Review

Meta-analysis of prospective studies evaluating breast cancer detection and interval cancer rates for digital breast tomosynthesis versus mammography population screening

Nehmat Houssami [a,*], Sophia Zackrisson [b], Katrina Blazek [a], Kylie Hunter [c], Daniela Bernardi [d], Kristina Lång [b], Solveig Hofvind [e]

[a] *Sydney School of Public Health, Faculty of Medicine and Health, University of Sydney, Sydney, Australia*
[b] *Department of Translational Medicine, Diagnostic Radiology, Lund University, Skane University Hospital, Malmö, Sweden*
[c] *NHMRC Clinical Trials Centre, University of Sydney, Sydney, Australia*
[d] *Breast Imaging and Screening Unit, Department of Radiology, Humanitas Research Hospital, Milan, Italy*
[e] *Section of Breast Screening, Cancer Registry of Norway, Oslo, Norway*

Abstract *Introduction:* Breast cancer (BC) screening using digital breast tomosynthesis (DBT) has been shown to increase cancer detection compared with mammography; however, it is unknown whether DBT impacts *interval cancer rate* (ICR).

Methods: We systematically identified prospective DBT studies reporting data on screen-detected and interval BCs to perform a study-level meta-analysis of the comparative effect of DBT on ICR in population screening. Meta-analysis of cancer detection rate (CDR), ICR, and the differences between DBT and mammography in CDR and ICR pooled estimates, included random-effects. Sensitivity analysis examined whether study methods (imaging used, comparison group design, interval BC ascertainment) affected pooled estimates.

Results: Five eligible prospective (non-randomised) studies of DBT population screening reported on 129,969 DBT-screened participants and 227,882 mammography-only screens, including follow-up publications reporting interval BC data. Pooled CDR was 9.03/1000 (95% confidence interval [CI] 8.53—9.56) for DBT, and 5.95/1000 (95% CI 5.65—6.28) for mammography: the pooled *difference* in CDR was 3.15/1000 (95% CI 2.53—3.77), and was evident for the detection of invasive and in-situ malignancy. Pooled ICR was 1.56/1000 DBT screens (95% CI 1.22—2.00), and 1.75/1000 mammography screens (95% CI 1.46

* Corresponding author.
E-mail address: nehmat.houssami@sydney.edu.au (N. Houssami).

−2.11): the estimated pooled difference in ICR was −0.15/1000 (95% CI −0.59 to 0.29) and was not substantially altered in several sensitivity analyses.

Conclusions: Meta-analysis shows consistent evidence that DBT significantly increased CDR compared with mammography screening; however, there was little difference between DBT and mammography in pooled ICR. This could suggest, but does not demonstrate, some over-detection. Meta-analysis using individual participant data, randomised trials and comparative studies quantifying cumulative detection and ICR over repeat DBT screen-rounds would provide valuable evidence to inform screening programs.

1. Introduction

Breast cancer (BC) screening, using mammography, is widely implemented for population screening and has been shown to reduce BC mortality [1,2]. Digital breast tomosynthesis (DBT), a technology progressively adopted into breast imaging practice, has recently received conditional approval for BC screening in European guidelines, which also highlight evidence gaps regarding DBT's effect in screening [3]. DBT acquires multiple low-dose x-rays of the breast from varying angles, which are reconstructed into thin-slice images providing pseudo−three-dimensional imaging that reduces overlapping parenchyma and enhances detection of cancer [4−7]. Prospective and retrospective studies have shown that using DBT in addition or as replacement to mammography yields an incremental BC detection rate of 1.6/1000 screens compared with mammography [6].

Given that studies of DBT screening generally show higher *cancer detection rate* (CDR) compared with mammography screening, the question arises whether this will translate into additional screening benefit or whether the additional BC detection contributes to over-diagnosis. We previously outlined that BC mortality as an end-point requires very large studies with extended follow-up and may not be a feasible end-point in contemporary BC screening evaluation [8]. We proposed using a surrogate end-point, *interval cancer rate* (ICR) as an indicator of potential effectiveness [8]. As interval cancers are not detected at screening and usually emerge due to symptoms (before the next scheduled screen), and have a similar prognosis as clinically presenting BC, a reduction in interval cancers from a more sensitive screening technology would be expected to add screening benefit [8]. Furthermore, if a new screening test increases CDR (relative to standard screening) and has the effect of reducing ICR at follow-up, then it could be inferred that the additional cancer detection does not merely represent over-diagnosis.

