The Breast xxx (2015) 1–10

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Review

Simulation models in population breast cancer screening: A systematic review

Rositsa G. Koleva-Kolarova [a, *], Zhuozhao Zhan [a], Marcel J.W. Greuter [b], Talitha L. Feenstra [a, c], Geertruida H. De Bock [a]

[a] University of Groningen, University Medical Center Groningen, Department of Epidemiology, PO Box 30.001, 9700RB Groningen, The Netherlands
[b] University of Groningen, University Medical Center Groningen, Department of Radiology, PO Box 30.001, 9700RB Groningen, The Netherlands
[c] RIVM, PO Box 1, 3720BA Bilthoven, The Netherlands

ARTICLE INFO

Article history:
Received 8 February 2015
Received in revised form
17 March 2015
Accepted 24 March 2015
Available online xxx

Keywords:
Breast neoplasms
Mass screening
Computer simulation
Mortality
Cost–benefit analysis

ABSTRACT

The aim of this review was to critically evaluate published simulation models for breast cancer screening of the general population and provide a direction for future modeling. A systematic literature search was performed to identify simulation models with more than one application. A framework for qualitative assessment which incorporated model type; input parameters; modeling approach, transparency of input data sources/assumptions, sensitivity analyses and risk of bias; validation, and outcomes was developed. Predicted mortality reduction (MR) and cost-effectiveness (CE) were compared to estimates from meta-analyses of randomized control trials (RCTs) and acceptability thresholds. Seven original simulation models were distinguished, all sharing common input parameters. The modeling approach was based on tumor progression (except one model) with internal and cross validation of the resulting models, but without any external validation. Differences in lead times for invasive or non-invasive tumors, and the option for cancers not to progress were not explicitly modeled. The models tended to overestimate the MR (11–24%) due to screening as compared to optimal RCTs 10% (95% CI – 2–21%) MR. Only recently, potential harms due to regular breast cancer screening were reported. Most scenarios resulted in acceptable cost-effectiveness estimates given current thresholds. The selected models have been repeatedly applied in various settings to inform decision making and the critical analysis revealed high risk of bias in their outcomes. Given the importance of the models, there is a need for externally validated models which use systematical evidence for input data to allow for more critical evaluation of breast cancer screening.

Introduction

Although breast cancer screening has been implemented in several areas of the world, there is still debate on the effectiveness of regular screening and its impact on mortality reduction (MR) [1–5]. Randomized control trials (RCTs) are the golden standard for evaluating health interventions prior to their wide population-based implementation; however, in the case of breast cancer screening a long follow-up time and large groups of participants are needed to evaluate the effects on MR and the potential benefits and harms associated with regular mammographic screening [2,5,6]. Therefore, simulation models are often applied along with RCTs to ensure proper evaluation of the effects of screening.

Simulation models provide the opportunity to evaluate health interventions and compare scenarios to find the optimum policy without having to trial each variant. Simulation can extrapolate the results of RCTs to different population sub-groups and provide health technology assessment of screening interventions [7]. The modeling of breast cancer screening dates back to the 1980s, when the first models were developed to assess the effect of regular screening on MR [8]. Since then a number of models have been developed to study health-related, economic or aggregate outcomes [9]. As simulation models are used alongside the RCTs to inform decision making, the reliability and the accuracy of their

* Corresponding author. Tel.: +31 50 361 5992; fax: +31 50 361 4493.
E-mail addresses: r.koleva-kolarova@umcg.nl (R.G. Koleva-Kolarova), z.zhan01@umcg.nl (Z. Zhan), m.j.w.greuter@umcg.nl (M.J.W. Greuter), t.l.feenstra@umcg.nl (T.L. Feenstra), g.h.de.bock@umcg.nl (G.H. De Bock).

http://dx.doi.org/10.1016/j.breast.2015.03.013
0960-9776/

outcomes are of high importance. Guidelines have provided frameworks for quality assessment, however, clear criteria for what a good model should incorporate are yet lacking [7,10]. Several reviews have previously evaluated simulation models for breast cancer screening [11,12], however, these were performed by the group which developed the models and lacked the critical approach of the current work.

The aim of this review was to critically evaluate published simulation models for breast cancer screening of the general population and to provide a direction for future modeling.

Materials and methods

Search strategies and eligibility criteria

A systematic literature search was performed in PubMed and Embase for relevant studies (Fig. 1). Included were full-text original articles (irrespective of the year or language of publication) which described simulation studies in breast cancer screening of the general population. Studies were eligible if at least one of the tests was an imaging tool and their main outcomes were MR/life years gained (LYG) and/or cost-effectiveness (CE). Excluded were: studies evaluating breast cancer screening in population sub-groups, breast cancer follow-up or treatment; as well as studies in other type of cancers and non-simulation or statistical studies.

Selection of eligible studies and data extraction

Two authors (RK and ZZ) independently reviewed the potentially relevant publications by title and abstract, and full-text reading where indicated. The studies were selected on the basis of the eligibility criteria described earlier. Disagreement was resolved by discussion and/or arbitrage by a third party (GdB). The data was extracted by a checklist based on the developed framework. The references of the selected papers were examined for other eligible publications.

The framework for qualitative assessment of the simulation models

A framework for qualitative assessment of published simulation models for breast cancer screening of the general population was developed. The framework incorporated model type; input parameters; modeling approach, transparency of input data sources/ assumptions, sensitivity analyses and risk of bias; validation, and outcomes (Tables 1 and 2).

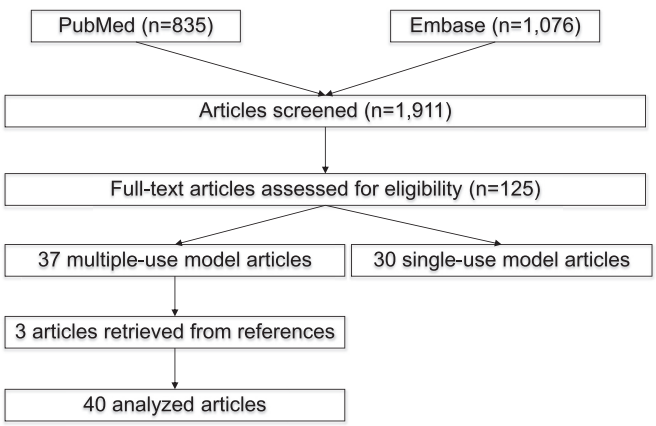

Fig. 1. Flow diagram – identification of eligible studies.

Only multiple-use models were included into the analysis. Models with a single publication were listed in a table and not investigated further (Appendix A). For multiple-use models, three types of publications were identified: the original model description, model applications' and model extensions' papers. The original model description was defined as the originally developed simulation model with distinctive input parameters and validity (Table 1). In case of several publications on the original model description, the paper which described the model the most completely was selected. When the original model was used to evaluate a different outcome or in a different population group without any modifications or extensions, it was classified as a model application. When the original model was modified or extended by adding input parameters, it was considered a model extension (Table 3).

