Cost-effectiveness of Digital Breast Tomosynthesis in Population-based Breast Cancer Screening: A Probabilistic Sensitivity Analysis

Valérie D. V. Sankatsing, MSc • Karolina Juraniec, MSc • Sabine E. Grimm, PhD • Manuela A. Joore, PhD • Ruud M. Pijnappel, MD, PhD • Harry J. de Koning, MD, PhD • Nicolien T. van Ravesteyn, PhD

From the Department of Public Health, Erasmus MC, University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, the Netherlands (V.D.V.S., K.J., H.J.d.K., N.T.v.R.); Department of Clinical Epidemiology and Medical Technology Assessment, School for Public Health and Primary Care (CAPHRI), Maastricht, the Netherlands (S.E.G., M.A.J.); Dutch Reference Center for Screening, Nijmegen, the Netherlands (R.M.P.); and Department of Radiology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands (R.M.P.). Received November 11, 2019; revision requested December 12; revision received May 22, 2020; accepted June 5. Address correspondence to V.D.V.S. (e-mail: v.sankatsing@erasmusmc.nl).

Supported by ZonMw and the National Evaluation Team for Breast Cancer Screening. Supported in part by the National Cancer Institute as part of the Cancer Intervention and Surveillance Modeling Network (CISNET) (U01 CA199218).

Conflicts of interest are listed at the end of this article.

See also the editorial by Slanetz in this issue.

Radiology 2020; 297:40–48 • <https://doi.org/10.1148/radiol.2020192505> • Content code: BR

Background: Digital breast tomosynthesis (DBT) is a promising screening test, but its outcomes and cost-effectiveness remain uncertain.

Purpose: To determine if biennial DBT is cost-effective in a screening setting, when compared with digital mammography (DM) in the Netherlands, and to quantify the uncertainty.

Materials and Methods: In this study, performed from March 2018 to February 2019, the Microsimulation Screening Analysis model was used to conduct a probabilistic sensitivity analysis (PSA), consisting of 10 000 model runs with 1 000 000 women simulated per run. The Bayesian Cost-Effectiveness Analysis package and the Sheffield Accelerated Value of Information tool were used to process PSA outcomes. Two simulated cohorts born in 1970 were invited to undergo biennial screening between ages 50 and 74 years—one cohort was assigned to DM screening, and one was assigned to DBT screening. DM input parameters were based on data from the Dutch breast cancer screening program. DBT parameters were based on literature and expert opinion. Willingness-to-pay thresholds of €20 000 (\$22 000) and €35 000 (\$38 500) per life-year gained (LYG) were considered. Effects and costs were discounted at 3.5% per year.

Results: DBT resulted in a gain of 13 additional life-years per 1000 women invited to screening (7% increase, 13 of 193), followed over lifetime, compared with DM and led to 2% (four of 159) fewer false-positive results. DBT screening led to incremental discounted lifetime effects of 5.09 LYGs (95% confidence interval: −0.80, 9.70) and an increase in lifetime costs of €137 555 (\$151 311) per 1000 women (95% confidence interval: €31 093 [\$34 202], €263 537 [\$289 891]) compared with DM, resulting in a mean incremental cost-effectiveness ratio of €27 023 (\$29 725) per LYG. The probability of DBT being more cost-effective was 0.36 at €20 000 and 0.66 at €35 000 per LYG.

Conclusion: Switching from digital mammography to biennial digital breast tomosynthesis is not cost-effective at a willingness-to-pay threshold of €20 000 per life-year gained, but digital breast tomosynthesis has a higher probability of being more cost-effective than digital mammography at a threshold of €35 000 per life-year gained.

© RSNA, 2020

Online supplemental material is available for this article.

Several European trials and observational studies have demonstrated that screening with digital breast tomosynthesis (DBT) as a stand-alone modality, combined with digital mammography (DM), with synthetic two-dimensional images reconstructed from the three-dimensional DBT acquisitions improves the breast cancer detection rate compared with DM alone (1–6). It is, however, not clear if DBT also improves sensitivity and the positive predictive value (PPV). Only a few studies have reported findings related to interval cancers, which are required to calculate the sensitivity of DBT (1,7). Also, the few reported sensitivity estimates were based on interval breast cancers in the first screening round only, as there were no long-term trial results available. Estimates of the PPV of DBT in a screening setting vary from similar to those of DM to a twofold higher PPV (1,3,5,6).

In considering a new screening modality, it is important to assess how its implementation would affect screening outcomes and cost-effectiveness. A model study from the United States showed that the addition of annual DBT screening to DM would be cost-effective compared with DM screening only, whereas another more recent study concluded that DBT screening would not be cost-effective at current reimbursement rates in the United States (8,9).