We planned a collaborative individual participant data (IPD) meta-analysis to assess the impact of DBT screening on ICR in population-based screening; pooling data sets allows more precise estimation of this relatively infrequent end-point than individual studies. A preliminary study to the proposed IPD meta-analysis was performed as a *study-level* meta-analysis aiming to examine the effect of DBT versus mammography screening on ICR, based on *prospective* DBT studies that report outcomes for both screen-detected and interval cancers in population BC screening.

2. Methods

In April 2020, a systematic search was performed using MEDLINE for studies of DBT versus mammography population screening. Two investigators (NH and KH) screened titles and abstracts against pre-defined study eligibility criteria adapted from a published protocol [8]. Online Appendix-1 summarises the literature search, study selection process and details of the included [4,9−17] and excluded studies [18−27].

2.1. Study eligibility

Studies were eligible for inclusion in this meta-analysis if they met eligibility criteria adapted from an IPD meta-analysis protocol [8], given that the rationale and screening context for the study-level meta-analysis aligned with those of the planned IPD meta-analysis, as follows: studies of population BC screening that investigated DBT (interpreted alone, or with acquired or synthetic 2D-images) in comparison to digital mammography (DM); used a *prospective* design for DBT screening (prospective recruitment or inclusion of participants); reported data on BC detection (number of cancers, number of screens and/or CDR) *and* reported data on interval cancers (numbers or rates) *or* stated that interval cancer data would be reported at follow-up [8]; and used a predominantly *biennial* screening interval aligning with practice in organized population-based screening programs [8].

Studies were *not* eligible for inclusion if they used a retrospective design to investigate DBT screening, if they were based on annual screening, or if participants were classified as having increased BC risk (studies

selecting participants with specific risk factors) or were symptomatic populations. Studies using cancer-enriched imaging data sets (such as multi-reader studies) and case series were not eligible as these do not represent population screening.

2.2. Comparison cohort

As some prospective trials of DBT screening used a paired design [6], where screen-reading using DBT or mammography was compared *within-woman*, ICR from those studies would represent outcomes for participants screened with DBT (none of the participants would have received mammography-alone screening). For those studies, we sought updated publications providing interval cancer data for a comparison group screened with DM only, using our literature search and by directly checking with investigators of the DBT trials or IPD protocol [8].

2.3. Data extraction

Data were extracted from eligible publications by two investigators (NH and KH). Extracted data were *independently* checked by one investigator representing each of the eligible studies to verify the data or identify discrepancies (if any existed, these were resolved via email discussion and consensus). This approach was used because fully independent extraction was not feasible given that results of the studies were well-known to the authorship group, and several investigators had extracted (and were familiar with) data from these trials as part of an overview examining CDR [6].

The following variables were extracted: study characteristics (design, imaging used, timeframe, comparison cohort, aggregate-level age); number of screens (per imaging modality); reading protocol; cancer detection data (cancers detected, number in-situ or invasive, detection by modality *within* paired study design); and interval cancer data (number, number of screens with follow-up and ascertainment methods).

2.4. Study quality

We extracted information regarding study characteristics to guide interpretation of quality in the context of population screening, specifically: methods of the prospective DBT study (whether randomisation used); design of the comparison cohort (whether concurrent or historical group; whether from same screening program); and the methods used to ascertain interval cancers (whether population cancer registry checked). We detailed this information in a study-specific summary of methods and quality (Table 1; Online Appendix-2), and used it in sensitivity analyses to explore its effect on estimates.

2.5. Statistical methods

Study characteristics were summarised descriptively. Forest plots were used to display study-specific data (number of cancers and screens) and rates per 1000, and pooled estimates for CDR and ICR. Because studies were from different populations, we used random-effects models to allow for both within-study sampling variability and heterogeneity between studies when calculating pooled estimates. Pooled CDRs were also estimated and displayed in subgroups defined by the DBT imaging used. For the main analysis comparing *independent groups* (unpaired data), differences between DBT and mammography in CDR and ICR were pooled as risk differences using the Mantel−Haenszel method. For within-woman comparison of CDR (in studies reporting paired data), the difference was calculated as the pooled difference of the number detected by imaging modality in each study. Sensitivity analysis was used to examine the effect of the design of the comparison group, the DBT imaging used, and interval cancer ascertainment method, on estimates of the pooled difference between DBT and DM. Analyses used the *meta* package (version 4.11−0) [28] for R (version 3.6.3) [29].

