessment of model performance as it is comparing small groups at the extremes of a risk score distribution to the bulk of the study population, which would naturally be expected to diverge in their observed risk. These must therefore be interpreted in the context of the other sources of information regarding model performance when provided. A key external validation study of four key models in a cohort of 15,732 women from Australia, Canada and the US (519 cases of breast cancer) demonstrated c-statistics of 0.70 for BOADICEA (95% confidence interval [CI]: 0.68–0.72), 0.71 for IBIS (95% CI: 0.69–0.73), 0.68 for BRCAPRO (95% CI: 0.65–0.70) and 0.60 for BCRAT (0.58–0.62) [71]. Assessment of calibration was limited to the O/E ratio: BOADICEA 1.05 (95% CI: 0.97–1.14), IBIS 1.03 (95% CI: 0.96–1.12), BRCAPRO 0.68 (95% CI: 0.65–0.70), and BCRAT 0.79 (95% CI: 0.73–0.85).

Recently, there has been increasing interest in ‘machine learning’ prediction modelling for healthcare. Whilst arguably perceived to be more flexible (e.g. better at capturing non-linear, complex interactions), less reliant on assumptions than traditional regression and capable of handling some forms of data that regression models cannot, machine learning has not been shown to be inherently better than traditional statistical modelling approaches [72]. Datasets used for ML modelling should capture clinical reality, i.e. reflect the target population, and the architecture of any algorithm should be reported, given their structural flexibility. Clarity of reporting model development can be problematic [73] and validation/performance assessment approaches may not always be appropriate or transparent, especially when comparing different approaches. One recent study compared the performance of the BOADICEA model with a Markov chain Monte Carlo generalised linear mixed model, an adaptive boosting model and a random forest model developed using data from a single oncogenetic institution in Switzerland that focusses on counselling and testing for hereditary cancer syndromes [51]. Whilst the machine learning models were declared to outperform BOADICEA, no robust evaluation of model calibration was performed, and the effective comparison was an external validation of BOADICEA versus an internal validation of the new model using the data they were derived from. This used cross-validation with a low number of repeats ( $n=20$ ), which presumably was used for hyperparameter tuning as well as performance evaluation (not elaborated in paper), an option that is optimistically biased [74, 75]. Further work in this area of comparing different model-building strategies for predicting risk should focus on more robust, meaningful comparisons.

Overall, a range of clinical prediction models has been developed that could be used to guide risk-stratified screening, some of which are undergoing evaluation in trials of personalised screening. The ability of models to guide risk-stratified screening by predicting incident breast cancer risk in asymptomatic women is uncertain [64], even if integrating clinical, genetic and imaging-derived variables [30, 47, 49, 68]. There is no single accepted benchmark for a given performance metric to render a model suitable for guiding personalised screening, and decisions regarding optimal models should not be made on a single metric. Instead, models need to be robustly assessed in terms of discrimination, calibration and potential clinical utility in the

**Table 2.** Details regarding study data, modelling strategy and performance metrics of notable published risk prediction models for breast cancer or their 'updates' identified during the scoping review.

Study first author and year	Study design and setting	Risk trajectory modelled	Study participants (n) and cases	Risk model covariates/data type combinations	Validation strategy	Discrimination metrics reported (95% confidence interval)	Calibration metrics reported (95% confidence interval)	Examination of performance heterogeneity
'Tyrer-Cuzick' or 'BIS' model and updates or variations								
Tyrer et al. [29]	Construction of the model from first principles/informed by other published data	Diagnosis of breast cancer	N/A	Age, <i>BRCA</i> genotype, family history of breast cancer, (including relationship and ages), menarche, age at first birth, menopausal status, atypical hyperplasia, lobular carcinoma in situ, height, BMI	None	None	None	None
Brentnall et al. [47]	Cohort study (Kaiser Permanente Washington registry), US, women aged 40–73 years	Diagnosis of breast cancer within 19 years	132,139 (2699 breast cancer cases)	Tyrer-Cuzick model Tyrer-Cuzick model + breast density category	Evaluated pre-designed risk calculation mechanism in the whole cohort	Separation of cumulative risk curves for pre-determined risk groups Upper tenth, lower tenth and 'middle' 8 tenths compared with hazard ratios	O/E: 1.02 (0.98–1.06) O/E: 0.98 (0.94–1.02)	O/E of models by three age groups (<50, 50–59, 60+ years)
van Veen et al. [30]	Cohort study (PROCAS), England, women aged 46–73 years	Diagnosis of breast cancer within 10 years	9363 (466 cases, 196 incident)	Tyrer-Cuzick model T-C + mammographic density T-C + mammographic density + polygenic risk score (18 SNPs)	Evaluated pre-designed risk calculation mechanism in the whole cohort	AUC 0.58 (0.52–0.62) AUC 0.64 (0.60–0.68) AUC 0.67 (0.62–0.71)	O/E: 1.50 (1.33–1.70) O/E: 0.98 (0.69–1.28)	Discrimination and calibration in terms of tumour ER status, and by invasive/DCIS tumour type
'Gail' or 'BCRAT' model and updates or variations								
Gail et al. [167]	Case-control study, US, White women aged 50 years and over	Projected breast cancer probabilities within 10, 20 and 30 years of follow-up	5998 (2852 cases)	Age at menarche, age at first live birth, number of previous breast biopsies, number of first-degree relatives with breast cancer	None	None	None	None
Gail et al. [168]	Case-control study ('CARE'), US, African-American women aged 35–64 years	Projected breast cancer probabilities within 10, 20 and 30 years of follow-up	3254 (1607 breast cancer cases)	As for BCRAT, re-fitted in data from African-American women	External validation Women's Health Initiative cohort	C-index 0.555 (0.535–0.575)	O/E: 1.08 (0.97–1.20)	O/E ratio by age groups, age at menarche, number of biopsies, and number of first-degree relatives c-index by age groups
Tice et al. [169]	Cohort study, US, women aged 35 years and over	Diagnosis of invasive breast cancer (no explicit horizon, median follow-up 5.1 years)	81,777 (955 breast cancer cases)	BCRAT BCRAT + BI-RADS breast density category	Apparent performance in study dataset	C-index 0.67 (0.65–0.68) C-index 0.68 (0.66–0.70)	None	None
Zhang et al. [170]	Nested case-control study, US, women aged 34–70 years	Diagnosis of invasive breast cancer within 5 years	11,880 (4006 cases)	BCRAT risk score + polygenic risk score (67 SNPs) + mammographic density + estrone	Internal: cross-validation	All women: AUC 0.65 (0.64–0.66) Described 5-year risk per predicted risk percentile	None	Age group-specific changes in AUC with sequential addition of variables to model

Table 2 continued

Study first author and year	Study design and setting	Risk trajectory modelled	Study participants (n) and cases	Risk model covariates/ data type combinations	Validation strategy	Discrimination metrics reported (95% confidence interval)	Calibration metrics reported (95% confidence interval)	Examination of performance heterogeneity
BCSC model (Breast Cancer Surveillance Consortium)								
Tice et al. [171]	Cohort study, US, women undergoing mammography aged 35 years and over	Diagnosis of invasive breast cancer within 5 and 10 years	1,135,977 (17,908 breast cancer cases)	Age, race/ethnicity, family history of breast cancer, breast biopsy history, benign breast disease, BI-RADS breast density	Update of an earlier model developed using a sample of cohort [49], five-fold cross-validation	AUC 0.665	E/O ratio 1.04 (1.03–1.06)	E/O across 5-year age groups, ethnic groups, levels of predictor parameters
Vachon et al. [48]	Nested case-control study, US women undergoing mammography; two further case-control studies (US and Germany)	Diagnosis of invasive breast cancer within 5 years	1622 (set 1) 1529 (set 2) 879 (set 3) (total 1634 breast cancer cases)	BCSC + PRS (76 loci) + BI-RADS breast density	Evaluation of models formed using datasets 2 and 3 using dataset 1	AUC 0.69 (0.67–0.71)	Hosmer–Lemeshow test of calibration	None
Rosner-Colditz model and updates or variations								
Rosner and Colditz, 1996 [172]	Cohort study, US, Caucasian women aged 30–64 years	Diagnosis of breast cancer (no specific horizon)	89,132 (2249 breast cancer cases)	Age, age at menarche, age at menopause, age at first birth, age at subsequent births	None	None	None	None
Rosner et al. [173]	Cohort study, US, Caucasian women aged 30–64 years	Diagnosis of breast cancer (no specific horizon)	66,145 (1559 breast cancer cases)	Age, age at menarche, age at menopause, age at first birth, age at subsequent births, benign breast disease history, HRT use, first-degree family history of breast cancer, weight, BMI, alcohol intake, oestradiol levels	None (assessed fit of log-incidence model)	AUC 0.645	None	None
Zhang et al. [170]	Nested case-control study, US, women aged 34–70 years	Diagnosis of invasive breast cancer within 5 years	11,880 (4006 cases)	BCRAT risk score + polygenic risk score (67 SNPs) + mammographic density + estrone sulfate + testosterone + prolactin	Internal: cross-validation	All women: AUC 0.678 (0.666–0.690) Described 5-year risk per predicted risk percentile	None	Age group-specific changes in AUC with the sequential addition of variables to model
Other models								
Hippisley-Cox and Coupland 2015 [50]	Cohort study in primary care (England), women aged 25–84 years	Diagnosis of breast cancer within 10 years	3,318,258 (41,315 breast cancer cases)	Age, BMI, deprivation score, ethnicity, alcohol intake, family history of breast cancer, benign breast disease, oral contraceptive use, oestrogen-containing HRT use, manic depression/schizophrenia, previous blood cancer, previous	Split-sample validation	C-index 0.761 (0.758–0.765) D-statistic 1.088 (1.058–1.119)	Calibration plots (by a tenth of predicted risk)	Performance by age groups