These conflicting results may occur owing to the unavailability of long-term DBT screening data, leading to uncertainty in model input parameters. Microsimulation models can be used to predict long-term screening outcomes and account for all uncertainty in input parameters through probabilistic sensitivity analysis (PSA). These models recalculate the output multiple times by incorporating

Abbreviations

DBT = digital breast tomosynthesis, DCIS = ductal carcinoma in situ, DM = digital mammography, ICER = incremental cost-effectiveness ratio, LYG = life-year gained, MISCAN = Microsimulation Screening ANalysis, PERT = program evaluation and review technique, PPV = positive predictive value, PSA = probabilistic sensitivity analysis, WTP = willingness to pay

Summary

Breast cancer screening with digital breast tomosynthesis compared with digital mammography was predicted to increase life-years gained by 7% and decrease false-positive results by 2% and was cost-effective at a threshold of €35 000 per life-year gained.

Key Results

- Probabilistic sensitivity analysis indicated that screening with digital breast tomosynthesis would lead to incremental discounted lifetime effects of 5 life-years gained and an increase in lifetime costs of €137 555 (\$151 311) compared with digital mammography per 1000 women.
- At a cost-effectiveness threshold of €35 000 (\$38 500) per life-year gained, tomosynthesis was the more cost-effective option compared with digital mammography in the majority of runs with use of the Microsimulation Screening ANalysis model.

probability distributions for input parameters and sampling from those distributions based on a large number of runs. From PSA outcomes, the generated cost-effectiveness acceptability curves show the probability that a strategy will be cost-effective, given a particular willingness-to-pay (WTP) threshold (10).

To our knowledge, no studies have conducted a PSA to assess the uncertainty associated with the cost-effectiveness of biennial DBT screening in a population-based breast cancer screening program. We used the Microsimulation Screening ANalysis (MISCAN) model to predict long-term benefit, harm, and costs associated with DBT versus DM in the Netherlands. By performing a PSA, we estimated the mean incremental cost-effectiveness of DBT versus DM and the probability that DBT would be the more cost-effective option for different WTP thresholds.

Materials and Methods

Model Overview

We used the MISCAN model, which has been described previously (11). MISCAN simulates individual life histories and the natural history of breast cancer as progression through preclinical screen-detectable (either preceded by ductal carcinoma in situ [DCIS] or not) T1A, T1B, T1C, and T2+ tumors. When screening is modeled, tumors can be detected in the preclinical phase. MISCAN models actual DM screening participation and therapy use (11). An overview of the model parameters is shown in Table E1 (online). MISCAN is exempt from institutional review board approval.

This study was performed from March 2018 to February 2019. Given the scarcity of DBT data at the time, we used expert elicitation to obtain estimates for three input parameters and their distributions—test sensitivity, PPV, and screening costs with DBT—according to guidelines (12,13). Experts

were asked to give estimates for a situation with stand-alone DBT. Expert elicitation is valid for use early in the development of new techniques when input parameters are not available from literature or are scarce or when published values are inconsistent (12,14).

Invited experts were researchers with extensive experience in breast cancer screening, 23 years on average, who were involved in DBT trials or affiliated with the Dutch Expert Centre for Screening. One expert, a professor of radiology with 20 years of experience in the field, is an author of this study (R.M.P.). Elicitation sessions were conducted by using the Excel-based EXPert eLICitation (abbreviated as EXPLICIT) tool (14). Nine of the 13 experts who were invited responded.

During the expert elicitation, two studies were published regarding sensitivity with DBT based on interval cancers from the first screening round: the Oslo Tomosynthesis Screening Trial, or OTST, and the Screening with Tomosynthesis or Standard Mammography, or STORM, trial (1,7). Estimates for the increase in sensitivity with DBT were 6% and 11%, respectively, although the difference between DBT and DM was not significant for either study. We decided not to dismiss the expert estimates but to pool them with the trial results instead. We did not use only trial results because (a) there was a lack of long-term data (only one screening round); (b) one trial (7) had a very low number of interval cancers with DBT, making the estimate less reliable; and (c) the use of combined DM and DBT was compared with DM alone. Combined DM and DBT considerably increases the radiation dose and acquisition and reading time; therefore, it might not be ideal to implement this combination in a screening program (15–19). Reconstruction of a two-dimensional image from DBT data seems like a good alternative (4–6,20).

The PPV of recall and screening costs with DBT were based on expert opinion only, as synthesis of the literature was difficult because published PPV estimates vary widely (1,3,5,6). Screening costs with DBT were also uncertain and therefore were based on expert elicitation only. Increased reading time (reading the stack of images and interpreting the synthetic two-dimensional images) would probably increase costs the most compared with DM, but studies on plausible DBT reading strategies are scarce.

Assumptions.—We assumed no difference in treatment rates between DM and DBT. For DM, we assumed that all women would undergo diagnostic mammography after recall. Because this is less likely after DBT, we assumed that only 30% of the women would undergo mammography after recall after DBT (21). Some studies have shown that the biopsy rate during multiple screening rounds is similar for DM and DBT (22), so we used the same frequencies and costs for other diagnostic tests. Finally, we assumed that switching from DM to DBT would increase sensitivity.