The classification by Brennan et al. [13] was applied to characterize the model type. All original simulation models belonged to the individual sampling model group [13]. The models were further classified into sub-groups. The first sub-group was the discrete-time individual event history models, using Markovian state transitions and discrete-time periods. The second sub-group was the discrete individual simulation models, which were non-Markovian and used discrete-time cycles. The third sub-group was the simulated patient-level Markov models, which simulated individuals' histories in continuous time, assuming Markov property on all events.

The input parameters encompassed three domains: disease (onset, incidence rate, pattern of progression, stage shifts, mortality rate); population/cohort (dynamic or static cohort, risk factors, general mortality rate), and intervention (sensitivity/specificity of mammography, dissemination of mammography, attendance/recall rate).

The modeling approach was assessed from statements in the studies regarding the modeling of the disease progression. Based on the framework for evaluation of modeling studies of Carter et al. [14] main sources of input data and assumptions, as well as sensitivity analyses for uncertain variables were reported for each original model to assess the transparency and the risk of bias.

The validation of the models was assessed from statements in the publications reporting internal, external and/or cross validation [15,16]. Internal validation was defined as comparison of the output of the model to the data used for deriving and calibrating the input parameters. External validation was defined as comparison of the output of the model to independent data sources, which were different than the ones used for input calibration. Cross validation was defined as comparison of the output of the model to the output of other models [15,16].

Model outcomes considered expected benefits (number of screen detected and interval tumors, tumor size, mortality decline, prolonged survival, costs/cost-effectiveness) and potential harms (false-positives, false-negatives, overdiagnosis of indolent [17], in-situ and invasive tumors [14]/overtreatment, radiation-induced tumor risk).

Comparison of predicted MR to RCTs

To compare model predicted MR to estimates of MR from RCTs, the most recent meta-analyses were identified [5,18–20]. The Cochrane review was used as a reference as it encompassed the most relevant breast cancer screening RCTs performed in the Western World and assessed the quality of the RCTs by distinguishing between sub-optimally and optimally randomized

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

Table 1
Base case models.

Model	General population USA	General population NL	Birth cohorts	Breast cancer incidence	Tumor progression	Sensitivity (and specificity)	Treatment dissemination	Mortality rate	Cross validation	Internal validation	Transparency score%	Reference
MISCAN		x	x	x	x	x	x	x		x	91	[23]
MISCAN – Fadia	x		x	x	x	x	x	x	x	x	91	[24]
SPECTRUM	x		x	x	x	x	x	x		x	91	[25]
UWBCS	x		x	x	x	x	x	x		x	91	[26]
Stanford University	x		x	x	x	x	x	x		x	91	[27]
Dana Farber Cancer Institute	x		x	x	x	x	x	x		x	91	[28]
MD Anderson Cancer Center	x		x	x		x	x	x			73	[29]

NL – The Netherlands; USA – United States of America.
MISCAN – the "micro-simulation analytic model". SPECTRUM – the Georgetown University "Simulating Population Effects of Cancer Control interventions – Race and Understanding Mortality". UWBCS – the "Wisconsin breast cancer epidemiology simulation model".

Table 2
Summary of evidence for the original models: source and transparency (systematical evaluation) of data/assumptions, external validation, sensitivity analyses and overall risk of bias [14].

Model (reference)	Data/assumptions sources: (1) Breast cancer incidence (2) General mortality (3) Tumor progression (4) Sensitivity/specificity of mammography (5) Transparency of data/assumptions sources?[a]	(1) External validation; (2) Sensitivity analyses (3) Key uncertain variables[b]	Overall risk of bias
MISCAN (23)	(1) Dutch screening program in Nijmegen and Utrecht (2) Dutch life-table for 1982–83 (3) Tumor diameter obtained from cancers diagnosed outside the screening program; exponential distribution of the duration of preclinical invasive stages (4) Sensitivity depends on tumor diameter and specificity estimated from screening program (5) No	(1) No (2) Yes (3) Preclinical period; Sensitivity of mammography	High
MISCAN – Fadia (24)	(1) SEER (2) USA life-table for each cohort (3) Tumor growth rate based on lognormal distribution, fatal tumor diameter on Weibull distribution (4) Sensitivity based on threshold tumor diameter (5) No	(1) No (2) Yes (3) Continuous tumor growth model	High
SPECTRUM (25)	(1) SEER (2) USA life-table for each cohort (3) Age-dependent sojourn and dwell time exponentially distributed and based on published data (4) Sensitivity based on published age-specific estimates, assumptions about variation of sensitivity with regards to age and screening round (5) No	(1) No (2) Yes (3) Sojourn time; sensitivity of mammography	High
UWBCS (26)	(1) SEER (2) USA life-tables from the Center of Disease Control and Prevention (3) Tumor growth based on Gompertzian distribution; a fraction limited malignant potential tumors with maximum size and sojourn time assumed (4) Sensitivity based on tumor diameter (5) No	(1) No (2) Yes (3) Tumor progression	High
Stanford University (27)	(1) SEER and Connecticut Tumor Registry (2) USA life-table for each cohort (3) Tumor growth based on exponential distribution (4) Sensitivity based on tumor diameter threshold (5) No	(1) No (2) Yes (3) Treatment efficacy	High
Dana Farber Cancer Institute (28)	(1) SEER (2) Age-specific mortality rate for birth cohort (3) Preclinical sojourn time based on exponential distribution (4) Sensitivity age-dependent based on Breast Cancer Surveillance Consortium publication (5) No	(1) No (2) Yes (3) Sensitivity of mammography; tumor stage shifts	High
MD Anderson Cancer Center (29)	(1) Uniform distribution (0 – constant incidence over time; 1 – incidence estimated from an age-period-cohort model) (2) USA life-table for each cohort (3) No (4) Probability of tumor detection based on National Breast and Cervical Cancer Early Detection Program data (5) No	(1) No (2) Yes (3) Treatment efficacy	High

[a] The transparency of data/assumptions sources is defined as a systematical evaluation and selection of data/assumption sources.
[b] The key uncertain variables are the variables which have the largest effect on the model predicted mortality reduction.

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

Table 3
Applications (A) and extensions (E) of the original models.