Table 2 continued

Study first author and year	Study design and setting	Risk trajectory modelled	Study participants (n) and cases	Risk model covariates/ data type combinations	Validation strategy	Discrimination metrics reported (95% confidence interval)	Calibration metrics reported (95% confidence interval)	Examination of performance heterogeneity
Pal Choudhury et al. (iCARE models) [77]	Cohort study in UK, women aged 35–74 years	Diagnosis of invasive breast cancer within 5 years	UK-based validation: 64,874 (863 breast cancer cases) US-based validation for iCARE-Lit: 47,279 (1008 breast cancer cases)	lung cancer, previous ovarian cancer iCARE-BPC3: age at menopause, age at first birth, parity, age at first birth, height, alcohol intake, first-degree relative with breast cancer, ever/never smoker, BMI, HRT use (type), current HRT use, the interaction between BMI and HTY type iCARE-Lit: age at menopause, age at first birth, OCP use, BMI, height, alcohol intake, benign breast disease history, first-degree family history of breast cancer, the interaction between BMI and HRT use	External validation of models	UK: AUC over 50 s: 0.602 (0.580–0.624) AUC under 50 s: 0.654 (0.621–0.687) AUC over 50 s: 0.622 (0.600–0.645)	Over 50 s: E/O: 1.00 (0.93–1.09) CS: 0.960 (0.680–1.239) CI: 0.001 (–0.004 to 0.005) Over 50 s: E/O 1.13 (1.04–1.22) CS: 0.811 (0.668–0.954) CI: 0.001 (–0.001 to 0.004)	Discrimination and calibration metrics by age groups (<50, 50+ years)
Ming et al. [51]	Single centre (oncogenetic unit), women aged 20–80 years, Switzerland	Lifetime risk of breast cancer (classified into <17, 17–29, 30+ %)	112,587 (4911 breast cancer cases)	ML classification models: Markov chain Monte Carlo generalised linear mixed model, adaptive boosting and random forest Same inputs as BOADICEA	Internal: repeated cross-validation	ML models: AUC 0.843–0.889 BOADICEA: AUC 0.639	Not presented	None
Park et al. (KoBCRAT) [174]	Case-control study, Korea No age limits	5-year and lifetime risk of breast cancer	9248 (4601 breast cancer cases)	Same predictors as Gail model, re-developed in the Korean population	External validation in further Korean cohort studies	AUC 0.61 (0.49–0.72) in one cohort, AUC 0.89 (0.85–0.93 in the other)	E/O: 0.97 (0.67–1.40) in the first cohort, 0.96 (0.70–1.37) in second cohort	AUCs in women aged <50 and 50+ years
Abdolell et al. [175]	Nested case-control study from a population-based screening program in Canada, women aged 40–75 years	Risk of breast cancer diagnosis at mammography	7770 (1882 breast cancer cases)	Percentage mammographic density, breast volume, age, core biopsy history, first-degree family history, number of births, menopausal status, HRT use	Apparent performance in study dataset (no validation)	Imaging features model: AUC 0.597 Imaging features + biopsy history: AUC 0.660 Full model: AUC 0.665	Not presented	None
Wang et al. [176]	Systematic review to pool odds ratios from observational studies to produce a risk score. Evaluated in Chinese women	Risk of breast cancer within 5 years	62,875 (15 cases of breast cancer)	Age at menarche, age at first birth, benign breast disease history, family history of breast cancer, history of breastfeeding, number of terminations of pregnancy	Evaluated pre-designed risk calculation mechanism in the whole cohort	AUC 0.64 (0.50–0.78)	None	None

Table 2 continued

Study first author and year	Study design and setting	Risk trajectory modelled	Study participants (n) and cases	Risk model covariates/ data type combinations	Validation strategy	Discrimination metrics reported (95% confidence interval)	Calibration metrics reported (95% confidence interval)	Examination of performance heterogeneity
Ueda et al. [28]	Case-control study, Japanese women at a single institution	Risk of breast cancer within 10 years, 20 years, and until age 84 years	376 breast cancer patients (no information on the number of controls)	Age at menarche, age at first delivery, family history of breast cancer, BMI (if post-menopausal)	None	None	None	None
Eriksson et al. [93]	Nested case-control study, Swedish women aged 31–79 years undergoing screening	Risk of breast cancer within 2 years	2165 (433 cancer cases)	(Different model weights for pre- and post-menopausal women): BMI, HRT use, family history of breast cancer, breast density, absolute difference in density between breasts, absolute difference in microcalcifications between breasts, an interaction term for % breast density and masses	Cross-validation (number of folds not specified)	AUC 0.71 (0.69–0.73)	None	None
Barlow et al. [177]	Cohort study, US women aged 35–84 years that underwent a previous mammogram in the preceding 5 years	Risk of invasive breast cancer or DCIS following a screening mammogram within 1 year	1,007,600 (11,638 breast cancer cases)	Pre-menopausal women: age, breast density (BI-RADS category), family history of breast cancer, and a prior breast procedure Postmenopausal women: age, breast density, race/ethnicity, family history of breast cancer, prior breast procedure, BMI, natural menopause, HRT, prior false-positive mammogram	Split-sample validation	Pre-menopausal women: AUC 0.629 (0.603–0.656) Post-menopausal women: AUC 0.626 (0.615–0.637)	Hosmer–Lemeshow goodness of fit	None (separate models fitted for pre- and post-menopausal women)

In terms of discrimination metrics, the *c*-statistic or *c*-index is reported as the AUC and *c*-statistic are identical for binary outcomes, whereas *c*-index is specifically for survival/time-to-event data; we have named the metrics using the appropriate nomenclature dictated by the statistical approach used, regardless of the name used in the original publication. It is important to note that the AUC/*c*-indices for each model development paper are not directly comparable, as each study population varies by age, region and setting. The above information is intended as a narrative overview of extant models as they are reported in their respective study populations, and does not constitute a formal comparison of model performance.

*N/A* not applicable, *O/E ratio* observed to expected ratio, *E/O ratio* expected to observed ratio, *BI-RADS* Breast Imaging-Reporting and Data System Classification, *AUC* area under the receiver operating curve, *BMI* body mass index, *HRT* hormone replacement therapy, *ML* machine learning, *ER+* oestrogen receptor positive, *ER-* oestrogen receptor negative, *PR+* progesterone receptor positive, *PR-* progesterone receptor negative, *SVP* single-nucleotide polymorphisms, *CS* calibration slope, *CI* calibration intercept (i.e. calibration in the large).

target populations. Whilst a model with an AUC of 0.5 in its target population cannot be informative, a high AUC in a model development/evaluation study does not guarantee utility in guiding risk-based screening. Poorly calibrated models may cause harm, and those with unstable performance across sub-groups may raise concerns regarding 'algorithmic fairness/bias'. Some of these models are not fully developed using individual-level data, rather, are pre-determined systems of weights that are then applied to a test dataset to assess performance [30, 47]. Crucially, however, weak or non-existent calibration assessment, non-examination of performance heterogeneity, or the lack of consideration of geographical and temporal transportability [55, 76] are notable limitations. The QCancer (Breast) [50], IBIS [47] and iCARE [77] models are some examples wherein exploration of performance heterogeneity is performed according to age groups or other clinically relevant sub-populations (see Table 2). It has also been suggested that in order to minimise harms from overdiagnosis of indolent tumours, modelling the risk of developing lethal breast cancers could be more appropriate than modelling the diagnosis of any breast cancer in order to risk-stratify screening [11]. This is another avenue for further exploration and analysis.

### Epidemiological analyses and retrospective evaluations of risk-stratified screening

The real-world effects of implementing risk-stratified screening strategies [78] warrant evaluation in prospective studies and trials.

Whilst trials have recently been initiated (see later section), there have also been several explorations of the possible benefits and harms of using epidemiological approaches in a jurisdiction where multiple forms of mammography screening are available, namely in Taiwan. This complements studies modelling breast cancer risk in large cohort studies relative to age, or retrospectively simulating the effect of implementing different screening practices in screening cohorts. Table 3 summarises the evidence from such epidemiological papers or retrospective clinical evaluations simulating the possible effects of risk decision rules using patient data.

A large Taiwanese study ( $n > 1.4$  million) exploited the natural experiment of the concurrent availability of three screening approaches in the country's population, namely annual clinical breast examination as the baseline (women aged 35 years and over), risk-stratified biennial mammography screening or universal mammography (both for women aged 50–69 years) [79]. The existence of three available approaches was predicated by a low breast cancer incidence rate in 2002–2004 and concerns about healthcare system capacity for whole population screening at the time, although the rates have increased since [80]. Risk stratification used a 'risk score' derived from reproductive/menstrual history and family history data obtained during attendances for clinical breast examination between 1999 and 2001, with the median of the risk scores used as the cut-off for eligibility for biennial mammography. Using propensity score methodology to try to adjust for disparities in baseline risk factors across the three

**Table 3.** Comparison of studies evaluating risk-stratified screening using simulations on retrospective data, or epidemiological studies.

Study	Country and setting	Description of modelling processes	Key results
Retrospective evaluations using clinical data, or epidemiological modelling			
van den Broek et al. [178]	US Women aged 30–50 years	Breast cancer simulation models: average-risk women, screened according to USPSTF guideline Family history strategy Polygenic breast cancer risk model (313 SNPs) Family history + polygenic risk	<i>Per 1000 women screened, during lifetime:</i> 118 life-years gained, 7 deaths averted, 15 overdiagnoses, 920 false positives 125 life-years gained, 6.9 deaths averted, 14.9 overdiagnoses, 1000 false positives 141 life-years gained, 7.4 breast cancer deaths averted, 16.0 overdiagnoses, 1156 false positives 154 life-years gained, 7.9 deaths averted, 16.6 overdiagnoses, 1169 false positives
Mukama et al. [82]	Sweden, $n = 5,099,172$	10-year cumulative risk curves for breast cancer Analysed risk levels of women with parity and first birth age: risk-adapted starting age of screening based on reproductive profiles	Women with first birth at age <25 years and one child attained the same level of risk as average 50-year-old female at age 51 years; those with parity of 4 or more met this threshold at 59
Mukama et al. [81]	Sweden, $n = 5,099,172$	Modelled age at which women with specific permutations of family history variables attained the risk level of the average 50-year-old woman (age at which screening starts)	If screening would be advised to start at 50 years, women with one first-degree relative diagnosed with breast cancer aged <40 years met the benchmark level of risk at 36 years of age
Lee et al. [88]	US Screening mammograms from 2,647,315 women	Separated women into risk groups based on 5-year age bracket, family history of breast cancer, personal history of breast cancer and dense breasts	Women aged 30–34 years had similar cancer detection rates and recall rates as those aged 40–49 years, suggesting earlier screening in women at higher risk may be appropriate
Burnside et al. [87]	US Screening mammograms from 10,280 women	Cross-sectional study comparing two scenarios: standard age-based screening versus risk-based, defined as having 5-year risk greater than average 50-year-old	Age-based screening diagnosed more cancers than risk-based (68 versus 26%), more false positives (50.3 versus 12.1%)
Prospective epidemiological studies			
Yen et al. [79]	Taiwan, population-based cohort study, $n = 1,429,890$	Cohort study of three screening strategies, adjusting for propensity score for participation: clinical breast examination, risk-based mammography and universal mammography (aged 50–69 years)	No overdiagnosis compared to clinical examination for risk-based screening, versus 13% overdiagnosis with universal screening 41% reduction in breast cancer mortality with universal screening (adjusted for year of birth and propensity score), non-significant reduction with risk-stratified screening