Costs.—Costs are reported in euros (€), with U.S. dollars (\$) in parentheses (€1 = \$1.10; May 19, 2020). Screening costs are based on data from the Dutch Screening Organizations (23) for DM or expert opinion for DBT. Costs of breast cancer diagnosis after recall and treatment are from a published Dutch study (Table E2 [online]) (24). Total costs with

screening are presented as net costs, compared with a situation without screening. All costs and effects are discounted at 3.5% per year.

Analysis

Two cohorts, each consisting of 500 000 women born in 1970, were simulated and followed during their lifetime, starting at age 50 years. It was simulated that cohort 1 was invited to biennial DM screening between ages 50 and 74 years (current practice). Cohort 2 was simulated to be invited to biennial DBT screening between ages 50 and 74 years, with the assumption that DBT would fully replace DM. For both scenarios, we assumed an attendance rate of 80%, reflecting actual participation over 2013–2018 in the Netherlands (23).

Benefits and harms.—Undiscounted benefits and harms were based on a single run and included life-years gained (LYGs), breast cancer deaths averted, and false-positive findings—that is, screening examinations that led to further diagnostics but not to breast cancer diagnosis, calculated by using the PPV of recall. The PPV was defined as the percentage of breast cancers detected, among all recalls.

PSA method.—Input parameters were varied simultaneously in 10 000 runs, sampling from fixed probability distributions, generating 10 000 incremental cost-effect pairs for DBT versus DM (Tables 1, 2). In each run, 1 million women were simulated. In MISCAN, model parameters are calibrated with the Nelder-Mead algorithm (25), which does not provide distributions around calibrated values. However, we could not obtain uncertainty measures from the literature for most calibrated parameters. To account for uncertainty, we used a previously described approach (Appendix E1 [online]) (26,27).

For parameters not calibrated, distributions were chosen on the following bases: (a) unobservable parameters have a uniform distribution and vary by 5% around the mode (fixed value); (b) parameters representing costs or PPVs have a beta-program evaluation and review technique (PERT) distribution and vary by 10%; and (c) bounds for DBT parameters, based on expert opinion, have a beta-PERT distribution (14). The beta-PERT distribution is a version of the beta distribution (interval [0, 1]) defined by the minimum, most likely, and maximum values of a parameter and was used because expert estimates were defined by the lowest, most likely, and highest parameter values. Individual expert estimates were combined through unweighted linear pooling by averaging the lowest, highest, and most likely values, respectively, generating an average most likely value and distribution (13). We assumed that sensitivity of DBT is higher than that of DM (1,7). Therefore, we induced correlation between sensitivity of DBT and DM, after sampling, by using the method of Goldhaber-Fiebert and Jalal (28). We adopted a tolerance threshold of 5% to allow for some ordering violation to occur.

PSA outcomes were mean incremental costs and effects of DBT versus DM, the mean incremental cost-effectiveness ratio (ICER), and mean incremental net benefit. The ICER was calculated by dividing the difference in costs between DBT

and DM by the difference in LYGs. A screening scenario was considered cost-effective if the ICER was below the WTP threshold. The common Dutch WTP threshold is €20 000 (\$22 000) per (quality-adjusted) LYG (29), generally used with 4% annual discounting of costs and 1.5% of effects. However, we used 3.5% discounting of both costs and effects because this is recommended by the National Institute for Health and Care Excellence (30), which uses a higher threshold range of £20 000–£30 000 per LYG and thus a maximum of around €35 000 (\$38 500) per LYG. We also used this threshold to account for a more internationally representative WTP threshold and because the WTP threshold of €20 000 per LYG has been argued to be low.

The mean incremental net benefit of DBT versus DM was calculated as the difference between the expected net benefit (ENB) of DBT and DM:

$$\text{ENB} = \text{WTP LYG} - \text{Cost}$$

$$\text{INB} = \text{ENB}_{\text{DBT}} - \text{ENB}_{\text{DM}}$$

Cost-effectiveness acceptability curves were calculated by selecting the screening modality with the highest expected net benefit in each of the 10 000 runs. Subsequently, we determined the proportion of runs for which the modality has the highest net benefit. The curves for DBT and DM show the probability of each being cost-effective, given a particular WTP threshold (10).

Scenario Analyses

To account for different DBT scenarios we performed one-way scenario analyses, based on a single run. First, the percentage of diagnostic mammography after recall from DBT (30% in base case) was varied between 0% and 100%. Second, biopsy after recall was assumed to be 38% higher with DBT than with DM (21,31). Further, attendance with DBT was varied between 70% and 90% but was kept constant with DM, based on observed rates. Finally, PPV of recall with DBT was simulated to be 50% higher than with DM (32). DBT parameters were varied by using the boundaries from the distribution obtained from expert opinion.