Model	Country/year	Application (A) or extension (E)	Mortality reduction (%)	Life years gained per 1000 screened women	ICER in 1000 currency units	Potential harms of screening	Cross validation	Internal validation	Reference
MISCAN	NL/2011	A	26–30			RIBCD 1.3–14.4			[35]
	NL/2011	A	20.8–24.8			OBC 2–25%		x	[36]
	Switzerland/2009	A	13–35		€11–28			x	[37]
	India/2008	A	24.5–25.8		$3–19	FP 5775/$0.75 mln		x	[38]
	Canada/2004	A	13.6–34.1					x	[39]
	UK, NL/1994	A	18–29					x	[40]
	USA/1999	A		8–83	$8.3–38.6				[41]
	UK/1998	A	12.8–16.4	66–81	£2.5–2.7			x	[42]
	Spain/1998	A	12–14.9						[43]
	Germany/1997	A	11		DM 15–21	CFPB 232–576DM		x	[44]
	Spain/1997	A	17–23		Pts 427–555			x	[45]
	Italy/1997	A			£6.587–11.518				[46]
	Sweden/1995	A	7–30						[47]
	NL/1995	A		395–497	£8.4–36				[48]
	Italy/1995	A	13					x	[49]
	Germany/1994	A	11 (7–18)		DM 18.8–25.3			x	[50]
	UK, NL, Spain, France/1993	A		61–252	£1.8–9.7				[51]
	NL/1991	A		180–310	$3–6				[52]
	NL/1991	A		259					[53]
	NL/1998	E	17			RIBCD 0.7–10.1			[62]
	Australia/1993	E			$*13–27				[63]
Miscan-Fadia	USA/2012	A		18–93		FP 795			[54]
	USA/2011	A	13–49			FP 340–2250/UB 24–158	x		[55]
	USA/2009	A		49–227		FP 340–2250/UB 24–158	x		[56]
	USA/2005	A			$58–151 (41–368)			x	[57]
	USA/2004	A			$53–124	FP 0.7/$93		x	[58]
	USA/2010	E	8–27.4					x	[64]
	USA/2013	E	10.7–39.9			FP 19.8–53.5		x	[65]
SPECTRUM	USA/2012	A		18–108		FP 939			[54]
	USA/2011	A	11–32				x		[55]
	USA/2009	A		61–192			x		[56]
	USA/2013	E	11.5–32.1			FP 20.9–57.5		x	[65]
UWBCS	USA/2012	A		21–111		FP 890			[54]
	USA/2011	A	12–54				x		[55]
	USA/2009	A		39–202			x		[56]
	USA/2008	E			$26.5–272 (21–536)	FP $160–371			[66]
	USA/2006	E			$27–58 (25–71)				[67]
Stanford University	USA/2011	A	9–35				x		[55]
	USA/2009	A		52–210			x		[56]
	USA/2011	A		56–206					[59]
Dana-Farber Cancer Institute	USA/2012	A		8–113		FP 877			[54]
	USA/2011	A	11–38				x		[55]
	USA/2009	A		51–177			x		[56]
	Spain/2011	A			€3–715				[60]
	Spain/2009	A	20–30						[61]
MD Anderson Cancer Center	USA/2011	A	10–29				x		[55]
	USA/2009	A		43–128			x		[56]

Outcomes are presented in terms of point estimate (single number) or minimum–maximum predicted values (range) for the different tested screening scenarios.
NL – The Netherlands; UK – United Kingdom; USA – United States of America. € – euro; $ – US dollar; £ – British pound; DM – German mark; £ – Italian lira; $* – Australian dollar; Pts – Spanish pesetas.
Potential harms of screening: OBS – overdiagnosed breast cancers [36], RIBCD – radiation-induced breast cancer deaths [35,62], FP – false positives (including costs) [38,54–56,58,65,66], UB – unnecessary biopsies [55,56], and CFPB – costs of false positive biopsies [44].

RCTs and their relevant contribution to breast cancer MR from screening [5]. The Cochrane review reported 10% (95% CI – 2–21%) MR from three optimal RCTs, while in four suboptimal RCTs the reported MR was higher (25%, 95% CI 17–33%), with a mean of 19% (95% CI 13–26%) of all seven trials [5].

Point estimates and confidence intervals of MR as predicted by the simulation models were compared to the estimates from of RCTs (Fig. 2).

Comparison of predicted CE to acceptability thresholds

The predicted incremental CE ratios (ICERs) were compared to the World Health Organization (WHO) criteria, based on the gross domestic product (GDP) per capita of the country [21]. The originally used currency was converted to international dollars and divided by the GDP per capita for the relevant year to derive the GDP threshold [22]. Only point estimates simulated by the models

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

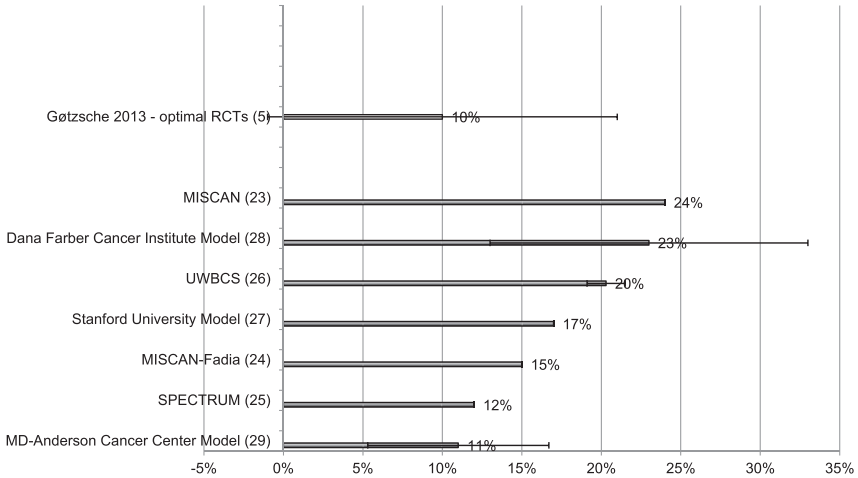

Fig. 2. Comparison of mortality reduction due to breast cancer screening predicted by the original models and estimates from RCTs' meta-analysis (95% CI error bars). 95% CI were derived from the relevant publications where available.

were used. A ratio less than the GDP per capita was considered very cost-effective, a ratio – 1–3 times the GDP per capita was cost-effective and a ratio above 3 times the GDP per capita was not cost-effective [21].

Results

The qualitative assessment of the simulation models

Seven original models were identified which all belonged to the Cancer Intervention and Surveillance Modeling Network (CISNET) Breast Cancer group (Table 1) [23–29]. The original models were developed to study the effect of implementing regular screening and relevant therapy on MR. The MISCAN ("micro-simulation analytic model") originated from the Netherlands and was later adopted by the CISNET as the MISCAN-Fadia. There were several publications related to the MISCAN [23,30–32] and the Dana-Farber [28,33,34] original models, 40 applications described in 27 publications [35–61] and 7 extensions in 6 papers (Table 3) [62–67]. The original models were extended by additional parameters related to modeling of the disease (radiation-induced risk, receptor status) [62,65] and the screening modality (age- and BMI-specific sensitivity, breast density, digital mammography sensitivity and specificity) [65,66].