groups (age at menarche, parity, breastfeeding and BMI), Cox models were used, where screening modality was modelled as a time-dependent covariate. Compared to clinical examination, universal biennial mammography had a higher breast cancer detection rate, was associated with a downwards stage migration of detected cancers, a 13% overdiagnosis rate (95% CI: 8–18%), a 30% reduction in stage II + breast cancers (hazard ratio [HR] 0.70, 95% CI: 0.66–0.74), and a 41% reduction in breast cancer mortality (95% CI: 27–52%) when adjusting for propensity score and year of birth [79]. Compared to clinical examination, the overdiagnosis with risk-based screening was negligible (HR for diagnosis 0.97, 95% CI: 0.92–1.03), there was an 8% reduction in stage II + breast cancers (HR 0.92, 95% CI: 0.86–0.99) and a ‘non-significant’ reduction in breast cancer mortality of 14% (HR 0.86, 95% CI: 0.73–1.03) [79]. However, the risk-stratification mechanism in this study was unclear—data were not provided on the modelling methodology used, the risk score covariates, the risk score distribution in the population seeking to opt into risk-based screening (or across the three arms) or performance evaluation to assess if this approach was suitable for clinical use. Further detail is needed to make meaningful inferences on the performance of risk-based screening versus ‘standard’ screening in this analysis. In addition, given the relatively low, albeit increasing, breast cancer incidence rate in this population, risk stratification beyond age and sex may have different proportional benefits in Taiwan in comparison to other nations.

Rather than analysing the effects of altering screening intensity or avoiding screening in low-risk women, a large Swedish cohort with linkage to several national databases ( $n > 5,000,000$  women) was used to assess whether earlier screening starting ages could be appropriate for some women. By using 10-year cumulative risk estimations, the risk level of the ‘average’ 50-year-old woman that would be offered screening was calculated as a benchmark. The ages at which other women would attain the same 10-year risk were compared, based on patterns of family history [81] or personal reproductive history (parity and age at first birth) [82]. Both studies found that either approach could identify women who, despite not being eligible to start age-based screening, had the same 10-year risk estimate as 50-year-old women who would be invited to screen, or indeed may only attain that same threshold of risk after age 50 years. For example, women who had their first baby aged under 25 years met the benchmark aged 51 years, whereas women who had four births by age 25 years met this at 59 years of age [82]. Furthermore, women with one first-degree relative diagnosed with breast cancer before the age of 40 years met the average risk of women starting age-based screening at age 36 years [81]. Therefore, despite debate around the benefits of universally expanding screening to younger age groups such as the lack of long-term effect seen in Age UK [83], selected women with selected risk factors may be suitable for earlier or delayed commencement of early detection strategies. The optimal way to assess risk would require elucidation as reliance on two albeit important risk factors may inadequately capture risk.

In the radiological literature, some commentators have voiced criticism of the potential harms offered by risk-stratified screening [84–86], typically fuelled by retrospective studies applying risk factor-based decision rules to cohorts of women that partook in service screening [86–89]. For example, Lee et al. examined recall rates, cancer detection rates and positive predictive values for biopsy recommendation and the fact of biopsy across age groups, when accounting for breast cancer family history, personal breast cancer history and having dense breasts in a cohort of >2.6 million women [88]. The recall and cancer detection rates in 30–39-year-old women were the same with these risk factors undergoing incidence screening as the 40–49-year-old ‘average-risk’ women undergoing screening; thus, they concluded that such a higher-risk women may benefit from earlier screening starting age. Other institutional studies have expressed concern that risk-stratified

approaches have the potential to miss 75.6–88% of cancers if screening was purely based on family history, 56–86% of cancers if density was the sole determinant or 43.5–76% if access to screening was dictated by positive family history and breast density [86, 89]. Such approaches are not powered to assess the effects of varying screening approaches on stage at detection nor breast cancer mortality, but most crucially, they rely on relatively simplistic determinations of relative risk. Prospective clinical trials are not evaluating such approaches, rather, more nuanced methods for estimating risk.

Overall, currently available epidemiological evaluations or retrospective clinical estimations of breast cancer screening guided by individualised risk are insufficient to inform the utility of risk-stratified screening.

### Prospective cohort studies

Three notable cohort studies are currently exploring multiple aspects pertaining to the feasibility and acceptability of personalised risk assessment in the general population: specifically, the Personalised RiSk-based MAMma screening study (PRISMA), the Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA) and the Predicting the Risk of Cancer at Screening study (PROCAS).

PRISMA is a Dutch collaboration between institutions including Radboud UMC Nijmegen and the North, East, West and South Screening Programmes. In 2014, PRISMA started recruiting asymptomatic women aged 50–75 years in the general population eligible for the national screening programme for data collection via questionnaires, blood and saliva samples and mammograms for assessing breast density. It has a target of 90,000 women with regards to risk factor questionnaire data collection and imaging, and 27,000 blood samples. It aims to not only develop risk prediction models as a fulcrum for investigating risk-based screening strategies but undertake a robust assessment of the acceptability of risk-based screening from ethical, psychological, legal and logistical perspectives. Outputs from PRISMA thus far include multi-cohort qualitative research incorporating individuals from KARMA and PROCAS, which identified a preference for risk-tailored assessment results communication (e.g. letters for below average and average risk, face-to-face appointments for higher risk), the need for standardised risk assessments within national policies and detailed information needs for women in different European countries [90, 91].

The KARMA prospective screening cohort is developing an extensive resource of banked biological, mammogram and lifestyle/clinical factor information from over 70,000 women [92]. Its aims include identification of novel circulating risk markers, genetic risk factors and imaging protocols, assessment of high-throughput breast density measurement, trials of pharmacological prevention therapies such as lower-dose anti-oestrogen therapy and risk communication as well as the development of new risk prediction models with an evaluation of how these could be implemented within screening routines. Notable outputs include the CAD2Y model, which integrates mammographic features such as computer-detected microcalcifications with ‘clinical’ factors for short-term risk prediction [93], risk estimation integrating mammographic density and polygenic risk [48, 94], and contributions to the identification and understanding of genomic breast cancer risk by multi-centre consortia [95–98].

The PROCAS 2 study began in 2015 following PROCAS 1, which recruited over 50,000 women eligible for screening mammography at the Great Manchester NHS Breast Screening Programme. Lifestyle, reproductive history and other clinical informations were collected via questionnaires, with imaging assessment of mammographic density and DNA obtained for polygenic risk analysis [99]. Numerous studies have been undertaken within the remit of PROCAS, such as the evaluation of Tyrer–Cuzick and Gail risk prediction models in a screening population, the predictive

impact of the inclusion of mammographic density [68] and/or polygenic risk score components into these [30, 100], extensive assessment of risk feedback and perception [101–103] and probing associations between ethnicity and mammographic density [99]. PROCAS outputs have been central to supporting the feasibility of population-based breast cancer risk assessment [101, 104], including identifying no major psychological harms of providing 10-year risk estimates from different forms of risk algorithms [103].

### Prospective evaluations and trials of risk-stratified screening

Prospective clinical studies and ideally randomised trials should evaluate risk-based screening practices in terms of outcomes, balance of harms and benefits, cost-effectiveness and acceptability. Several key studies are underway [105–108], including NCT04359420 [108]. This is a non-randomised, counterbalanced study across seven screening sites in the United Kingdom, in which women on an invitation to the NHS Breast Screening Programme will either be offered the standard programme or the additional invitation to use BC-Predict, an automated system for offering a breast cancer risk assessment (to include questionnaires, breast density measurement and polygenic risk) on an invitation to screen [108]. Its aims include assessing risk assessment uptake after offer, uptake of risk consultation, chemoprevention or additional mammography, as well as risks of potential cancer worry, anxiety or health service costs [108]. Many others do not strictly seek to evaluate the outcomes of screening intensity/eligibility decisions based on individualised risk estimation. Indeed, some are also exploring how best to communicate personal risk (e.g. PROSPR/PCIPS 3, NCT01879189), promote breast screening uptake based on risk factor-specific educational materials (e.g. NCT00416975) or identify optimal imaging modalities for women at specific levels of predicted risk (e.g. NCT00003736).

The Women Informed to Screen Depending on Measures of risk (WISDOM) trial is a preference-tolerant randomised trial of a risk-based screening algorithm versus standard screening practice in the United States that started in 2016 [105]. Absolute breast cancer risk estimates are generated using the BCSC model [49], modified by a polygenic risk score incorporating 96 SNPs using Bayesian principles [105] and testing for nine high- or moderate-risk genes (*BRCA1*, *BRCA2*, *TP53*, *STK11*, *PTEN*, *CHD1*, *ATM*, *PALB2* and *CHEK2*). Predicted risks at 5 years and age dictate the screening strategy. For those aged 40–49 years: women with a 5-year risk of <1.3% are not being offered screening, those with a 5-year risk of 1.3% or greater are undergoing biennial mammography, whilst women with extremely dense breasts or who are carriers of *ATM/PALB2/CHEK2* mutations without a positive family history are undergoing annual mammography. For women aged 50–74 years, all are undergoing biennial mammography unless they are carriers of *ATM/PALB2/CHEK2* mutations without a positive family history in which case they receive annual mammography. Regardless of age group, annual mammography with adjunct MRI is being deployed in carriers of *BRCA1/BRCA2/TP53/PTEN/STK11/CDH1* mutations regardless of family history, carriers of *ATM/PALB2/CHEK2* mutations with a positive family history, those who had chest irradiation between the ages of 10–30 years or have a 5-year breast cancer risk of at least 6%. With target recruitment of 100,000 women, it has been projected that 75% of women aged 40–49 years will be allocated to ‘no screening’, whereas 91% of women aged 50–74 years will undergo biennial mammography [105]. The primary endpoints are non-inferiority to standard screening regarding the number of late-stage breast cancers diagnosed (>stage IIB), and rates of recall and breast biopsy. Secondary endpoints include the rate of stage IIB and interval cancers, recall rates, rates of DCIS diagnosis, rates of chemoprevention use, cancer incidence rate, PROMIS anxiety score and rates of systemic therapy use between arms. Importantly, the design is

inherently adaptive so that risk assessment methodology and screening strategies are adjustable in line with future evidence under a ‘continuous improvement’ framework [18].