Statistical Analysis

MISCAN was used for the cost-effectiveness analysis with 10 000 iterations. The Sheffield Accelerated Value of Information tool (33) and the Bayesian Cost-Effectiveness Analysis package (34) were used to postprocess the results of the 10 000 runs, to obtain mean incremental costs and effects with confidence intervals and the ICER, and to assess uncertainty in constructing cost-effectiveness planes and acceptability curves. We used R, version 3.4.1 (R Foundation for Statistical Computing, Vienna, Austria) for the Bayesian cost-effectiveness analysis package and the Sheffield Accelerated Value of Information tool, version 2.1.2 (School of Health and Related Research, Sheffield, England) (R Shiny Server application) (33,34).

Table 1: Input Parameter Values, Distribution Types, and Boundaries for PSA

Parameter	Distribution	Calibrated Value or Mode (fixed value)	Minimum	Maximum
Calibrated input parameters				
Stage-specific sensitivity of digital mammography				
DCIS	Log normal	0.87	0.69	1
T1A	Log normal	0.55	0.43	0.70
T1B	Log normal	0.48	0.38	0.61
T1C	Log normal	0.86	0.66	1
T2+	Log normal	1	0.76	1
Breast cancer onset	Beta	0.291	0.279	0.305
Onset hazard by age				
30 y	Log normal	7.049×10^{-5}	3.222×10^{-5}	1.348×10^{-5}
50 y	Log normal	0.010	0.007	0.014
70 y	Log normal	0.019	0.014	0.026
100 y	Log normal	0.024	0.018	0.032
Stage-specific duration* (y) screen-detectable preclinical stage				
DCIS	Log normal	3.5	2.85	4.37
T1A	Log normal	0.1	0.08	0.12
T1B	Log normal	0.5	0.39	0.63
T1C	Log normal	1.5	1.21	1.85
T2+	Log normal	0.9	0.70	1.16
Probability of DCIS progressing into next stage (T1A)				
Immediately	Beta	0.842	0.415	0.996
Slowly	Beta	0.038	0.01	0.09
Noncalibrated input parameters				
Relative sensitivity [†] of younger ages, digital mammography	Beta-Pert	0.75	0.60	0.90
Relative sensitivity [†] of younger ages, tomosynthesis	Beta-Pert	0.85	0.70	1
Factor to vary, period-specific increase in onset of breast cancer	Uniform	Original [‡] value	10% lower	10% higher
Factor to vary, stage- and age-specific probability of transition node negative to positive	Uniform	Original value	5% lower	5% higher
Factor to vary, stage- and age-specific probability of diagnosis	Uniform	Original value	5% lower	5% higher
Relative improvement in prognosis, screen vs clinical detection <70 y	Uniform	Original value	10% lower	10% higher
Relative improvement in prognosis, screen vs clinical detection >70 y	Uniform	Original value	10% lower	10% higher
Costs of late-stage breast cancer	Beta-PERT	€18 000 (\$19,800)	€16,200 (\$17,820)	€19,800 (\$21,780)
Stage-specific treatment costs				
DCIS	Beta-PERT	€4569 (\$5026)	€4112 (\$4523)	€5026 (\$5529)
T1A, N–	Beta-PERT	€4333 (\$4766)	€3900 (\$4290)	€4766 (\$5243)
T1B, N–	Beta-PERT	€5057 (\$5563)	€4551 (\$5006)	€5563 (\$6119)
T1C, N–	Beta-PERT	€11 146 (\$12 261)	€10 031 (\$11 034)	€12 261 (\$13 487)
T2+, N–	Beta-PERT	€10 815 (\$11 897)	€9 734 (\$10 707)	€11 897 (\$13 087)
T1A, N+	Beta-PERT	€16 103 (\$17 713)	€14 493 (\$15 942)	€17 713 (\$19 484)
T1B, N+	Beta-PERT	€6 744 (\$7 418)	€6 070 (\$6 677)	€7 418 (\$8 160)
T1C, N+	Beta-PERT	€20 822 (\$22 904)	€18 740 (\$20 614)	€22 904 (\$25 194)
T2+, N+	Beta-PERT	€15 063 (\$16 569)	€13 557 (\$14 913)	€16 569 (\$18 226)
Positive predictive value				
Recall with digital mammography screening (%)	Beta-PERT	28	27	29
Biopsy (%)	Beta-PERT	70	64	75
Recall by a general practitioner (%)	Beta-PERT	58.3	52.5	64

Note.—DCIS = ductal carcinoma in situ, N+ = lymph node positive, N– = lymph node negative, PERT = program evaluation and review technique, PSA = probabilistic sensitivity analysis.

* Original value differs by stage and age.

[†] The relative sensitivity is 1 for age 55 years and older, 0.85 for younger than 50 years, and interpolated between 0.85 and 1 for ages 50–55 years.

[‡] Values for the stage-specific durations represent durations for women aged 50 years (in the model, different durations for women aged 40, 50, 60, and 70 years are used).