3. Results

We identified five eligible studies of DBT population screening (*Appendix-1*) that reported cancer detection *and* interval cancer data at follow-up across 10 papers [4,9−17] (total 357,851 DBT or DM screens): these were prospective non-randomised trials, or prospective cohort studies, that investigated DBT in population-based screening programs in Europe. Characteristics of these studies are summarised in Table 1, including details of comparison groups screened with DM, showing some heterogeneity in methods. Additional study-specific methods are shown in flow-diagrams (online Appendix-2) highlighting the potential for confounding given that none of the studies randomised participants to DBT or DM screening. Three studies used DBT and DM acquisitions (which allowed within-woman comparison of screen-reading by imaging modality) [4,10,12,13] and two studies used DBT with synthetic 2D-images [15,17]. There were no randomised controlled trials reporting both CDR and interval cancer data for DBT screening.

3.1. Cancer detection rates

Collectively, the five studies [4,9−17] of DBT screening reported that 1174 women had screen-detected BC (1182 counting bilateral BC) amongst 129,969 participants who had DBT screening. For consistency, we report CDR counting one cancer per woman, so estimates may slightly differ from reports where bilateral cancer was

Table 1
Characteristics of eligible studies comparing digital breast tomosynthesis (DBT) and digital mammography (DM) population screening.

Study [publications from which data extracted]	Design	Timeframe	Age (median or mean)	Double-reading and recall process	Comparison group[a] screened with DM only		Interval cancer ascertainment and follow-up (same for comparison?)
					Design	Age	
STORM trial [Ciatto [4]; Houssami [9]]	Prospective non-randomised trial comparing screen-reading of DM with DM + DBT (paired data) in women ≥48 years	August 2011–June 2012	58 years	Independent double-read, recall if either reader recalls	Concurrent	Not reported	Regional cancer registry, and pathology and hospital databases; 2-year follow-up (yes)
Oslo trial (OTST) [Skaane [10]; Skaane [11]; Skaane [12]]	Prospective non-randomised trial comparing screen-reading using DM versus DM + DBT (paired data) in women 50–69 years at screening	November 2010–December 2012	59 years	Independent double-read, consensus (arbitration) meeting if discordant scores between readers	Historical	Not reported	Cancer Registry of Norway - Population cancer registry linkage; 2-year follow-up (yes)
Malmo trial (MBTST) [Zackrisson [13]; Johnson [14]]	Prospective non-randomised trial comparing screen-reading using one-view DBT with 2-view DM (paired data) in women >40 years	January 2010–February 2015	57 years	Independent double-read, consensus meeting if score >3 by either reader	Concurrent	53 years	Population cancer registry linkage; 2-year follow-up, 18 months in subgroup where this interval recommended (yes)
OVVV study [Hofvind [15]; Hovda [16]]	Prospective cohort study of DBT[b] screening versus concurrent (DM-screened) cohort in women 50–69 years at screening	February 2014–January 2016	59 years	Independent double-read, consensus meeting if discordant scores between readers	Concurrent	59 years	Cancer Registry of Norway Population cancer registry linkage; 2-year follow-up (yes)
Trento pilot study [Bernardi [17]]	Prospective cohort study of DBT[b] screening versus historical (DM-screened) cohort in women ≥50 years	October 2014–October 2016	58 years	Independent double-read, third read if discordant readings	Historical	58 years	Regional cancer registry, and pathology and hospital databases; 2-year follow-up (unclear)

STORM, screening with tomosynthesis or mammography; OTST, Oslo Tomosynthesis Screening Trial; MBTST, Malmo Breast Tomosynthesis Screening Trial; OVVV, Oslo, Vestfold & Vestre Viken study.