The MISCAN, the MISCAN-Fadia, the SPECTRUM and the Stanford University models were classified as simulated patient-level continuous time Markov models, the UWBCS and the MD-Anderson models as non-Markovian discrete individual simulation models and the Dana-Farber model as discrete-time individual event history model [13].

The original models shared common input parameters regarding the modeling of the disease, the screened population and the screening intervention (Table 1). All original models based the incidence of breast cancer only on the age of the simulated populations.

The modeling approach was based on simulating the tumor progression which was modeled either exponentially or by using a Gompertz distribution [23–28]. Only the MD-Anderson model did not utilize the tumor progression approach, but distinguished a preclinical and a clinical stage [29]. Models differed in the amount of detail used to model disease progression: the SPECTRUM [25] and the MISCAN-Fadia [24] explicitly used the transition from ductal carcinoma in situ (DCIS) to local, then regional and distant disease, while the Stanford University model [27] used the DCIS transition to local or advanced disease. The models mainly utilized stage shifts and tumor dwelling times to define the preclinical and the clinical period of the tumors. The differences in lead times for invasive and non-invasive tumors, as well as the option that a tumor does not progress or cause any adverse consequences during a lifetime were not explicitly modeled. Only the UWBCS model assumed that there may be a fraction of tumors which does not surface clinically during a lifetime. These tumors were called limited malignant potential (LMP) tumors and were assigned a maximum sojourn time after reaching a maximum of LMP size. After this time, if not discovered, the LMP tumors disappeared the next simulation round [26].

Data about general mortality and breast cancer incidence derived from the same epidemiological databases was used to calibrate the models, but different assumptions were applied to model disease progression and the sensitivity/specificity of screening mammography (Table 2) [23–29]. The use of different data sources and assumptions was clearly described in the papers, however, the evidence for their values was not evaluated systematically, which bears a high risk of bias in these models. The SPECTRUM model systematically selected data for tumor progression and sensitivity of mammography, however, assumptions about the variation in sensitivity with regards to age and screening round were made (Table 2) [25]. The uncertainty of model variables was tested in sensitivity analyses and the assumptions regarding tumor growth [23–26,28], sensitivity of mammography [23,25,28] and treatment efficacy [27,29] were found to have the largest effect on the model predicted MR (Table 2).

None of the models reported external validation. The MD-Anderson model did not report any validation while the remainder six original models reported internal validation [23–28]. Internal validation was also reported by 12 applications [36–40,42,44,45,49,50,57,58] and 3 extensions [64,65]. In 3 studies some validation efforts were observed [41,51,59]. Cross validation was reported by the original MISCAN-Fadia against the original MISCAN [24] and by the model applications in 2 studies [55,56].

The outcomes of the original models were focused on expected benefits as all models reported MR due to regular screening. None of the original models reported potential harms

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

due to screening, however, potential harms due to regular breast cancer screening were reported by later publications of models' applications and extensions. These potential harms included overdiagnosed breast cancers [36], radiation-induced breast cancer deaths [35,62], false positives test results (including costs) [38,54–56,58,65,66], unnecessary biopsies [55,56], and costs of false positive biopsies [44].

Comparison of predicted MR to RCTs

The MR estimates of the original models varied from 11% to 24% [23–29] for the same scenario of introducing screening compared to a reference scenario of no screening. Comparing the models' MR to optimal RCTs 10% (95%CI – 2–21%) MR [5], all models tended to overestimate MR due to screening with 5 models being within the 95% CI [24–27,29]. However, original model estimates compared relatively well to MR from suboptimal trials (25%, 95% CI 17–33%) and from all trials (19%, 95% CI 13–26%) (Fig. 2) [5].

Comparison of predicted CE to acceptability thresholds

ICERs were reported by both model applications and extensions, and varied across screening scenarios in different countries (Table 3, Fig. 3) [37,38,41,42,44–46,48,50–52,57,58,60,63,66,67]. The main question in these studies was which screening scenario would be the most cost-effective one given the particular characteristics of the population and the context in which they were screened.

Most reported screening regimens fulfilled the WHO criteria with the exception of some very intensive USA, Spanish and Indian scenarios.

Discussion

By selecting models on breast cancer screening published more than once, we retrieved 7 original simulation models that shared common input parameters and modeling approach. Six models were internally and/or cross validated, no model was externally validated. Age was the only risk factor for breast cancer incidence considered. Differences in lead times for invasive or non-invasive tumors, and the option for cancers not to progress were not explicitly modeled. The original models tended to overestimate the MR (11–24%) due to screening as compared to the optimal RCTs 10% MR (95% CI – 2–21%). Only recently, potential harms due to regular breast cancer screening were reported in simulation studies. Most evaluated screening scenarios resulted in cost-effectiveness estimates that would be considered value for money given WHO criteria.

Model type could affect the accuracy of the models' output; therefore simplicity and flexibility are advisable when choosing the modeling type [13]. All original models belonged to the individual level group which were reported to be more flexible than cohort ones. A disadvantage of the individual models is that achieving sufficient stability of the outcomes requires repeated runs of the model [13].

The original models used the same epidemiological database to calibrate their input parameters and applied different assumptions and methods for calibration, which could explain the similarities as well as the differences in the predicted MR. However, tuning the models to a particular database could be a considerable limitation as models by definition will be predetermined to produce results compatible with observed epidemiological trends.

The main outcome of the original models was the MR due to regular screening; therefore applying the aggregated breast cancer incidence rate at a population level to model the disease incidence was a reasonable approach. However, age-dependent incidence fails to encompass the change of advanced breast cancer incidence due to risk factors related to increased age of first birth, alcohol consumption and smoking, oral contraceptive usage, and BMI [68,69]. The aggregated approach can be further improved by incorporating risk factors at an individual level, and assessing and comparing screening effectiveness amongst different risk groups.