Table 2: Input Parameter Values, Distribution Types, and Boundaries for PSA for Parameters Based on Expert Opinion

Input Parameter Based on Expert Opinion	Distribution	Mode (Fixed Value)	Minimum	Maximum
Stage-specific sensitivity of tomosynthesis				
DCIS	Beta-PERT	1	0.92	1
T1A	Beta-PERT	0.65	0.58	0.72
T1B	Beta-PERT	0.57	0.51	0.63
T1C*	Beta-PERT	1	0.91	1
Positive predictive value of recall with tomosynthesis screening (%)	Beta-PERT	30	23	37
Screening costs, tomosynthesis	Beta-PERT	€91 (\$100)	€72 (\$79)	€125 (\$138)

Note.—DCIS = ductal carcinoma in situ, PERT = program evaluation and review technique, PSA = probabilistic sensitivity analysis.

* T2+ was not varied in the probabilistic sensitivity analysis, as the lower limit of the distribution was already 1 (100%).

Results

Expert Opinion

On the basis of the pooled results from the expert elicitation and the trials, test sensitivity was estimated to increase by 18% in switching from DM to DBT screening (distribution range, 6%–31%) (Table 3) (1,7). The pooled expert estimates were a PPV with DBT of 30% (range, 23%–37%) and DBT screening costs of €91 (\$100) (range, €72 [\$79]–€125 [\$138]), respectively.

Undiscounted Benefits and Harms

Biennial DBT screening between 50 and 74 years of age, compared with biennial DM, was expected to result in 13 additional LYGs (7% increase, 13 of 193) and 0.9 additional breast cancer death averted (6% increase, 0.9 of 14.5) per 1000 women invited to screening followed over their lifetime (Table 4). There were four fewer false-positive findings predicted per 1000 women invited to screening with DBT (2.5% decrease, four of 159) compared with DM.

Probabilistic Sensitivity Analysis

On the basis of 10 000 model runs, the mean lifetime incremental cost of DBT versus DM was €137 555 (\$151 311) per 1000 women invited (95% confidence interval: €31 093 [\$34 202], €263 537 [\$289 891]) and the mean incremental discounted lifetime effect was 5.09 LYGs per 1000 women invited (95% confidence interval: –0.80 to 9.70). The mean ICER of DBT versus DM was estimated at €27 023 (\$29 725) per LYG (Table 4). The uncertainty around the ICER estimate is represented by the cost-effectiveness plane (Fig 1a, 1b). The majority of the incremental cost-effect pairs of DBT, relative to DM, fell within the northeast quadrant of the cost-effectiveness plane, but a small number fell within the northwest quadrant. This indicates that DBT is more costly than DM and, for a high proportion of the iterations, more effective (94.7%).

Figure 2 shows the cost-effectiveness acceptability curves for DBT and DM. At a WTP threshold of €20 000 (\$22 000) per LYG, DBT had a low probability of being (most) cost-effective (0.36) and a negative incremental net benefit of –35 749 compared with DM (Table 4). At €35 000 (\$38 500) per LYG, the incremental net benefit was positive (40 606) and the probability of cost-effectiveness was 0.66, which means that DBT had

the highest expected net benefit in 66% of the 10 000 iterations (Fig 2).

One-Way Scenario Analyses

The ICER slightly ranged between €27 281 and €30 710 per LYG (\$30 009–\$33 782) when the percentage of diagnostic mammography after recall from DBT, the biopsy rate with DBT, or the PPV of DBT was varied (Fig E1 [online]). Ranging attendance with DBT resulted in considerable changes in the ICER (€51 898 [\$57 088] per LYG with 70% attendance vs €19 722 [\$21 694] with 90% attendance). Substantial ranges in the ICER of €1276–€76 685 (\$1404–\$84 353) and €22 841–€96 745 (\$25 125–\$106 419) were found when screening costs were varied between €72 and €125 (\$79 and \$138) and the increase in sensitivity with DBT between 31% and 6%, respectively (Fig E1 [online]).

Discussion

Because only short-term outcomes of screening with digital breast tomosynthesis (DBT) have been published, we predicted long-term costs and effects of DBT in a population-based screening setting; we performed a probabilistic sensitivity analysis to account for the uncertainty in model parameters. With current data, biennial DBT screening between 50 and 74 years of age is expected not to be cost-effective, considering the Dutch willingness-to-pay threshold (WTP) of €20 000 (\$22 000) per life-year gained (LYG) and 3.5% annual discounting. However, Dutch discounting recommendations (1.5% annually for effects, 4% for costs) would likely change this conclusion (12). Internationally, a threshold of €20 000 per LYG is rather low, which also puts the conclusion into a different perspective. Switching to biennial DBT in the Netherlands is likely to be cost-effective when a higher threshold of €35 000 (\$38 500) per LYG is used, as DBT was most cost-effective in 66% of the iterations with this threshold. In converting our findings to U.S. dollars and with a WTP threshold of \$100 000, biennial screening with DBT between 50 and 74 years of age would be cost-effective. However, to provide accurate estimates for the U.S. setting, U.S. data on demography, breast cancer incidence and screening (for example recall rates) would be required.