The population-based Tailored Breast Screening Trial (NCT02619123) was initiated in Italy in 2013, and is randomising pre-menopausal women aged 44 years and older to invitation to ‘tailored screening’ or an active comparator [107]. The target recruitment is 33,200 women and the estimated study completion date will be early 2022. In the tailored arm, those with BI-RADS grade C–D breast density receive annual mammogram invitations until age 50 years and then standard population screening; those with lower density breasts are invited 2-yearly until age 50 years and then standard population screening. In the active comparator arm, women are invited to annual mammography until age 50 years, followed by usual population screening. The primary outcome measures are the difference in cumulative interval cancers between arms (also by density group) and the cumulative incidence of >T2 or node-positive breast cancers by arms (also by density group). Secondary endpoints include a comparison of false-positive rates between arms, the cumulative incidence of all breast cancer cases and attendance to screening. However, the ramifications of this trial on clinical practice may be limited, due to the basis for stratification (dense versus non-dense breasts), the small divergences in screening strategy (annual or biennial mammography) and the short period in which the screening intensity will be altered.

The other key trial is My Personalized Breast Screening (NCT03672331); an international study seeking to recruit 85,000 women aged 40–70 years, in which screening strategy in the experimental arm will be dictated by risk assessment incorporating age, family history, previous benign breast disease, hormone/reproductive history and a polygenic risk score. Specifically, women with one or no first-degree relatives with breast or ovarian cancer will utilise the MammoRisk® model, otherwise, the Tyrer–Cuzick model will be used (see above). No data regarding the proprietary MammoRisk® algorithm structure itself or results of performance evaluation is accessible on the owning company’s website (<https://www.predilife.com/en/home-2/>, accessed 1 November 2020) or identifiable on Medline, although studies of acceptability/ease of software use are available [109, 110]. In the comparator arm, women will be screened with mammography, tomosynthesis or MRI in accordance with extant national guidelines, whereas in the active arm, an estimated 5-year risk will inform mammography and/or tomosynthesis screening every 1 to 4 years (with or without ultrasound depending on breast density). The trial has a non-inferiority design and the primary outcome is the incidence of stage >II breast cancers for the risk-stratified arm. Secondary outcome measures include a superiority analysis regarding the incidence rate of stage >II cancers, rates of false positives and benign biopsies, subject anxiety, health-related quality of life according to the EQ-5D and cumulative breast cancer diagnosis rates.

Overall, three key randomised trials are underway to assess the outcomes associated with different screening intervals, starting age, or imaging modalities based on individualised risk assessment, although there is diversity in the robustness of the risk assessment mechanism. The initial results from these trials are likely to emerge within the next 3 years, and it is notable that the largest and most comprehensive is adaptable [105], in that newer methods of risk assessment or amended screening strategies can be incorporated should novel evidence emerge.

### Health economic evaluations of stratified screening

Economic simulations of risk-stratified breast screening have modelled the clinical outcomes and cost-effectiveness of a range of scenarios in health systems such as the United Kingdom’s NHS Breast Screening Programme [16, 111], the United States [112, 113], Germany [114, 115], the Netherlands [116] and China

[117]. Models have evaluated the stratification of screening intensity based on classical risk factors such as family history, class of breast density, age or relative risk based on polygenic risk scores. Across models, analytical approaches have been compared, e.g. relative numbers of breast cancer deaths avoided, rates of overdiagnosis, incremental costs or incremental quality-adjusted life-years and incremental cost-effectiveness ratios.

Whilst it appears that such modelling does lend support to the general concept of altering screening practices on the basis of risk to strike a more favourable balance of benefits (reduction in breast cancer deaths) and harms (overdiagnosis and unnecessary treatment) or to formulate a more cost-effective approach, a cohesive narrative regarding a particular algorithm-informed strategy is difficult to synthesise. Table 4 summarises the approaches and key results from such studies. These existing studies may diverge from real-world practice in terms of the ascertainment of individual risk (in terms of nuance and approach) or may be limited by incomplete information on risk distributions in the target population.

More simplistic estimations of relative risk on a small number of covariates tend to be used as the stratification mechanism, such as 'positive family history' or breast density category, or use blunter approaches to ascertain 'high-risk' women, such as the relative risk of 2.0 or greater. Multivariable risk assessment seeks to offer more nuanced estimations of risk, which are not recapitulated in many studies. Modelling the implementation of a polygenic risk assessment to inform relative risk may be limited by the relatively small absolute contribution that genetic factors may play in the majority of women that do not have high- or medium-penetrant mutations. As some studies have highlighted, the lack of data on the true distribution of risk groups in the target population as a whole limits their assessment of the impact of risk-based screening practices in a health system.

The approaches to and thresholds used to risk-stratify screening appear to differ between studies. Some seek to find the economically optimal screening interval or starting age for screening in cohorts simulated as having set risk levels, or even have intensified screening with additional imaging modalities as one pathway in their models. Few compare existing age-based methods with risk-stratified approaches in the same target population, and fewer evaluate a more robust view of risk-adapted screening, namely not offering screening to those at the lowest risk. Furthermore, the direct comparator in some studies is 'no screening' rather than an evaluation of transitioning from age-based screening to truly risk-adapted screening.

### Qualitative research

In order to accept screening strategies tailored to individual risk, women need to be able to access and comprehend accurate risk estimations [118–120], and indeed, many women are interested to understand and discuss their risk [90, 101, 102, 121, 122]. Risk communication and risk perception are multifaceted and complex [119], yet it is striking that as few as 10% of women have accurate perceptions of personal risk with an otherwise roughly even split between under- and over-estimators [123, 124]. The variable use of absolute and relative risks can have major influences on screening intentions or in some cases be misleading [125].

A growing body of qualitative research has synthesised evidence from focus groups, semi-structured interviews and other methodologies regarding stakeholder views on the implementation and acceptability of risk-stratified screening [90, 126–136]. Particularly key considerations are those pertaining to communication of risk and risk-based pathways across languages and cultural groups [127], socioeconomic groups and those that have lower engagement with preventive healthcare. Generally, perceptions of risk-based screening appear to be favourable, whether based on genetic risk [133, 137] or other factors. Whilst it felt to be acceptable in principle by many women, the evidence base used

to support these approaches needs clear articulation to secure buy-in [126] and reasons for heterogeneous policies for different groups need to be transparent [126]. Importantly, perceptions of risk stratification as a euphemism for service funding reductions may arise [128], and there should be cognisance of anxiety around self-directed risk assessment if used, such as via websites or apps [138] (albeit, at low levels) [134].

### DISCUSSION

Genetic, pharmacological and environmental or lifestyle factors affect breast cancer risk, with heterogeneity encountered in terms of individual women's risks and the risks posed by individual tumours. As one-size-fits-all approaches are increasingly disparaged in breast cancer treatment, detection and prevention under the auspices of 'personalised healthcare' [21], the use of age as the single precision factor for guiding early detection strategies in women may be over-reductionist [19–21]. Approximately one-quarter of all breast cancers are diagnosed in women ineligible for screening due to age, only around 30% of breast cancers in the United Kingdom are detected by the triennial UK screening programme [139] (43% of all breast cancers in the United States [140]), and current methods may predicate overdiagnosis [3, 5, 8, 11]. In the United States, 73% of breast cancers in biennial screeners are detected by mammography, and 78% of breast cancers are detected by screening in annual screeners [141]. Approaches that tailor screening intensity and personalise modalities to those that stand most to gain and minimise unnecessary interventions in those with little to gain require robust evaluation.

It is relevant to distinguish between the use of genetic, lifestyle and other factors in tools for long-term risk prediction to guide imaging strategies over a woman's 'screening lifetime' and estimating the risk of an underlying cancer being present at the time of screening. Compared to settings such as lung cancer screening, where smoking history has the predominant effect on decision-making and may approach immediate diagnostic relevance, the long-term prediction of breast cancer relies on appropriate multi-factorial assessment using data points that have 'weaker' effects on risk.

Approaches to risk-stratified screening that have been explored include altering screening intensity, screening starting age or screening technology used and therefore a single consensus definition has not emerged, nor on the optimal form risk-stratified screening could take. One concern is that a paradigm withholding access to screening with a sensitive imaging modality in those deemed 'low risk' by a less sensitive predictive model may be inappropriate. As such, there must be judicious consideration of whether a 'screen-only high-risk' model is ethically, economically and clinically appropriate in comparison to other models such as the current effective approach, or those that seek to deliver screening that is 'tailored according to risk', where risk assessment influences the factors mentioned above without removing women from the screening pool. It is not yet clear which if any of the currently available risk models or risk-based strategies deliver on this concept—ongoing trials and other continued progress in optimising individualised risk estimation and screening strategy should begin to deliver clinically informative answers in the coming years [105–107]. Further, individualised risk estimation may not only be useful for stratifying screening strategy to reduce harms but to also reduce the substantial public health burden of breast cancer, potentially through identifying women at previously unknown high risk suitable for prevention therapy (e.g. Anastrozole or Tamoxifen).

Alongside the elucidation of the optimal degree of nuance for breast screening pathways, it would be essential to explore risk communication strategies, and continually monitor such programs' effectiveness both clinically and economically. Another avenue for

**Table 4.** Summary of health economic and outcomes models evaluating risk-stratified breast screening identified during the scoping review.