All models except one were built on the tumor growth approach. Although the tumor growth rate was based on sound statistical assumptions, in reality tumor progression may not be chronological and the lead time for invasive and non-invasive tumors differs. Part of this disadvantage was tackled by applying

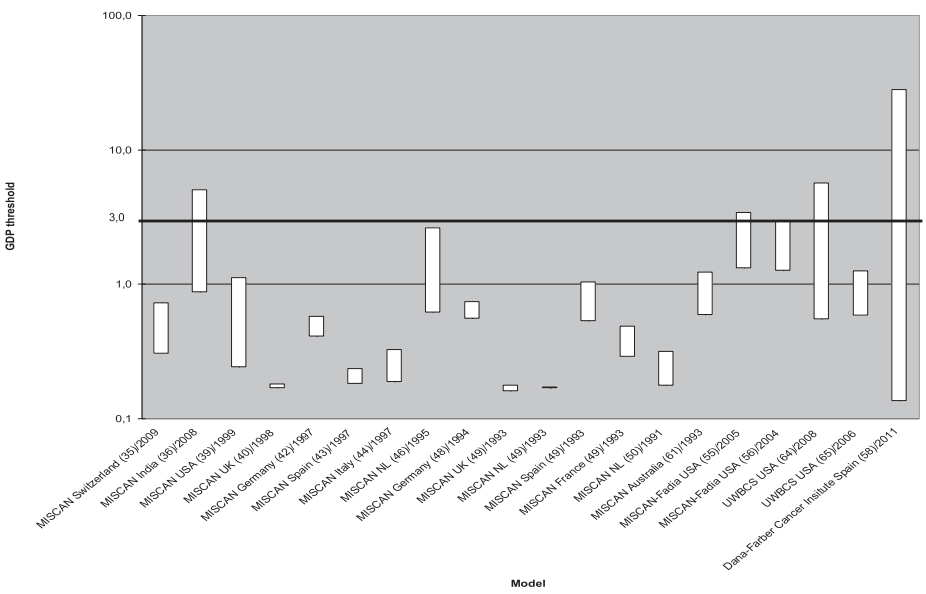

Fig. 3. Cost-effectiveness ratios predicted by the simulation models (in GDP thresholds). The lines at 1,0 and 3,0 represent the WHO threshold value of one to three times GDP per capita [21].

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

a Markovian property with predefined transition probabilities between tumor states, however, the Markovian property does not take into account the previous history and therefore tumor growth is solely based on the current status. As tumor progression influenced model outcomes, the differences in lead times for invasive and non-invasive tumors should be incorporated into the models to represent more realistically the progression of the disease.

In addition, the option for cancers not to progress (or even regress) should be explicitly modeled in the original models. These cancers do not really have a lead time, as they do not cause symptoms. They may be screen detected and contribute to higher incidence rates of breast cancer thus causing bias when estimating lead time distributions from screening program data [17].

Current modeling guidelines recommended validating a simulation model at least once [7], while other authors recommended validation to every use of the model [70]. This review showed that despite the recommendations validation was often omitted in breast cancer screening modeling. Only internal and cross validation were used to evaluate the performance of the analyzed models as these were relatively the easiest and less time-consuming methods due to the availability of the database for calibration and validation, and the use of a similar modeling approach and the simulation of the same sequence of scenario rounds. While internal validation can be used as a method to verify the assumptions made in building the model, it is insufficient to prove the generalizability of the results to a broader or a different population group. External validation can be considered a better approach when generalizability of results is important, but it requires more time and resources. However, external validation was not reported by the models. This could be due to a lack of reliable independent databases. The lack of external validation combined with the lack of systematically selected input data/assumptions can be considered a flaw in current models as it carries the risk for biased model output.

The outcomes reported by the original models regarded only expected benefits of breast cancer screening, e.g. the MR due to regular breast cancer screening. Although several applications and extensions of the models analyzed the potential harms of breast cancer screening, none of the original models reported any potential harms of regular screening. This can be considered as a drawback, because it is well known that false-positives, false-negatives, overdiagnosis/overtreatment and radiation-induced tumors can negatively affect quality of life, as well as the cost-effectiveness of the screening intervention. Further, not considering overdiagnosis could result in biased estimation of the reduction in breast cancer mortality attributable to regular screening.

The original models reported MRs from breast cancer due to screening which tended to overestimate the optimal RCTs estimate. This could be explained by the uncertainty involved in calibrating the models' input parameters and their probability to change over time, as well as the lack of external validation. Another explanation could be that the MR in RCTs is underestimated due to the quality of the included RCTs or the selection of optimal and sub-optimal RCTs.

The comparison of predicted ICERs in absolute values would be unrealistic as they analyzed different populations and policy scenarios. Overall breast cancer screening was reported to be a cost-effective health intervention compared to no screening with the exception of some very intensive screening regimens. The simulated breast cancer screening scenarios in the European countries appeared to be more cost-effective than the screening regimens in the USA and India (Fig. 3).

An advantage of this review is that it presented a selection of comprehensive simulation models and provided a framework for their qualitative assessment. However, this selection might seem limited in scope and give the impression of CISNET dominance over the modeling in breast cancer screening. The selection of only multiple-use models was based on the assumption that these models are well tested. In addition, these models have been repeatedly applied in various settings to inform decision makers in breast cancer screening and can be regarded as very influential, therefore close critical analysis is warranted. Furthermore, only MR/LYG and CE of breast cancer screening in the general population were considered, as these outcomes are the most important when evaluating a health intervention. Simulation models could in principle produce various other outputs in different population sub-groups.

Conclusion

The reviewed simulation models were developed to evaluate the potential harms and benefits of introducing regular breast cancer screening. All models belonged to the individual simulation group and the modeling approach was mainly based on the tumor progression. The models were internally and cross validated, but external validation was not performed which carries the risk for biased model output. Potential harms due to regular breast cancer screening were reported only in later publications of model applications and extensions. The predicted MR (11–24%) tended to overestimate the optimal RCTs' estimate (10%, 95% CI − 2–21%). Most scenarios resulted in cost-effectiveness estimates that would be considered acceptable given current thresholds. The critical evaluation of the selected original models revealed that there is a high risk of bias in their output due to the lack of transparency in the selection of sources for input data/assumptions and the lack of external validation. Given the repeated application of these models in various setting and their importance for policy makers, healthcare professionals and patients, there is a need for externally validated models which use systematical evidence for their input data to allow for more critical evaluation of breast cancer screening.

Conflict of interest statement

None declared.

Funding source

Not applicable.

Ethical approval

No ethical approval was needed for this study.

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

Appendix A

Table A.1
Single-use models.