In agreement with a study by Lowry et al (9), we found that the incremental ICER was most sensitive to costs of DBT. Our

Table 3: Individual Estimates of Nine Experts and Two Trials and Unweighted Linearly Pooled Value

Source	Increase in Sensitivity			PPV, DBT			Screening Cost, DBT (€)		
	Most Likely	Lowest	Highest	Most Likely	Lowest	Highest	Most Likely	Lowest	Highest
Expert 1	50	5	80	30	23	45	95	70	130
Expert 2	10	5	25	32	22	38	95	75	105
Expert 3	15	5	25	30	25	35	130	70	200
Expert 4	15	5	30	35	30	45	90	80	150
Expert 5	8	4	15	30	26	35	80	70	90
Expert 6	13	4	23	24	15	28	85	75	150
Expert 7	30	10	40	28	20	35	75	67	90
Expert 8	20	0	50	30	25	35	75	70	85
Expert 9	22	15	25	30	25	35	...	...	...
OTST	5	1	11	...	...	...	...	...	...
STORM	11	7	12	...	...	...	...	...	...
Pooled value	18	6	31	30	23	37	91 (\$100)	72 (\$79)	125 (\$138)

Note.—Unless otherwise specified, data are percentages. Pooled values reflect the mean of all estimates; ellipses indicate that an estimate was not given or, in case of the trials, that these data were not used. Values for increase in sensitivity are based on estimates of nine experts and the estimates of the OTST (15) and STORM trial (16). DBT = digital breast tomosynthesis, OTST = Oslo Tomosynthesis Screening Trial, PPV = positive predictive value, STORM = Screening with Tomosynthesis or Standard Mammography.

Table 4: Long-term Screening Outcomes and Results of PSA per 1000 Women Invited to Screening

Outcome	DM	DBT	Incremental*
Screening outcomes (single run)			
Life-years gained [†]	179.9	192.5	12.6
Breast cancer deaths averted [†]	13.6	14.5	0.9
False-positive findings	158.8	155.2	−3.6
PSA outcomes (3.5% discounted)			
Mean costs [†]	€305 829 (\$336 412)	€443 385 (\$487 724)	€137 555 (\$151 311)
Mean effect (life-years gained) [†]	67	73	5
Mean ICER (DBT vs DM)	...	...	€27,023 (\$29 725)
ENB (WTP threshold = €20 000)	1 043 980	1 008 231	−35 749
ENB (WTP threshold = €35 000)	2 056 337	2 096 943	40 606

Note.—Women were invited to biennial screening between 50 and 74 years of age and were followed during their lifetimes. DBT = digital breast tomosynthesis, DM = digital mammography, ENB = expected net benefit, ICER = incremental cost-effectiveness ratio, INB = incremental net benefit, PSA = probabilistic sensitivity analysis, WTP = willingness-to-pay.

* DBT relative to DM.

[†] Relative to no screening.

results were that at a WTP threshold of €35 000 per LYG, biennial DBT between 50 and 74 years of age was more cost-effective than DM in the majority of runs. With use of a higher WTP threshold of \$100 000, Kalra et al (8) showed that in the United States, annual DBT starting at 40 years of age was also more cost-effective than DM in the majority of runs.

Some DBT trials and observational studies have shown lower false-positive rates with use of DBT compared with DM, whereas others have suggested higher rates (4,15,35,36). In contrast to a study from the United States, which demonstrated a 24%–28% reduction in false-positive results, we found only a slight decrease because we assumed a minor increase in the PPV with DBT relative to DM (9). Recall and false-positive rates with DM in the Netherlands are relatively low compared with other countries; therefore, a reduction when switching to DBT may be less likely than

in other settings. Correspondingly, a meta-analysis evaluating recall rates with DBT compared with DM showed that reductions in recall rates with DBT were predominantly found in studies from the United States that reported high baseline recall (37). The decrease in recall rates and the number of false-positive results with DBT thus depends on the initial rate with DM.

Prospective studies have shown that DBT increases breast cancer detection compared with DM (1–3,5), which could contribute to overdiagnosis of breast cancer. Recent literature shows that DBT enables detection of a higher proportion of invasive cancers than DM, but it may be too soon to assess the effect on interval cancer rates (31,38).

Our findings show the impact of DBT compared with DM in a European screening setting, supplementing earlier published outcomes for the screening setting in the United States (8,9).

Figure 1: Plots show cost-effectiveness plane of tomosynthesis versus mammography at willingness-to-pay thresholds of €20 000 (\$22 000) and €35 000 (\$38 500) per life-year gained (costs and effects per 1000 women invited to screening). Gray dots show simulated values for joint distribution of effectiveness and cost differentials; red dot indicates average of distribution of outcomes (ICER); gray area reflects part of plane for which simulated values are below willingness-to-pay threshold. ICER = incremental cost-effectiveness ratio, k = willingness-to-pay threshold.