Study reference	Population modelled	Modelling approach utilised	Risk groups and screening scenarios simulated	Key results
Sankatsing et al. [114]	Netherlands—women without BRCA1/2 mutation Women born in 1974 Time horizon age 40–death	Microsimulation (MISCAN—microsimulation screening analysis; semi-Markov processes)	Risk groups: low, average and high, using 'common risk factors', excluding breast density Simulated: biennial screening aged 50–74 overall; biennial or triennial screening for low-risk women starting 50–60 to 64–74 years; annual or biennial screening for high-risk women starting 40–50 until 74–84 years Assumption of 100% screening attendance	Per 1000 women: <i>Biennial screening to all, 50–74:</i> 206 life-years gained, 16 deaths avoided, 187 false positives, 5 overdiagnosed cases <i>Triennial screening 50–71 for low risk:</i> 134 life-years gained, 10 breast cancer deaths avoided, 102 false positives, 3 overdiagnosed cases <i>Biennial screening 40–74 for high risk:</i> 380 life-years gained, 26 breast cancer deaths avoided, 371 false positives, 7 overdiagnosed cases
Arnold et al. [114]	Germany Women aged 50, followed to age 100 or death	Microsimulation Markov model	Risk factors: family history, personal history of biopsy, breast density Compared annual, biennial and triennial universal screening to risk-adapted strategies based on relative risk (three risk categories) Assumption of 54% adherence	Risk-stratified programmes may be more efficient, depending on mortality reduction or QALYs are strategic focus At 54% adherence, compared with no screening, screening women with relative risk >1 was projected to generate 8.63% mortality reduction, incremental QALYs of 0.023 and incremental costs of 211 Euros per woman (2017 prices).
Sun et al. [117]	China (urban population)	Prior natural history Markov model	High-risk defined: relative risk >2 High-risk women: screened using USS aged 40–44 years with subsequent mammography if indicated; with both modalities if 45–69 years Low-risk women: no screening (diagnosis after symptoms arise) Simulated complete treatment and also 70% treatment after diagnosis	<i>Compared with no screening, risk-adapted approach screening every 3 years, with full treatment:</i> Lifetime costs US\$184 per case (2014 prices), 22.99 QALYs, 0.0127 difference in QALY <i>Compared with no screening, annual screening and full treatment:</i> Lifetime costs US\$335.43 per case (2014 prices), 23.01 QALYs, 0.028 different in QALYs
Pashayan et al. [116]	UK Cohort $n = 364,500$ Women aged 50 years followed up to 85 years	Life-table model	Three cohorts with screening based on risk group: (1) No screening (2) Screen all women aged 50–69 years as per NHS BSP (3) Only women above risk polygenic risk threshold screened every 3 years from age 50–69 years	<i>Compared to no screening, risk-based screening:</i> Overdiagnosis: deaths prevented ratio increased from 0.07 to 0.99 as risk threshold lowered (from 99th to 71st percentile) Minimum ICER at 77th percentile of risk threshold (£11,911 per QALY gained), versus £66,445 when using 99th percentile as risk cut-off (price date not specified) At 32nd percentile of risk, risk-adapted screening generated an incremental cost of £20,066 (price date not specified), 450 more QALYs, and 7 fewer breast cancer deaths
Gray et al. [111]	UK Women eligible for NHS BSP (aged 50–70 years)	Discrete event simulation	Four stratification methods in NHS BSP: (1) Absolute 10-year risk as per of Brentnall (BCR 2015): <3.5% triennial screening, 3.5–8% biennial screening, >8% annual screening (2) Relative 10-year risk: low tertile = triennial, middle tertile = biennial, high tertile = annual	<i>Compared to the current NHS BSP strategy (screening 50–70 years of age ever 3 years):</i> Risk-stratification methods 1 and 2 were deemed cost-effective relative to threshold range of £20,000–30,000 per QALY ICER for method 1 versus UK BSP: £16,689 (2015 prices)

Table 4 continued

Study reference	Population modelled	Modelling approach utilised	Risk groups and screening scenarios simulated	Key results
Trentham-Dietz et al. [112]	US Women aged 50+ Lifetime horizon	Microsimulation models ×3	(3) Supplemental US for women with high breast density (4) Approach 1 plus the supplemental US as in (3)  Examined various combinations of breast density (four categories) and relative risk for factors other than density (1.0, 1.3, 2.0 and 4.0) Settings of annual/biennial/triennial screening for women aged 50–74 years, and also for women 65–74 years Assumed 100% adherence	ICER for method 2 versus UK BSP: £23,924 (2015 prices)  Per 1000 women with fatty breasts/scattered fibroglandular density + RR 1 or 1.3: Biennial screening (50–74): 5.1 deaths averted Triennial screening (50–74): 3.4 death averted Triennial screening (65–74): 4.1 deaths averted Triennial screening for average-risk women with low-density breasts provided favourable balance of harms and benefits and is cost-effective Annual screening for higher risk (RR 2.0 or 4.0) with heterogeneously or very dense breasts has favourable balance of benefits and harms and is cost-effective
Schousboe et al. [113]	US Women aged 40–49, 50–59, 60–69 and 70–79 (initial mammography at 40) Lifetime horizon	Markov cost-utility microsimulation model	Modelled risk based on BI-RADS breast density category, and up to 2 risk factors (family history or previous biopsy) Examined annual, biennial, triennial, 3–4 yearly mammography or no mammography	A range of cost-effective strategies for women of different age groups, breast density and presence of up to 2 risk factors were identified (assuming \$100,000 and \$50,000 cost-effectiveness thresholds), e.g. at a 50,000 cost-effectiveness threshold: BI-RADS B-D, or BI-RADS A + 1–2 risk factors: biennial screening 50–59 years, reassess at age 60 BI-RADS A + 0/1 risk factors: 3–4 yearly mammography 50–59 years, reassess at age 60

BI-RADS Breast Imaging-Reporting and Data system classification, USS ultrasound scan, NHS BSP National Health Service Breast Screening Program (in the United Kingdom), ICER incremental cost-effectiveness ratio, QALY quality-adjusted life year.

future work will be further assessment of stated preferences, such as through discrete choice experiments [142, 143] or contingent valuation studies, and the trade-offs between benefits, risks and costs in making such preferences in a stratified programme.

Concerns have been raised that risk-stratified screening poses the danger that many breast cancers may be 'missed' by not screening women believed to be at low risk [84, 144]. However, many such studies ignore the inherently multivariable nature of the best performing currently available risk models [30, 47, 50], compare current age-based screening against a straw man of blunt risk estimation that is not advocated, and are not powered to identify the effects on important outcomes such as the proportion of late-stage cancer detected, or cancer survival. Stratification may lead to changes in imaging strategy, such as increased use of supplemental MRI, which randomised trial evidence suggests reduces rates of interval cancers in women with extremely dense breast tissue [145]. Recent evidence of supplemental abbreviated MRI in women at average risk with dense breasts and negative digital tomosynthesis results appears to increase the prevalent cancer detection rate (up to 27.4 per 1000 women), but the survival benefits are yet to be quantified [146].

Overdiagnosis and overtreatment have been widely acknowledged in prostate cancer screening for years [147–150], yet a progression towards risk-stratification of screening itself or risk-guided management of detected neoplasms is far more mature than in the breast cancer field [149, 151]. The debate around differing estimates of screening mammography's benefits and harms has become increasingly polarised over recent decades [4, 18, 152–158], with consistent disagreement over the interpretation of decades-old trials, the reliability of specific randomised studies, statistical approaches used to interpret them or epidemiological studies [11, 158–160]. However, whilst the clearly prevailing consensus is that screening mammography reduces breast cancer mortality, whether this can be further improved is a worthy avenue to explore. Recently, trials have been designed to provide evidence to inform this. Should their results be positive, they must be followed by careful consideration of whether the 'new normal' would be acceptable to healthcare systems, policy-makers, clinicians and members of the public.

## DATA AVAILABILITY

No original data were generated for this review. The literature search strategy is provided.

## REFERENCES

- Balmana J, Diez O, Rubio IT, Cardoso F, Group EGW. BRCA in breast cancer: ESMO Clinical Practice Guidelines. *Ann Oncol*. 2011;22 Suppl 6:vi31–34.
- Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. *Nat Rev Dis Prim*. 2019;5:66.
- Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer*. 2013;108:2205–40.
- Gotzsche PC. Mammography screening is harmful and should be abandoned. *J R Soc Med*. 2015;108:341–5.
- Miller AB, Wall C, Baines CJ, Sun P, To T, Narod SA. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. *BMJ*. 2014;348:g366.
- Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med*. 2012;367:1998–2005.
- Kalager M, Loberg M, Bretthauer M, Adami HO. Comparative analysis of breast cancer mortality following mammography screening in Denmark and Norway. *Ann Oncol*. 2014;25:1137–43.
- Loberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res*. 2015;17:63.
- Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med*. 2010;363:1203–10.
- Jorgensen KJ, Gotzsche PC, Kalager M, Zahl PH. Breast Cancer Screening in Denmark: a cohort study of tumor size and overdiagnosis. *Ann Intern Med*. 2017;166:313–23.
- Autier P, Boniol M. Mammography screening: a major issue in medicine. *Eur J Cancer*. 2018;90:34–62.
- Paci E, Group EW. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. *J Med Screen*. 2012;19 Suppl 1:5–13.
- Njor SH, Garne JP, Lynge E. Over-diagnosis estimate from The Independent UK Panel on Breast Cancer Screening is based on unsuitable data. *J Med Screen*. 2013;20:104–5.
- Baker SG, Prorok PC. Breast cancer overdiagnosis in stop-screen trials: more uncertainty than previously reported. *J Med Screen*. 2020. <https://doi.org/10.1177/0969141320950784>.
- Baum M. Should routine screening by mammography be replaced by a more selective service of risk assessment/risk management? *Women's Health*. 2010;6:71–6.
- Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: a Life-Table Model. *JAMA Oncol*. 2018;4:1504–10.
- Onega T, Beaber EF, Sprague BL, Barlow WE, Haas JS, Tosteson AN, et al. Breast cancer screening in an era of personalized regimens: a conceptual model and National Cancer Institute initiative for risk-based and preference-based approaches at a population level. *Cancer*. 2014;120:2955–64.
- Esserman LJ, Study W, Athena I. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer*. 2017;3:34.
- Bitencourt AG, Rossi Saccarelli C, Kuhl C, Morris EA. Breast cancer screening in average-risk women: towards personalized screening. *Br J Radiol*. 2019;92:20190660.
- Kerlikowske K, O'Kane ME, Esserman LJ. Fifty years of age-based screening: time for a new risk-based screening approach. *Evid Based Med*. 2014;19:183.
- Pashayan N, Antoniou AC, Ivanus U, Esserman LJ, Easton DF, French D, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. *Nat Rev Clin Oncol*. 2020;17:687–705.
- Harkness EF, Astley SM, Evans DG. Risk-based breast cancer screening strategies in women. *Best Pract Res Clin Obstet Gynaecol*. 2020;65:3–17.
- Allweis, TM, Hermann N, Berenstein-Molho R, Guindy M. Personalized screening for breast cancer: rationale, present practices, and future directions. *Ann Surg Oncol*. 2021. <https://doi.org/10.1245/s10434-020-09426-1>.
- Ryan R. Cochrane Consumers and Communication Review Group. Data synthesis and analysis. 2013. <http://cccr.cochrane.org>.
- Engmann NJ, Golmakani MK, Miglioretti DL, Sprague BL, Kerlikowske K, Breast Cancer Surveillance, C. Population-attributable risk proportion of clinical risk factors for breast cancer. *JAMA Oncol*. 2017;3:1228–36.
- Amadou A, Hainaut P, Romieu I. Role of obesity in the risk of breast cancer: lessons from anthropometry. *J Oncol*. 2013;2013:906495.
- Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. *CA Cancer J Clin*. 2017;67:378–97.
- Ueda K, Tsukuma H, Tanaka H, Ajiki W, Oshima A. Estimation of individualized probabilities of developing breast cancer for Japanese women. *Breast Cancer*. 2003;10:54–62.
- Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med*. 2004;23:1111–30.
- van Veen EM, Brentnall AR, Byers H, Harkness EF, Astley SM, Sampson S, et al. Use of single-nucleotide polymorphisms and mammographic density plus classic risk factors for breast cancer risk prediction. *JAMA Oncol*. 2018;4:476–82.
- Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer*. 2008;98:1457–66.
- Antoniou AC, Hardy R, Walker L, Evans DG, Shenton A, Eeles R, et al. Predicting the likelihood of carrying a BRCA1 or BRCA2 mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics. *J Med Genet*. 2008;45:425–31.
- Antoniou AC, Pharoah PP, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer*. 2004;91:1580–90.
- Antoniou AC, Durocher F, Smith P, Simard J, Easton DF, members, I. B. P. BRCA1 and BRCA2 mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families. *Breast Cancer Res*. 2006;8:R3.
- Eccles SA, Aboagye EO, Ali S, Anderson AS, Armes J, Berditchevski F, et al. Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Res*. 2013;15:R92.
- Yanes T, Young MA, Meiser B, James PA. Clinical applications of polygenic breast cancer risk: a critical review and perspectives of an emerging field. *Breast Cancer Res*. 2020;22:21.
- Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet*. 2013;45:353–61. 361e351–52

38. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet.* 2015;47:373–80.
39. Mavaddat N, Pharoah PD, Michailidou K, Tyrer J, Brook MN, Bolla MK, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst.* 2015;107:djv036.
40. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet.* 2019;104:21–34.
41. Winkel RR, von Euler-Chelpin M, Nielsen M, Petersen K, Lillholm M, Nielsen MB, et al. Mammographic density and structural features can individually and jointly contribute to breast cancer risk assessment in mammography screening: a case-control study. *BMC Cancer.* 2016;16:414.
42. Sprague BL, Conant EF, Onega T, Garcia MP, Beaber EF, Herschorn SD, et al. Variation in mammographic breast density assessments among radiologists in clinical practice: a multicenter observational study. *Ann Intern Med.* 2016;165:457–64.
43. Arieno A, Chan A, Destounis SV. A review of the role of augmented intelligence in breast imaging: from automated breast density assessment to risk stratification. *Am J Roentgenol.* 2019;212:259–70.
44. Ekpo EU, Mello-Thoms C, Rickard M, Brennan PC, McEntee MF. Breast density (BD) assessment with digital breast tomosynthesis (DBT): agreement between Quantra and 5th edition BI-RADS(R). *Breast.* 2016;30:185–90.
45. Ekpo EU, McEntee MF. Measurement of breast density with digital breast tomosynthesis—a systematic review. *Br J Radiol.* 2014;87:20140460.
46. Magny SJ, Shikhman R, Keppke AL. Breast imaging reporting and data system. In: *StatPearls.* StatPearls Publishing: Treasure Island; 2020.
47. Brentnall AR, Cuzick J, Buist DSM, Bowles EJA. Long-term accuracy of breast cancer risk assessment combining classic risk factors and breast density. *JAMA Oncol.* 2018;4:e180174.
48. Vachon CM, Pankratz VS, Scott CG, Haeberle L, Ziv E, Jensen MR, et al. The contributions of breast density and common genetic variation to breast cancer risk. *J Natl Cancer Inst.* 2015;107:dju397.
49. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med.* 2008;148:337–47.
50. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ Open.* 2015;5:e007825.
51. Ming C, Viassolo V, Probst-Hensch N, Dinov ID, Chappuis PO, Katapodi MC. Machine learning-based lifetime breast cancer risk reclassification compared with the BOA-DICEA model: impact on screening recommendations. *Br J Cancer.* 2020;123:860–7.
52. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD). *Ann Intern Med.* 2015;162:735–6.
53. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162:W1–73.
54. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol.* 2016;69:245–7.
55. Austin PC, van Klaveren D, Vergouwe Y, Nieboer D, Lee DS, Steyerberg EW. Validation of prediction models: examining temporal and geographic stability of baseline risk and estimated covariate effects. *Diagn Progn Res.* 2017;1:12.
56. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J.* 2014;35:1925–31.
57. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating diagnostic, t. et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17:230.
58. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ.* 2016;352:i6.
59. Steyerberg EW, Vickers AJ, Cook NR, Gerdts T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21:128–38.
60. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation.* 2007;115:928–35.
61. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. *Stat Med.* 2004;23:723–48.
62. Wynants L, Riley RD, Timmerman D, Van Calster B. Random-effects meta-analysis of the clinical utility of tests and prediction models. *Stat Med.* 2018;37:2034–52.
63. Riley RD, Ensor J, Snell KI, Debray TP, Altman DG, Moons KG, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ.* 2016;353:i140.
64. Louro J, Posso M, Hilton Boon M, Roman M, Domingo L, Castells X, et al. A systematic review and quality assessment of individualised breast cancer risk prediction models. *Br J Cancer.* 2019;121:76–85.
65. Jaime Caro J, Eddy DM, Kan H, Kaltz C, Patel B, Eldessouki R, et al. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health.* 2014;17:174–82.
66. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med.* 2019;170:W1–W33.
67. Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Srinakarin J, Woodtichartpreecha P, Hirunpat S, et al. Development and validation of a breast cancer risk prediction model for Thai women: a cross-sectional study. *Asian Pac J Cancer Prev.* 2014;15:6811–7.
68. Brentnall AR, Harkness EF, Astley SM, Donnelly LS, Stavrinou P, Sampson S, et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Res.* 2015;17:147.
69. Vilmun BM, Vejborg I, Lyng E, Lillholm M, Nielsen M, Nielsen MB, et al. Impact of adding breast density to breast cancer risk models: a systematic review. *Eur J Radio.* 2020;127:109019.
70. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol.* 2016;74:167–76.
71. Terry MB, Liao Y, Whittemore AS, Loeoe N, Buchsbaum R, Zeinomar N, et al. 10-year performance of four models of breast cancer risk: a validation study. *Lancet Oncol.* 2019;20:504–17.
72. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019;110:12–22.
73. Haibe-Kains B, Adam GA, Hosny A, Khodakarami F, Massive Analysis Quality Control Society Board of, D, Waldron L, et al. Transparency and reproducibility in artificial intelligence. *Nature.* 2020;586:E14–6.
74. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* 2006;7:91.
75. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PLoS ONE.* 2019;14:e0224365.
76. Austin PC, van Klaveren D, Vergouwe Y, Nieboer D, Lee DS, Steyerberg EW. Geographic and temporal validity of prediction models: different approaches were useful to examine model performance. *J Clin Epidemiol.* 2016;79:76–85.
77. Pal Choudhury P, Wilcox AN, Brook MN, Zhang Y, Ahearn T, Orr N, et al. Comparative validation of breast cancer risk prediction models and projections for future risk stratification. *J Natl Cancer Inst.* 2020;112:278–85.
78. Taplin SH, Thompson RS, Schnitzer F, Anderman C, Immanuel V. Revisions in the risk-based Breast Cancer Screening Program at Group Health Cooperative. *Cancer.* 1990;66:812–8.
79. Yen AM, Tsau HS, Fann JC, Chen SL, Chiu SY, Lee YC, et al. Population-based breast cancer screening with risk-based and universal mammography screening compared with clinical breast examination: a propensity score analysis of 1 429 890 Taiwanese women. *JAMA Oncol.* 2016;2:915–21.
80. Liu FC, Lin HT, Kuo CF, See LC, Chiou MJ, Yu HP. Epidemiology and survival outcome of breast cancer in a nationwide study. *Oncotarget.* 2017;8:16939–50.
81. Mukama T, Kharazmi E, Xing X, Sundquist K, Sundquist J, Brenner H, et al. Risk-adapted starting age of screening for relatives of patients with breast cancer. *JAMA Oncol.* 2019;6:68–74.
82. Mukama T, Fallah M, Tian Y, Sundquist K, Sundquist J, Brenner H, et al. Risk-tailored starting age of breast cancer screening based on women's reproductive profile: a nationwide cohort study. *Eur J Cancer.* 2020;124:207–13.
83. Duffy SW, Vulkan D, Cuckle H, Parmar D, Sheikh S, Smith RA, et al. Effect of mammographic screening from age 40 years on breast cancer mortality (UK Age trial): final results of a randomised, controlled trial. *Lancet Oncol.* 2020;21:1165–72.
84. Feig SA. Personalized screening for breast cancer: a wolf in sheep's clothing? *Am J Roentgenol.* 2015;205:1365–71.
85. Joe BN, Hayward JH. More lives risked with risk-based versus age-based breast cancer screening. *Radiology.* 2019;292:329–30.
86. Neal CH, Rahman WT, Joe AI, Noroozian M, Pinsky RW, Helvie MA. Harms of restrictive risk-based mammographic breast cancer screening. *Am J Roentgenol.* 2018;210:228–34.
87. Burnside ES, Trentham-Dietz A, Shafer CM, Hampton JM, Alagoz O, Cox JR, et al. Age-based versus risk-based mammography screening in women 40–49 years old: a cross-sectional study. *Radiology.* 2019;292:321–8.
88. Lee CS, Ashih H, Sengupta D, Sickles EA, Zuley M, Pisano E. Risk-based screening mammography for women aged <40: outcomes from the national mammography database. *J Am Coll Radiol.* 2020;17:368–76.