[a] Studies were not randomised—comparison groups were from the same population-based screening programs allowing comparison of interval cancer rates (OVVV compared different services within the Norwegian breast screening program).
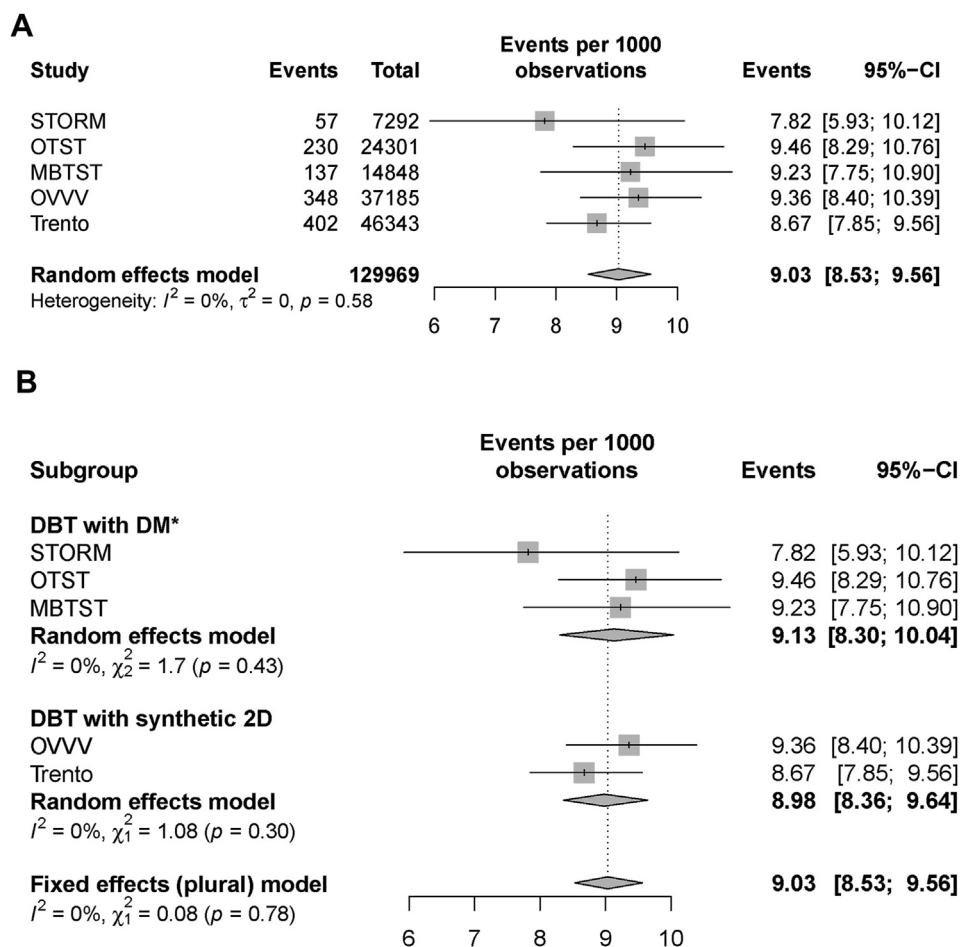
[b] DBT with synthetic 2D images reconstructed from the DBT acquisition.

counted. Fig. 1(a and b) shows study-specific, subgroup and overall pooled CDR for DBT counting all detected (invasive and in-situ) cancers: pooled CDR was 9.03 (95% CI 8.53−9.56) per 1000 screens. Subgroup estimates were similar for studies using DBT with DM, and those that used DBT with synthetic 2D-images (Fig. 1b). Fig. 2(a and b) shows study-specific detection data and pooled estimates for DBT screening for invasive cancer (pooled CDR 7.65/1000; 95% CI 7.19−8.14) and ductal carcinoma in-situ ((DCIS) pooled CDR 1.37/1000; 95% CI 1.13−1.68). In the subset of three studies [4,10,12,13] reporting within-woman comparison of DBT and DM detection, DBT's higher detection was evident in a pooled CDR *difference* of 2.22/1000 (95% CI 1.83−2.69) shown in online Appendix-3.

For the five studies, *comparison groups* screened with DM only had 1357 screen-detected cancers amongst 227,882 participants [4,9−13,15,17]: Fig. 3 shows study-specific CDR and a pooled CDR of 5.95/1000 (95% CI 5.65−6.28). In the subset of four studies [10−13,15,17] reporting data by cancer type, study-specific and pooled CDR for invasive cancer and DCIS for DM-screening are shown in online Appendix-3.