Our study had some limitations. First, because trial estimates of DBT sensitivity were based on interval cancers of the first screening round only (1,7), and the performance of combined DBT and DM was compared with DM alone (1,2), the actual effect of DBT alone could be smaller than reported.

We did not account for increased radiation dose with DBT compared with DM. We also did not specify our results for different breast density levels, because there was not sufficient evidence from the trials that DBT performs better in dense breasts than in nondense breasts compared with DM (2).

Furthermore, we assumed no differences in frequency and costs of diagnostics after recall between DBT and DM. However, our scenario analyses showed that the diagnostic mammography and biopsy rate hardly influenced the ICER.

Finally, we relied on expert opinion for some DBT inputs, which may have induced bias. Still, our inputs fairly represent the latest literature. The Oslo Tomosynthesis Screening Trial (32) recently reported a 28% increase in sensitivity with combined DBT and synthetic two-dimensional imaging, compared with DM. In addition, an Italian screening pilot study found an increase of 11% (31). When these new estimates are averaged with the earlier published 11% increase from the Screening with Tomosynthesis or Standard Mammography trial (7), the average increase in sensitivity with DBT compared with DM is 17%. This is similar to the 18% increase we used for the PSA, based on the expert estimates and trial results at the time of our study. Our estimate for the PPV also falls within the range of published estimates (1,5,6,35). Furthermore, although DBT costs in the

Figure 2: Graph shows cost-effectiveness acceptability curves for digital breast tomosynthesis and digital mammography. Dotted lines represent willingness-to-pay thresholds of €20 000 (\$22 000) and €35 000 (\$38 500) per life-year gained.

Dutch screening setting are unclear, costs of diagnostic DBT in university medical centers in the Netherlands (€94.58 [\$104]) (39) are similar to the €91 [\$100]) used in our analysis.

In conclusion, analysis of currently available preliminary data, based on one screening round with digital breast tomosynthesis, demonstrated that switching from biennial digital mammography to digital breast tomosynthesis screening in the Netherlands would not be cost-effective with use of a willingness-to-pay threshold of €20 000 (\$22 000) per life-year gained. Switching to screening with digital breast tomosynthesis could be cost-effective at a threshold of €35 000 (\$38 500) per life-year gained.

Author contributions: Guarantors of integrity of entire study, V.D.V.S., H.J.d.K., N.T.V.R.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, V.D.V.S., K.J., S.E.G., R.M.P.; clinical studies, R.M.P.; statistical analysis, V.D.V.S., K.J., S.E.G., M.A.J.; and manuscript editing, V.D.V.S., S.E.G., M.A.J., R.M.P., H.J.d.K., N.T.V.R.

Disclosures of Conflicts of Interest: V.D.V.S. disclosed no relevant relationships. K.J. disclosed no relevant relationships. S.E.G. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: institution has a consultancy on a pharmaceutical reimbursement application submission with Janssen. Other relationships: disclosed no relevant relationships. M.A.J. disclosed no relevant relationships. R.M.P. disclosed no relevant relationships. H.J.d.K. disclosed no relevant relationships. N.T.V.R. disclosed no relevant relationships.

References

- Skaane P, Sebuødegård S, Bandos AI, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. *Breast Cancer Res Treat* 2018;169(3):489–496.
- Ciatto S, Houssami N, Bernardi D, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013;14(7):583–589.
- Lång K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. *Eur Radiol* 2016;26(1):184–190.
- Bernardi D, Macaskill P, Pellegrini M, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. *Lancet Oncol* 2016;17(8):1105–1113.
- Hofvind S, Hovda T, Holen AS, et al. Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program. *Radiology* 2018;287(3):787–794.
- Caumo F, Zorzi M, Brunelli S, et al. Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program. *Radiology* 2018;287(1):37–46.
- Houssami N, Bernardi D, Caumo F, et al. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. *Breast* 2018;38:150–153.
- Kalra VB, Wu X, Haas BM, Forman HP, Philpotts LE. Cost-Effectiveness of Tomosynthesis in Annual Screening Mammography. *AJR Am J Roentgenol* 2016;207(5):1152–1155.
- Lowry KP, Trentham-Dietz A, Schechter CB, et al. Long-term Outcomes and Cost-effectiveness of Breast Cancer Screening with Digital Breast Tomosynthesis in the United States. *J Natl Cancer Inst* 2020. 10.1093/jnci/djz184. Published online June 1, 2020.
- Barton GR, Briggs AH, Fenwick EA. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfect information (EVPI). *Value Health* 2008;11(5):886–897.
- Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteijn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in the Netherlands. *Int J Cancer* 2015;137(8):1990–1999.
- Bojke L, Grigore B, Jankovic D, Peters J, Soares M, Stein K. Informing Reimbursement Decisions Using Cost-Effectiveness Modelling: A Guide to the Process of Generating Elicited Priors to Capture Model Uncertainties. *Pharmacoeconomics* 2017;35(9):867–877.
- Iglesias CP, Thompson A, Rogowski WH, Payne K. Reporting Guidelines for the Use of Expert Judgement in Model-Based Economic Evaluations. *Pharmacoeconomics* 2016;34(11):1161–1172.
- Grigore B, Peters J, Hyde C, Stein K. EXPLICIT: a feasibility study of remote expert elicitation in health technology assessment. *BMC Med Inform Decis Mak* 2017;17(1):131.
- Skaane P, Bandos AI, Gullien R, et al. Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *Eur Radiol* 2013;23(8):2061–2071.
- Skaane P, Bandos AI, Gullien R, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 2013;267(1):47–56.
- Svahn TM, Houssami N, Sechopoulos I, Mattsson S. Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. *Breast* 2015;24(2):93–99.
- Tagliafico AS, Calabrese M, Bignotti B, et al. Accuracy and reading time for six strategies using digital breast tomosynthesis in women with mammographically negative dense breasts. *Eur Radiol* 2017;27(12):5179–5184.