89. Price ER, Keedy AW, Gidwaney R, Sickles EA, Joe BN. The potential impact of risk-based screening mammography in women 40-49 years old. *Am J Roentgenol*. 2015;205:1360-4.
90. Rainey L, van der Waal D, Broeders MJM. Dutch women's intended participation in a risk-based breast cancer screening and prevention programme: a survey study identifying preferences, facilitators and barriers. *BMC Cancer*. 2020;20:965.
91. Rainey L, van der Waal D, Jervaeus A, Donnelly LS, Evans DG, Hammarstrom M, et al. European women's perceptions of the implementation and organisation of risk-based breast cancer screening and prevention: a qualitative study. *BMC Cancer*. 2020;20:247.
92. Gabrielson M, Eriksson M, Hammarstrom M, Borgquist S, Leifland K, Czene K, et al. Cohort profile: the Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA). *Int J Epidemiol*. 2017;46:1740-41.
93. Eriksson M, Czene K, Pawitan Y, Leifland K, Darabi H, Hall P. A clinical model for identifying the short-term risk of breast cancer. *Breast Cancer Res*. 2017;19:29.
94. Vachon CM, Scott CG, Tamimi RM, Thompson DJ, Fasching PA, Stone J, et al. Joint association of mammographic density adjusted for age and body mass index and polygenic risk score with breast cancer risk. *Breast Cancer Res*. 2019;21:68.
95. Ferreira MA, Gamazon ER, Al-Ejeh F, Aittomaki K, Andrulis IL, Anton-Culver H, et al. Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. *Nat Commun*. 2019;10:1741.
96. Escala-Garcia M, Guo Q, Dork T, Canisius S, Keeman R, Dennis J, et al. Genome-wide association study of germline variants and breast cancer-specific mortality. *Br J Cancer*. 2019;120:647-57.
97. Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet*. 2013;45:392-8. 398e391-92
98. Garcia-Closas M, Gail MH, Kelsey KT, Ziegler RG. Searching for blood DNA methylation markers of breast cancer risk and early detection. *J Natl Cancer Inst*. 2013;105:678-80.
99. Evans DG, Astley S, Stavrinou P, Harkness E, Donnelly LS, Dawe S, et al. Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. In: Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. NIHR Journals Library; 2016. <https://doi.org/10.3310/pgfar04110>.
100. Evans DGR, Harkness EF, Brentnall AR, van Veen EM, Astley SM, Byers H, et al. Breast cancer pathology and stage are better predicted by risk stratification models that include mammographic density and common genetic variants. *Breast Cancer Res Treat*. 2019;176:141-8.
101. Evans DG, Donnelly LS, Harkness EF, Astley SM, Stavrinou P, Dawe S, et al. Breast cancer risk feedback to women in the UK NHS breast screening population. *Br J Cancer*. 2016;114:1045-52.
102. Evans DG, Howell A. Can the breast screening appointment be used to provide risk assessment and prevention advice? *Breast Cancer Res*. 2015;17:84.
103. French DP, Southworth J, Howell A, Harvie M, Stavrinou P, Watterson D, et al. Psychological impact of providing women with personalised 10-year breast cancer risk estimates. *Br J Cancer*. 2018;118:1648-57.
104. Howell A, Astley S, Warwick J, Stavrinou P, Sahin S, Ingham S, et al. Prevention of breast cancer in the context of a national breast screening programme. *J Intern Med*. 2012;271:321-30.
105. Shieh Y, Eklund M, Madlensky L, Sawyer SD, Thompson CK, Stover Fiscalini A, et al. Breast Cancer screening in the precision medicine era: risk-based screening in a population-based trial. *J Natl Cancer Inst*. 2017;109:djw290.
106. Giordano L, Gallo F, Petracci E, Chiorino G, Segnan N. The ANDROMEDA prospective cohort study: predictive value of combined criteria to tailor breast cancer screening and new opportunities from circulating markers: study protocol. *BMC Cancer*. 2017;17:785.
107. Paci E, Mantellini P, Giorgi Rossi P, Falini P, Puliti D. Tailored Breast Screening Trial (TBST). *Epidemiol Prev*. 2013;37:317-27.
108. French DP, Astley S, Brentnall AR, Cuzick J, Dobrashian R, Duffy SW, et al. What are the benefits and harms of risk stratified screening as part of the NHS breast screening Programme? Study protocol for a multi-site non-randomised comparison of BC-predict versus usual screening (NCT04359420). *BMC Cancer*. 2020;20:570.
109. Uzan C, Ndiaye-Gueye D, Nikpayam M, Oueld Es Cheikh E, Lebegue G, Canlorbe G, et al. First results of a breast cancer risk assessment and management consultation. *Bull Cancer*. 2020;107:972-81.
110. Weigert J, Cavanaugh N, Ju T. Evaluating mammographer acceptance of MammoRisk Software. *Radio Technol*. 2018;89:344-50.
111. Gray E, Donten A, Karssemeijer N, van Gils C, Evans DG, Astley S, et al. Evaluation of a stratified National Breast Screening Program in the United Kingdom: an early model-based cost-effectiveness analysis. *Value Health*. 2017;20:1100-9.
112. Trentham-Dietz A, Kerlikowske K, Stout NK, Miglioretti DL, Schechter CB, Ergun MA, et al. Tailoring breast cancer screening intervals by breast density and risk for women aged 50 years or older: collaborative modeling of screening outcomes. *Ann Intern Med*. 2016;165:700-12.
113. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Ann Intern Med*. 2011;155:10-20.
114. Arnold M, Pfeifer K, Quante AS. Is risk-stratified breast cancer screening economically efficient in Germany? *PLoS ONE*. 2019;14:e0217213.
115. Arnold M, Quante AS. Personalized mammography screening and screening adherence—a simulation and economic evaluation. *Value Health*. 2018;21:799-808.
116. Sankatsing VDV, van Ravestein NT, Heijnsdijk EAM, Broeders MJM, de Koning HJ. Risk stratification in breast cancer screening: cost-effectiveness and harm-benefit ratios for low-risk and high-risk women. *Int J Cancer*. 2020;147:3059-67.
117. Sun L, Legood R, Sadique Z, Dos-Santos-Silva I, Yang L. Cost-effectiveness of risk-based breast cancer screening programme, China. *Bull World Health Organ*. 2018;96:568-77.
118. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. *J Natl Cancer Inst*. 2011;103:1436-43.
119. de Jonge ET, Vlasselaer J, Van de Putte G, Schobbens JC. The construct of breast cancer risk perception: need for a better risk communication? *Facts Views Vis Obgyn*. 2009;1:122-9.
120. Edwards AG, Naik G, Ahmed H, Elwyn GJ, Pickles T, Hood K, et al. Personalised risk communication for informed decision making about taking screening tests. *Cochrane Database Syst Rev*. 2013;2013:Cd001865.
121. Evans DG, Brentnall AR, Harvie M, Dawe S, Sergeant JC, Stavrinou P, et al. Breast cancer risk in young women in the national breast screening programme: implications for applying NICE guidelines for additional screening and chemoprevention. *Cancer Prev Res*. 2014;7:993-1001.
122. Yanes T, Kaur R, Meiser B, Scheepers-Joynt M, McInerney S, Barlow-Stewart K, et al. Women's responses and understanding of polygenic breast cancer risk information. *Fam Cancer*. 2020;19:297-306.
123. Printz C. Most women have an inaccurate perception of their breast cancer risk. *Cancer*. 2014;120:314-5.
124. Abittan B, Pachtman S, Herman S, Indelicato J, Herman J. Perception of breast cancer risk in over 11,000 patients during routine mammography exam. *J Cancer Educ*. 2020;35:782-7.
125. Gigerenzer G. Breast cancer screening pamphlets mislead women. *BMJ*. 2014;348:g2636.
126. McWilliams L, Woof VG, Donnelly LS, Howell A, Evans DG, French DP. Risk stratified breast cancer screening: UK healthcare policy decision-making stakeholders' views on a low-risk breast screening pathway. *BMC Cancer*. 2020;20:680.
127. Woof VG, Ruane H, French DP, Ulph F, Qureshi N, Khan N, et al. The introduction of risk stratified screening into the NHS breast screening Programme: views from British-Pakistani women. *BMC Cancer*. 2020;20:452.
128. He X, Schifferdecker KE, Ozanne EM, Tosteson ANA, Woloshin S, Schwartz LM. How do women view risk-based mammography screening? A qualitative study. *J Gen Intern Med*. 2018;33:1905-12.
129. Schifferdecker KE, Tosteson ANA, Kaplan C, Kerlikowske K, Buist DSM, Henderson LM, et al. Knowledge and perception of breast density, screening mammography, and supplemental screening: in search of "Informed". *J Gen Intern Med*. 2020;35:1654-60.
130. Ghanouni A, Sanderson SC, Pashayan N, Renzi C, von Wagner C, Waller J. Attitudes towards risk-stratified breast cancer screening among women in England: a cross-sectional survey. *J Med Screen*. 2020;27:138-45.
131. Ghanouni A, Waller J, Stoffel ST, Vlaev I, von Wagner C. Acceptability of risk-stratified breast screening: effect of the order of presenting risk and benefit information. *J Med Screen*. 2020;27:52-6.
132. Lippey J, Keogh LA, Mann GB, Campbell IG, Forrest LE. "A Natural Progression": Australian women's attitudes about an individualized Breast Screening Model. *Cancer Prev Res*. 2019;12:383-90.
133. Meisel SF, Pashayan N, Rahman B, Side L, Fraser L, Gessler S, et al. Adjusting the frequency of mammography screening on the basis of genetic risk: Attitudes among women in the UK. *Breast*. 2015;24:237-41.
134. van Erkelens A, Sie AS, Manders P, Visser A, Duijm LE, Mann RM, et al. Online self-test identifies women at high familial breast cancer risk in population-based breast cancer screening without inducing anxiety or distress. *Eur J Cancer*. 2017;78:45-52.
135. Fürst N, Kiechle M, Strahwald B, Quante AS. Mammography Screening 2.0 - How can risk-adapted screening be implemented in clinical practice?: results of a focus group discussion with experts in the RISIKOLOTSE.DE Project. *Geburtshilfe Frauenheilkd*. 2018;78:506-11.
136. Rainey L, van der Waal D, Donnelly LS, Evans DG, Wengström Y, Broeders M. Women's decision-making regarding risk-stratified breast cancer screening and prevention from the perspective of international healthcare professionals. *PLoS ONE*. 2018;13:e0197772.
137. Wong XY, Chong KJ, van Til JA, Wee HL. A qualitative study on Singaporean women's views towards breast cancer screening and Single Nucleotide