Country/study/year	Model type	Modeling approach	Validation	MR (%)/LYG*	CER	Reference
Brazil/2012	B	Y	NR		$13.6–2.904	[77]
China/2012	C	Y	NR	14.7–14.8%	¥217–248	[73]
China/2012	B	Y	NR		$73–204	[75]
China/2007	B	Y	C		$62–179	[84]
China/2007	B	N	NR		$90–2.894	[85]
Columbia/2012	B	Y	C	1250–8299*		[72]
Mexico/2009	A	Y	A		MXN75.3–171.1	[81]
Japan/2012	B	Y	NR		¥*310.8	[74]
Japan/1991	B	Y	NR		$14–40	[99]
New Zealand/1996	A	Y	NR		$*13–22	[95]
Slovenia/2008	B	Y	C*	27%	€4–42	[83]
Spain/2012	A	Y	C	2–43%		[76]
Spain/1997	B	Y	NR		ECU 2.1–7.3	[94]
Sweden/2001	B	Y	A	9–48%		[90]
Switzerland/2007	B	Y	A		$73–118	[86]
UK/2010	B	N	C	9–17%		[79]
UK/2010	B	N	NR		£17–30	[80]
UK/1988	B	Y	NR		£3.0–3.4	[100]
USA/2013	A	Y	A	6.2–10.5%	$17–180	[71]
USA/2011	B	Y	A		$8–363	[78]
USA/2009	A	Y	NR		$36–3.939	[82]
USA/2006	A	Y	A	19.8%		[87]
USA/2005	A	Y	A		$9–580	[88]
USA/2005	A	Y	NR		$49–92	[89]
USA/1999	C	Y	NR		$5–1.019	[91]
USA/1998	B	Y	NR		$10–27	[92]
USA/1997	B	Y	NR		$21–168	[93]
USA/1995	B	Y	NR		$16–32	[96]
USA/1994	B	Y	NR	17.3–41.4*		[97]
USA/1994	B	Y	NR	15%		[98]

Outcomes are presented in terms of point estimate (single number) or minimum–maximum predicted values (range) for the different tested screening scenarios.
Model type: A – Micro-simulation Markov/Non-Markovian model; B – Cohort Markov/Non-Markovian model; C – Computer simulation (unspecified).
Modeling approach: Y – Tumor progression; N – Not based on tumor progression.
Validation: A – Internal validation; B – Cross validation; C – Other type of validation efforts.
MR – mortality reduction; CER – represents incremental or marginal cost effectiveness ratio in thousands of the relevant currency units; LYG* - the numbers marked with an asterisk represent life-years gained (LYG).
NR – not reported.
UK – United Kingdom; USA – United States of America.
$ – US dollar; ¥ – Chinese yuan; ¥* – Japanese yen; MXN – Mexican pesos; ECU – European Currency Unit; $* – Australian dollar.
C* validating efforts with some external part as the estimates from the input database were compared to estimates of breast cancer screening RCTs.

References

[1] Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet 2006;368:2053–60.
[2] Djulbegovic B, Lyman GH. Screening mammography at 40–49 years: regret or no regret? Lancet 2006;368:2035–7.
[3] Glasziou P, Houssami N. The evidence base for breast cancer screening. Prev Med 2011;53:100–2.
[4] Senkus E, Kyriakides S, Penault-Llorca F, Poortmans P, Thompson A, Zackrisson S, et al. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013;24(Suppl. 6):vi7–23.
[5] Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013. http://dx.doi.org/10.1002/14651858.CD001877.pub5. Issue 6. Art. No.: CD001877. p.10.
[6] Moss S, Thomas I, Evans A, Thomas B, Johns L. Randomised controlled trial of mammographic screening in women from age 40: results of screening in the first 10 years. Br J Cancer 2005;92:949–54.
[7] Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices — overview: a report of the of the ISPOR-SMDM modeling good research practices task force-1. Value Heal 2012;15:796–803.
[8] Plevritis SK. A mathematical algorithm that computes breast cancer sizes and doubling times detected by screening. Math Biosci 2001;171:155–78.
[9] Plevritis SK, Salzman P, Sigal BM, Glynn PW. A natural history model of stage progression applied to breast cancer. Stat Med 2007;26:581–95.
[10] Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modeling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics 2006;24(4):355–71.
[11] Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784–92.
[12] Clarke LD, Plevritis SK, Boer R, Cronin KA, Feuer EJ. A comparative review of CISNET breast models used to analyze U.S. breast cancer incidence and mortality trends. J Natl Cancer Inst Monogr 2006;36:96–105.
[13] Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ 2006;15:1295–310.
[14] Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ 2015;350:g7773.
[15] Sargent RG. Verification, validation, and accreditation of simulation models. In: Joines JA, Barton RR, Kang K, Fishwick PA, editors. Proc. 2000 winter simul. conf; 2000. p. 50–9.
[16] Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM modeling good research practices task force-7. Med Decis Mak 2012;32:733–43.
[17] Feinleib M, Zelen M. Some pitfalls in the evaluation of screening programs. Arch Environ Health 1969;19(3):412–5.
[18] Independent UK panel on breast cancer screening. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380:1778–86.
[19] Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. preventive services task force. Ann Intern Med 2009;15:727–37.
[20] The Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in average-risk women aged 40–74 years. CMAJ 2011;183:1991–2001.
[21] WHO. Report of the Commission on Macroeceonomics and Health. Chaired by Jeffrey D. Sachs, Presented to Gro Harlem Brundtland, Director-General of the World Health Organization. December 2. Macroeconomics and health: investing in health for economic development, vol. 8; 2001., http://www.who.int/choice/costs/CER_thresholds/en/ [accessed 14.07.14].
[22] http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD [accessed 24.11.14].