19. Bernardi D, Ciatto S, Pellegrini M, et al. Application of breast tomosynthesis in screening: incremental effect on mammography acquisition and reading time. *Br J Radiol* 2012;85(1020):e1174–e1178.
20. Skaane P, Bandos AI, Eben EB, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiology* 2014;271(3):655–663.
21. Alsheik NH, Dabbous F, Pohlman SK, et al. Comparison of Resource Utilization and Clinical Outcomes Following Screening with Digital Breast Tomosynthesis Versus Digital Mammography: Findings From a Learning Health System. *Acad Radiol* 2019;26(5):597–605.
22. McDonald ES, Oustimov A, Weinstein SP, Synnestvedt MB, Schnall M, Conant EF. Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening. *JAMA Oncol* 2016;2(6):737–743.
23. National Evaluation Team Breast cancer screening. National evaluation of breast cancer screening in the Netherlands 2004-2014. Fourteenth evaluation report. Erasmus MC University Medical Center Rotterdam, Radboud University Medical Center Nijmegen 2019.
24. Saadatmand S, Tilanus-Linthorst MM, Rutgers EJ, et al. Cost-effectiveness of screening women with familial risk for breast cancer with magnetic resonance imaging. *J Natl Cancer Inst* 2013;105(17):1314–1321.
25. Nelder JA, Mead R. A simplex method for function minimization. *Comput J* 1965;7(4):308–313.
26. Wever EM, Draisma G, Heijnsdijk EA, et al. Prostate-specific antigen screening in the United States vs in the European Randomized Study of Screening for Prostate Cancer-Rotterdam. *J Natl Cancer Inst* 2010;102(5):352–355.
27. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender: estimates from the NLST and the PLCO trials. *Cancer Epidemiol Biomarkers Prev* 2015;24(1):154–161.
28. Goldhaber-Fiebert JD, Jalal HJ. Some Health States Are Better Than Others: Using Health State Rank Order to Improve Probabilistic Analyses. *Med Decis Making* 2016;36(8):927–940.
29. Zwaap JKS, van der Meijden C, Staal P, van der Heiden L. Rapport kosteneffectiviteit in de praktijk. Diemen, the Netherlands: Zorginstituut Nederland, 2015.
30. Appleby J, Devlin N, Parkin D. NICE's cost effectiveness threshold. *BMJ* 2007;335(7616):358–359.
31. Bernardi D, Gentilini MA, De Nisi M, et al. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: results of the Trento DBT pilot evaluation. *Breast* 2020;50:135–140.
32. Skaane P, Bandos AI, Niklason LT, et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial. *Radiology* 2019;291(1):23–30.
33. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. *Med Decis Making* 2014;34(3):311–326.
34. Baio G, Berardi A, Heath A. Bayesian Cost-Effectiveness Analysis with the R package BCEA. Cham, Switzerland: Springer International Publishing, 2017.
35. Lång K, Nergården M, Andersson I, Rosso A, Zackrisson S. False positives in breast cancer screening with one-view breast tomosynthesis: an analysis of findings leading to recall, work-up and biopsy rates in the Malmö Breast Tomosynthesis Screening Trial. *Eur Radiol* 2016;26(11):3899–3907.
36. Friedewald SM, Rafferty EA, Rose SL, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA* 2014;311(24):2499–2507.
37. Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall. *J Natl Cancer Inst* 2018;110(9):942–949.
38. Bahl M, Gaffney S, McCarthy AM, Lowry KP, Dang PA, Lehman CD. Breast Cancer Characteristics Associated with 2D Digital Mammography versus Digital Breast Tomosynthesis for Screening-detected and Interval Cancers. *Radiology* 2018;287(1):49–57.
39. Ziekenhuiskosten per ziekenhuis 2018. FBTO. <https://cdn.fbto.nl/-/media/documenten/zorgverzekering-ziekenhuiskosten-ziekenhuis-2018.pdf>. Accessed May 2020.