Comparison of DBT-screened and DM-screened co-horts (Fig. 4: unpaired data from independent groups) showed a significant pooled *difference* in CDR of 3.15/1000 screens (95%CI 2.53−3.77) favouring DBT. A sensitivity analysis including only the three studies that had concurrent comparison groups [9,13−16] gave a similar estimate for the *difference* in CDR between DBT and DM (pooled CDR difference: 3.11/1000; 95% CI 2.24−3.99). Another sensitivity analysis including only studies that used DBT with synthetic 2D-images [15,17] also gave a similar estimate (pooled CDR difference between DBT and DM: 3.21/1000 screens; 95% CI 2.40−4.02). In the four studies [10−13,15,17] reporting data by cancer type (Fig. 5(a and b)), pooled *differences* in CDR between DBT and DM were evident for the

Fig. 1. (A) DBT CDR. (B) DBT subgroup estimates of CDR. CDR, cancer detection rate; CI, confidence interval; DBT, digital breast tomosynthesis; DM, digital mammography; STORM, screening with tomosynthesis or mammography; OTST, Oslo Tomosynthesis Screening Trial; MBTST, Malmo Breast Tomosynthesis Screening Trial; OVVV, Oslo, Vestfold & Vestre Viken study.
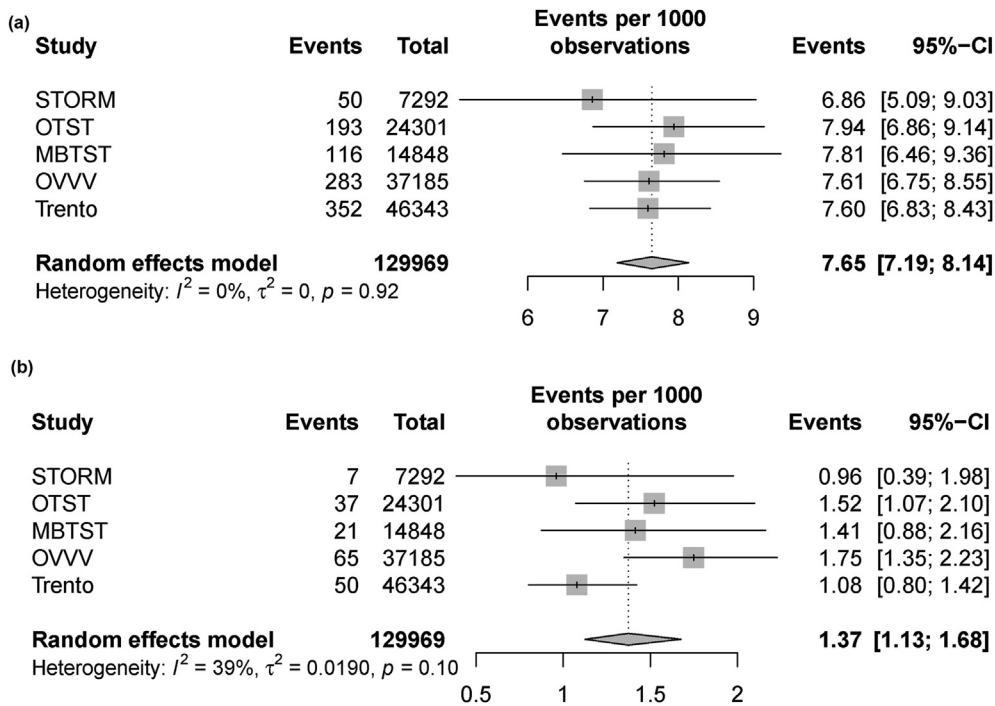
(a)

Study	Events	Total	Events per 1000 observations	Events	95%−CI
STORM	50	7292		6.86	[5.09; 9.03]
OTST	193	24301		7.94	[6.86; 9.14]
MBTST	116	14848		7.81	[6.46; 9.36]
OVVV	283	37185		7.61	[6.75; 8.55]
Trento	352	46343		7.60	[6.83; 8.43]
Random effects model		**129969**		**7.65**	**[7.19; 8.14]**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.92$

(b)

Study	Events	Total	Events per 1000 observations	Events	95%−CI
STORM	7	7292		0.96	[0.39; 1.98]
OTST	37	24301		1.52	[1.07; 2.10]
MBTST	21	14848		1.41	[0.88; 2.16]
OVVV	65	37185		1.75	[1.35; 2.23]
Trento	50	46343		1.08	[0.80; 1.42]
Random effects model		**129969**		**1.37**	**[1.13; 1.68]**

Heterogeneity: $I^2 = 39\%$, $\tau^2 = 0.0190$, $p = 0.10$

Fig. 2. (a) DBT invasive CDR. (b) DBT DCIS detection rate. DCIS, ductal carcinoma in-situ.