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

[23] Van Oortmarssen GJ, Habbema JDF, van der Maas PJ, de Koning HJ, Collette HJA, Verbeek ALM, et al. A model for breast cancer screening. Cancer 1990;66:1601–12.
[24] Tan SYGL, van Oortmarssen GJ, de Koning HJ, Boer R, Habbema JDF. The MISCAN-Fadia continuous tumor growth model for breast cancer. J Natl Cancer Inst Monogr 2006;36:56–65.
[25] Mandelblatt J, Schechter CB, Lawrence W, Yi B, Cullen J. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. J Natl Cancer Inst Monogr 2006;36:47–55.
[26] Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin breast cancer epidemiology simulation model. J Natl Cancer Inst Monogr 2006;36:37–47.
[27] Plevritis SK, Sigal BM, Salzman P, Rosenberg J, Glynn P. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. J Natl Cancer Inst Monogr 2006;36:86–95.
[28] Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. J Natl Cancer Inst Monogr 2006;36:79–86.
[29] Berry DA, Inoue L, Shen Y, Venier J, Cohen D, Bondy M, et al. Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. J Natl Cancer Inst Monogr 2006;36:30–6.
[30] Van der Maas PJ, de Koning HJ, van Ineveld BM, van Oortmarssen GJ, Habbema JDF, Lubbe KTN, et al. The cost-effectiveness of breast cancer screening. Int J Cancer 1989;43:1055–60.
[31] Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comput Meth Prog Biomed 1984;20:79–93.
[32] Habbema JDF, Lubbe JTN, van Oortmarssen GJ, van der Maas PJ. A simulation approach to cost-effectiveness and cost–benefit calculations of screening for the early detection of disease. Eur J Oper Res 1987;29:159–66.
[33] Lee SJ, Zelen M. Modelling the early detection of breast cancer. Ann Oncol 2003;14:1199–202.
[34] Lee S, Huang H, Zelen M. Early detection of disease and scheduling of screening examinations. Stat Methods Med Res 2004;13:443–56.
[35] De Gelder R, Draisma G, Heijnsdijk EAM, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. Br J Cancer 2011;104:1214–20.
[36] De Gelder R, Fracheboud J, Heijnsdijk EAM, den Heeten G, Verbeek ALM, Broeders MJM, et al. Digital mammography screening: weighing reduced mortality against increased overdiagnosis. Prev Med 2011;53:134–40.
[37] De Gelder R, Bulliard J-L, de Wolf C, Fracheboud J, Draisma G, Schopper D, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer 2009;45:127–38.
[38] Okonkwo QL, Draisma G, der Kinderen A, Brown ML, de Koning HJ. Breast cancer screening policies in developing countries: a cost-effectiveness analysis for India. J Natl Cancer Inst 2008;100:1290–300.
[39] Rijnsburger AJ, van Oortmarssen GJ, Boer R, Draisma G, To T, Miller AB, et al. Mammography benefit in the Canadian national breast screening study-2: a model evaluation. Int J Cancer 2004;110:756–62.
[40] Van den Akker-van Marle E, de Koning H, Boer R, van der Maas P. Reduction in breast cancer mortality due to the introduction of mass screening in the Netherlands: comparison with the United Kingdom. J Med Screen 1999;6:30–4.
[41] Boer R, de Koning HJ, van der Maas PJ. A longer breast carcinoma screening interval for women age older than 65 years? Cancer 1999;86:1506–10.
[42] Boer R, de Koning H, Threlfall A, Warmerdam P, Street A, Friedman E, et al. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. BMJ 1998;317:376–9.
[43] Borràs JM, Espinas JA, Beemsterboer PMM, Granados A, de Koning HJ. Anticipating the consequences for the primary therapy of breast cancer after introducing screening. A more global picture for health care policy making. Int J Technol Assess Health Care 1998;14:268–76.
[44] Warmerdam PG, de Koning HJ, Boer R, Beemsterboer PMM, Dierks ML, Swart E, et al. Quantitative estimates of the impact of sensitivity and specificity in mammographic screening in Germany. J Epidemiol Commun Health 1997;51:180–6.
[45] Van den Akker-van Marle ME, Reep-van den Bergh CMM, Boer R, Del Moral A, Ascunce N, de Koning HJ. Breast cancer screening in Navarra: interpretation of a high detection rate at the first screening round and a low rate at the second round. Int J Cancer 1997;73:464–9.
[46] Vanara F, Zappa M, del Turco MR, Segnan N, Paci E, Ponti A. Cost–benefit analysis of a mammography screening program extended to all the national territory. Epidemiol Prev 1997;21:118–28.
[47] De Koning HJ, Boer R, Warmerdam PG, Beemsterboer PMM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995;87:1217–23.
[48] Boer R, de Koning HJ, van Oortmarssen GJ, van der Maas PJ. In search of the best upper age limit for breast cancer screening. Eur J Cancer 1995;31A:2040–3.
[49] Paci E, Boer R, Zappa M, de Koning HJ, van Oortmarssen GJ, Crocetti E, et al. A model-based prediction of the impact on reduction in mortality by a breast cancer screening programme in the city of Florence, Italy. Eur J Cancer 1995;31A:348–53.
[50] Beemsterboer PMM, de Koning HJ, Warmerdam PG, Boer R, Swart E, Dierks ML, et al. Prediction of the effects and costs of breast-cancer screening in Germany. Int J Cancer 1994;58:623–8.
[51] Van Ineveld BM, van Oortmarssen GJ, de Koning HJ, Boer R, van der Maas PJ. How cost-effective is breast cancer screening in different EC countries? Eur J Cancer 1993;29A:1663–8.
[52] De Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JCJM, Collette HJA, Hendriks JHCL, et al. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991;49:531–7.
[53] De Haes JCJM, de Koning HJ, van Oortmarssen GJ, van Agt HME, de Bruyn AE, van der Maas PJ. The impact of a breast cancer screening programme on quality-adjusted life-years. Int J Cancer 1991;49:538–44.
[54] Van Ravesteyn NT, Miglioretti DL, Stout NK, Lee SJ, Schechter CB, Buist DSM, et al. What level of risk tips the balance of benefits and harms to favor screening mammography starting at age 40? Ann Intern Med 2012;156:609–17.
[55] Mandelblatt JS, Cronin KA, Berry DA, Chang Y, de Koning HJ, Lee SJ, et al. Modeling the impact of population screening on breast cancer mortality in the United States. Breast 2011;20(Suppl. 3):S75–81.
[56] Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738–47.
[57] Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Extermann M, et al. Toward optimal screening strategies for older women. Costs, benefits, and harms of breast cancer screening by age, biology, and health status. J Gen Intern Med 2005;20:487–96.
[58] Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Muennig P, et al. Benefits and costs of interventions to improve breast cancer outcomes in African American women. J Clin Oncol 2004;22:2554–66.
[59] Hoerger TJ, Ekwueme DU, Miller JW, Uzunangelov V, Hall IJ, Segel J, et al. Estimated effects of the national breast and cervical cancer early detection program on breast cancer mortality. Am J Prev Med 2011;40:397–404.
[60] Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Román R, et al. Cost-effectiveness of early detection of breast cancer in Catalonia (Spain). BMC Cancer 2011;11:192–203.
[61] Rue M, Vilaprinyo E, Lee S, Martinez-Alonso M, Carles M, Marcos-Gragera R, et al. Effectiveness of early detection on breast cancer mortality reduction in Catalonia (Spain). BMC Cancer 2009;9:326–37.
[62] Beemsterboer PMM, Warmerdam PG, Boer R, de Koning HJ. Radiation risk of mammography related to benefit in screening programmes: a favourable balance? J Med Screen 1998;5:81–7.
[63] Carter R, Glasziou P, van Oortmarssen G, de Koning H, Stevenson C, Salkeld G, et al. Cost-effectiveness of mammographic screening in Australia. Aust J Public Health 1993;17:42–50.
[64] Bailey SL, Sigal BM, Plevritis SK. A simulation model investigating the impact of tumor volume doubling time and mammographic tumor detectability on screening outcomes in women aged 40–49 years. J Natl Cancer Inst 2010;102:1263–71.
[65] Mandelblatt J, van Ravesteyn N, Schechter C, Chang Y, Huang AT, Near AM, et al. Which strategies reduce breast cancer mortality most? Collaborative modeling of optimal screening, treatment, and obesity prevention. Cancer 2013;119:2541–8.
[66] Tosteson ANA, Stout NK, Fryback DG, Acharyya S, Herman B, Hannah L, et al. Cost effectiveness of digital mammography breast cancer screening: results from ACRIN DMIST. Ann Intern Med 2008;148:1–10.
[67] Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst 2006;98:774–82.
[68] Soerjomataram I, Louwman MWJ, Ribot JG, Roukema JA, Coebergh JWW. An overview of prognostic factors for long-term survivors of breast cancer. Breast Cancer Res Treat 2008;107:309–30.
[69] Soerjomataram I, Pukkala E, Brenner H, Jan Willem W, Coebergh JWW. On the avoidability of breast cancer in industrialized societies: older mean age at first birth as an indicator of excess breast cancer risk. Breast Cancer Res Treat 2008;111:297–302.
[70] Vemer P, Krabbe PFM, Feenstra TL, van Voom GAK, Ramos IC, Ramos C. Improving model validation in health technology assessment: comments on guidelines of the ISPOR-SMDM modeling good research practices task force. Value Health 2013;16:1106–7.
[71] Melnikow J, Tancredi DJ, Yang Z, Ritley D, Jiang Y, Slee C, et al. Program-specific cost-effectiveness analysis: breast cancer screening policies for a safety-net program. Value Health 2013;16:932–41.
[72] González-Mariño MA. Evaluating the usefulness of a breast screening program in Bogotá, Colombia. Rev Salud Publica 2012;14:41–52.
[73] Zhang F, Luo LM, Bao XD, Chen B. Cost-effectiveness analysis of mammography screening for Chinese women. J Tumor 2012;32:440–7.
[74] Sato M, Kawai M, Nishino Y, Shibuya D, Ohuchi N, Ishibashi T. Cost-effectiveness analysis for breast cancer screening: double reading versus single + CAD reading. Breast cancer 2014;5:532–41.
[75] Wong IO, Tsang JW, Cowling BJ, Leung GM. Optimizing resource allocation for breast cancer prevention and care among Hong Kong Chinese women. Cancer 2012;118:4394–403.
[76] Zamora LI, Forastero C, Guirado D, Martinez-Luna RJ, Lallena AM. A Monte Carlo tool to study the mortality reduction due to breast screening programs. Med Phys 2012;39:7215–23.