- Polymorphisms (SNPs) gene testing to guide personalised screening strategies. *BMC Cancer*. 2017;17:776.
138. Elkin EB, Pocus VH, Mushlin AI, Cigler T, Atoria CL, Polaneczy MM. Facilitating informed decisions about breast cancer screening: development and evaluation of a web-based decision aid for women in their 40s. *BMC Med Inform Decis Mak*. 2017;17:29.
  139. NICE. The NHS Breast Screening Programme. 2017. <https://cks.nice.org.uk/topics/breast-screening/background-information/the-nhs-breast-screening-programme/>. Accessed 28th June 2021.
  140. Roth MY, Elmore JG, Yi-Frazier JP, Reisch LM, Oster NV, Miglioretti DL. Self-detection remains a key method of breast cancer detection for U.S. women. *J Women's Health*. 2011;20:1135–9.
  141. Miglioretti DL, Zhu W, Kerlikowske K, Sprague BL, Onega T, Buist DS, et al. Breast tumor prognostic characteristics and biennial vs annual mammography, age, and menopausal status. *JAMA Oncol*. 2015;1:1069–77.
  142. Vass CM, Rigby D, Payne K. Investigating the heterogeneity in women's preferences for breast screening: does the communication of risk matter? *Value Health*. 2018;21:219–28.
  143. Siscic J, Pelletier-Fleury N, Moumjid N. Women's benefits and harms trade-offs in breast cancer screening: results from a discrete-choice experiment. *Value Health*. 2018;21:78–88.
  144. Carbillon L, Bricou A, Sellier N. Challenges, benefits, and harms of risk-based screening mammography in women 40–49 years old. *Am J Roentgenol*. 2016;206:W50.
  145. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *N. Engl J Med*. 2019;381:2091–102.
  146. Weinstein SP, Korhonen K, Cirelli C, Schnall MD, McDonald ES, Pantel AR, et al. Abbreviated breast magnetic resonance imaging for supplemental screening of women with dense breasts and average risk. *J Clin Oncol*. 2020;38:3874–82.
  147. Pashayan N, Duffy SW, Chowdhury S, Dent T, Burton H, Neal DE, et al. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. *Br J Cancer*. 2011;104:1656–63.
  148. Pashayan N, Duffy SW, Neal DE, Hamdy FC, Donovan JL, Martin RM, et al. Implications of polygenic risk-stratified screening for prostate cancer on overdiagnosis. *Genet Med*. 2015;17:789–95.
  149. Vickers AJ. Redesigning prostate cancer screening strategies to reduce overdiagnosis. *Clin Chem*. 2019;65:39–41.
  150. Vickers AJ, Sjoberg DD, Ulmert D, Vertosick E, Roobol MJ, Thompson I, et al. Empirical estimates of prostate cancer overdiagnosis by age and prostate-specific antigen. *BMC Med*. 2014;12:26.
  151. Loeb S, Bjurlin MA, Nicholson J, Tammela TL, Penson DF, Carter HB, et al. Overdiagnosis and overtreatment of prostate cancer. *Eur Urol*. 2014;65:1046–55.
  152. Kopans DB. Arguments against mammography screening continue to be based on faulty science. *Oncologist*. 2014;19:107–12.
  153. Kopans DB. The faulty analysis of breast cancer screening data. *Acad Radiol*. 2016;23:116–7.
  154. Marmot MG. Sorting through the arguments on breast screening. *JAMA*. 2013;309:2553–4.
  155. Bleyer A, Thomas CR Jr, Baines C, Miller AB. Flawed assumptions used to defend screening mammography. *Cancer*. 2015;121:320–1.
  156. Gotzsche PC, Jorgensen KJ, Zahl PH. Breast screening: why estimates differ by a factor of 20–25. *J Med Screen*. 2010;17:158–9. author reply 159–60.
  157. Jorgensen KJ, Gotzsche PC. Breast cancer screening: benefit or harm? *JAMA*. 2016;315:1402.
  158. Jorgensen KJ, Keen JD, Zahl PH, Gotzsche PC. The Two-County breast screening trial cannot provide a reliable estimate of the effect of breast cancer screening. *Radiology*. 2012;262:729–30, author reply 730–31.
  159. Zahl PH, Jorgensen KJ, Gotzsche PC. Lead-time models should not be used to estimate overdiagnosis in cancer screening. *J Gen Intern Med*. 2014;29:1283–6.
  160. Zahl PH, Jorgensen KJ, Gotzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. *Br J Cancer*. 2013;109:2014–9.
  161. Department of Health & Social Care. NHS breast screening (BSP) programme. <https://www.gov.uk/health-and-social-care/population-screening-programmes-breast>. Accessed 27 Nov 2020.
  162. US Preventive Services Task Force. Breast cancer: screening. <https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/breast-cancer-screening>. Accessed 27 Nov 2020.
  163. Klarenbach S, Sims-Jones N, Lewin G, Singh H, Theriault G, Tonelli M, et al. Recommendations on screening for breast cancer in women aged 40–74 years who are not at increased risk for breast cancer. *CMAJ*. 2018;190:E1441–51.
  164. Rijksinstituut voor Volksgezondheid en Milieu. Bevolkingsonderzoek borsstkanker. <https://www.rivm.nl/bevolkingsonderzoek-borstkanker>. Accessed 27 Nov 2020.
  165. BreastScreen Australia. <http://www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/breast-screening-1>. Accessed 27 Nov 2020.
  166. National Health Commission Of The People's Republic Of China. Chinese guidelines for diagnosis and treatment of breast cancer 2018 (English version). *Chin J Cancer Res*. 2019;31:259–77.
  167. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989;81:1879–86.
  168. Gail MH, Costantino JP, Pee D, Bondy M, Newman L, Selvan M, et al. Projecting individualized absolute invasive breast cancer risk in African American women. *J Natl Cancer Inst*. 2007;99:1782–92.
  169. Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. *Breast Cancer Res Treat*. 2005;94:115–22.
  170. Zhang X, Rice M, Tworoger SS, Rosner BA, Eliassen AH, Tamimi RM, et al. Addition of a polygenic risk score, mammographic density, and endogenous hormones to existing breast cancer risk prediction models: a nested case-control study. *PLoS Med*. 2018;15:e1002644.
  171. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer. *J Clin Oncol*. 2015;33:3137–43.
  172. Rosner B, Colditz GA. Nurses' health study: log-incidence mathematical model of breast cancer incidence. *J Natl Cancer Inst*. 1996;88:359–64.
  173. Rosner B, Colditz GA, Iglehart JD, Hankinson SE. Risk prediction models with incomplete data with application to prediction of estrogen receptor-positive breast cancer: prospective data from the Nurses' Health Study. *Breast Cancer Res*. 2008;10:R55.
  174. Park B, Ma SH, Shin A, Chang MC, Choi JY, Kim S, et al. Korean risk assessment model for breast cancer risk prediction. *PLoS ONE*. 2013;8:e76736.
  175. Abdolell M, Payne J, Caines J, Tsuruda K, Barnes PJ, Talbot PJ, et al. Assessing breast cancer risk within the general screening population: developing a breast cancer risk model to identify higher risk women at mammographic screening. *Eur Radiol*. 2020;30:5417–26.
  176. Wang Y, Gao Y, Battsend M, Chen K, Lu W, Wang Y. Development of a risk assessment tool for projecting individualized probabilities of developing breast cancer for Chinese women. *Tumour Biol*. 2014;35:10861–9.
  177. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. *J Natl Cancer Inst*. 2006;98:1204–14.
  178. van den Broek JJ, Schechter CB, van Ravesteyn NT, Janssens A, Wolfson MC, Trentham-Dietz A, et al. Personalizing breast cancer screening based on polygenic risk and family history. *J Natl Cancer Inst*. 2020. <https://doi.org/10.1093/jnci/djaa127>.

## ACKNOWLEDGEMENTS

We thank the anonymous reviewers whose comments on earlier versions of these manuscripts have helped materially improve this paper.

## AUTHOR CONTRIBUTIONS

AKC conceived the study, undertook the literature review and wrote the first draft of the manuscript. DD provided clinical expertise-based input throughout the study and critically reviewed the manuscript. SL provided clinical expertise-based input throughout the study and critically reviewed the manuscript. SP provided health economics expertise-based input throughout the study and critically reviewed the manuscript. MB provided imaging and technological expertise-based input throughout the study and critically reviewed the manuscript. GSC provided statistical expertise-based input throughout the study and critically reviewed the manuscript. JH-C helped conceive the study, provided epidemiology expertise-based input throughout the study and critically reviewed the manuscript.

## FUNDING

This work was supported by Cancer Research UK [C2195/A31310 to AKC, C8225/A21133 to DD, C49297/A27294 to GSC, C5255/A18085 to JH-C], the NIHR Biomedical Research Centre, Oxford [GSC], the NIHR (SP receives support as an NIHR Senior Investigator [NF-SI-0616-10103]) and the NIHR Applied Research Collaboration Oxford and Thames Valley (SP).

## COMPETING INTERESTS

JH-C is an unpaid director of QResearch (a not-for-profit organisation, which is a partnership between the University of Oxford and EMIS Health, which supply the QResearch database) and is a founder and shareholder of ClinRisk Ltd and was its medical director until 31 May 2019 (ClinRisk produces open and closed source software to implement clinical risk algorithms, including two breast cancer risk models implemented in the NHS into clinical computer systems (outside of this work)). All the remaining authors declare no competing interests.

**ETHICS APPROVAL AND CONSENT TO PARTICIPATE**

Not applicable—this review did not use original/clinical data from human participants or tissue.

**ADDITIONAL INFORMATION**

**Correspondence** and requests for materials should be addressed to Ash Kieran Clift.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021