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013

[77] Peregrino AA, Vianna CM, de Almeida CE, Gonzáles GB, Machado SC, Costa e Silva FV, et al. Analysis of cost-effectiveness of screening for breast cancer with conventional mammography, digital and magnetic resonance imaging. Cien Saude Colet 2012;17:215–22.
[78] 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.
[79] Taylor P. Modelling the impact of changes in sensitivity on the outcomes of the UK breast screening programme. J Med Screen 2010;17:31–6.
[80] Madan J, Rawdin A, Stevenson M, Tappenden P. A rapid-response economic evaluation of the UK NHS cancer reform strategy breast cancer screening program extension via a plausible bounds approach. Value Health 2010;13:215–21.
[81] Valencia-Mendoza A, Sánchez-González G, Bautista-Arredondo S, Torres-Mejía G, Bertozzi SM. Cost-effectiveness of breast cancer screening policies in Mexico. Salud Publica Mex 2009;51:s296–304.
[82] Ahern CH, Shen Y. Cost-effectiveness analysis of mammography and clinical breast examination strategies: a comparison with current guidelines. Cancer Epidemiol Biomarkers Prev 2009;18:718–25.
[83] Rojnik K, Naversnik K, Mateović-Rojnik T, Primiczakelj M. Probabilistic cost-effectiveness modeling of different breast cancer screening policies in Slovenia. Value Health 2008;11:139–48.
[84] Wong IO, Kuntz KM, Cowling BJ, Lam CL, Leung GM. Cost effectiveness of mammography screening for Chinese women. Cancer 2007;110:885–95.
[85] Woo PP, Kim JJ, Leung GM. What is the most cost-effective population-based cancer screening program for Chinese women? J Clin Oncol 2007;25:617–24.
[86] Neeser K, Szucs T, Bulliard JL, Bachmann G, Schramm W. Cost-effectiveness analysis of a quality-controlled mammography screening program from the Swiss statutory health-care perspective: quantitative assessment of the most influential factors. Value Health 2007;10:42–53.
[87] Hanin LG, Miller A, Zorin AV, Yakovlev AY. The University of Rochester model of breast cancer detection and survival. J Natl Cancer Inst Monogr 2006;36:66–78.
[88] Carter KJ, Castro F, Kessler E, Erickson BA. Simulation of begin and end ages for mammography screening. J Healthc Qual 2005;27:40–7.
[89] Shen Y, Parmigiani G. A model-based comparison of breast cancer screening strategies: mammograms and clinical breast examinations. Cancer Epidemiol Biomarkers Prev 2005;14:529–32.
[90] Fett MJ. Computer modeling of the Swedish two county trial of mammographic screening and trade offs between participation and screening interval. J Med Screen 2001;8:39–45.
[91] Michaelson JS, Halpern E, Kopans DB. Breast cancer: computer simulation method for estimating optimal intervals for screening. Radiology 1999;212:551–60.
[92] Rosenquist CJ, Lindfors KK. Screening mammography beginning at age 40 years: a reappraisal of cost-effectiveness. Cancer 1998;82:2235–40.
[93] Salzmann P, Kerlikowske K, Phillips K. Cost-effectiveness of extending screening mammography guidelines to include women 40 to 49 years of age. Ann Intern Med 1997;127:955–65.
[94] Garuz R, Forcén T, Cabasés J, Antoñanzas F, Trinxet C, Rovira J, et al. Economic evaluation of a mammography-based breast cancer screening programme in Spain. Eur J Pub Health 1997;7:68–76.
[95] Szeto KL, Devlin NJ. The cost-effectiveness of mammography screening: evidence from a microsimulation model for New Zealand. Health Policy 1996;38:101–15.
[96] Lindfors KK, Rosenquist CJ. The cost-effectiveness of mammographic screening strategies. JAMA 1995;274:881–4.
[97] Griffiths RI, Griffiths CB, Powe NR. Simulated lifetime costs of three types of employer-based, periodic, breast cancer screening programs for working-age women. Am J Health Promot 1994;9:137–46.
[98] Rosenquist CJ, Lindfors KK. Screening mammography in women aged 40–49 years: analysis of cost-effectiveness. Radiology 1994;191:647–50.
[99] Okubo I, Glick H, Frumkin H, Eisenberg JM. Cost-effectiveness analysis of mass screening for breast cancer in Japan. Cancer 1991;67:2021–9.
[100] Knox EG. Evaluation of a proposed breast cancer screening regimen. BMJ 1988;297:650–4.

Please cite this article in press as: Koleva-Kolarova RG, et al., Simulation models in population breast cancer screening: A systematic review, The Breast (2015), http://dx.doi.org/10.1016/j.breast.2015.03.013