Study	Events	Total	Events per 1000 observations	Events	95%−CI
STORM	136	25058		5.43	[4.56; 6.42]
OTST	378	59877		6.31	[5.69; 6.98]
MBTST	259	43769		5.92	[5.22; 6.68]
OVVV	379	61742		6.14	[5.54; 6.79]
Trento	205	37436		5.48	[4.75; 6.28]
Random effects model		**227882**		**5.95**	**[5.65; 6.28]**

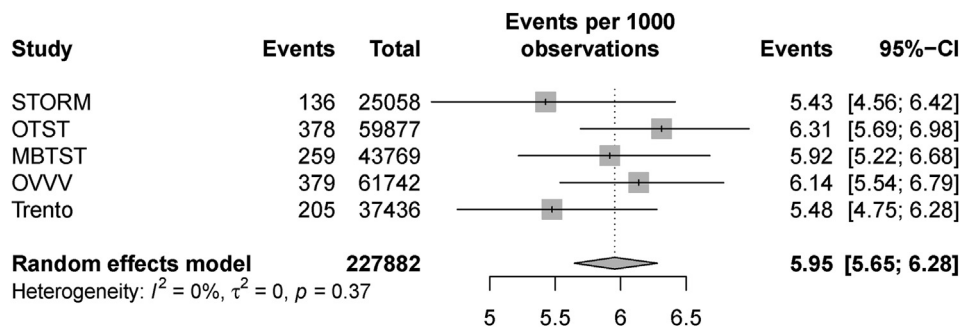Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.37$

Fig. 3. DM CDR.

detection of invasive BC (pooled difference in CDR: 2.72/1000; 95% CI 2.13−3.31) and DCIS (pooled difference in CDR: 0.49/1000; 95% CI 0.12−0.86).

3.2. Interval breast cancer rates

At final follow-up of the DBT-screened groups, a total of 201 interval cancers were reported amongst 127,425 screening participants (study-specific data shown in Fig. 6)—a pooled ICR of 1.56 per 1000 DBT screens (95% CI 1.22−2.00). In the comparison groups screened with DM, 402 interval cancers were reported amongst 223,903 screens (Fig. 6)—a pooled ICR of 1.75 per 1000 DM screens (95% CI 1.46−2.11). Comparing DBT-screened and DM-screened cohorts (Fig. 6) showed a pooled *difference* in ICR of −0.15 per 1000 screens (95% CI −0.59 to 0.29). A sensitivity analysis including only the three studies that had concurrent comparison groups

[9,13−16] gave a similar estimate of the *difference* in ICR (pooled difference −0.24/1000; 95% CI −1.12 to 0.63) between DBT and DM. Additional sensitivity analyses retaining three studies that used population cancer registry linkage to ascertain interval cancers [11,13−16] (pooled difference in ICR −0.08/1000 screens; 95% CI −0.85 to 0.68), or including only studies that used DBT with synthetic 2D-images [15,17] (pooled ICR difference 0.07/1000 screens; 95% CI −0.63 to 0.77) did not substantially change results (no significant difference in ICR in any sensitivity analyses).

4. Discussion

DBT has brought an opportunity to improve the effectiveness of population BC screening, initially realised through higher CDR compared with mammography. We argued that if DBT increased CDR, evidence that
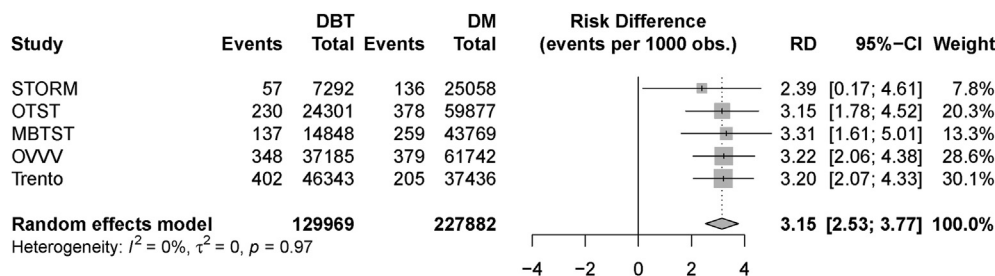
Fig. 4. Difference in CDR between DBT-screened and DM-screened cohorts (unpaired data). RD, risk difference.
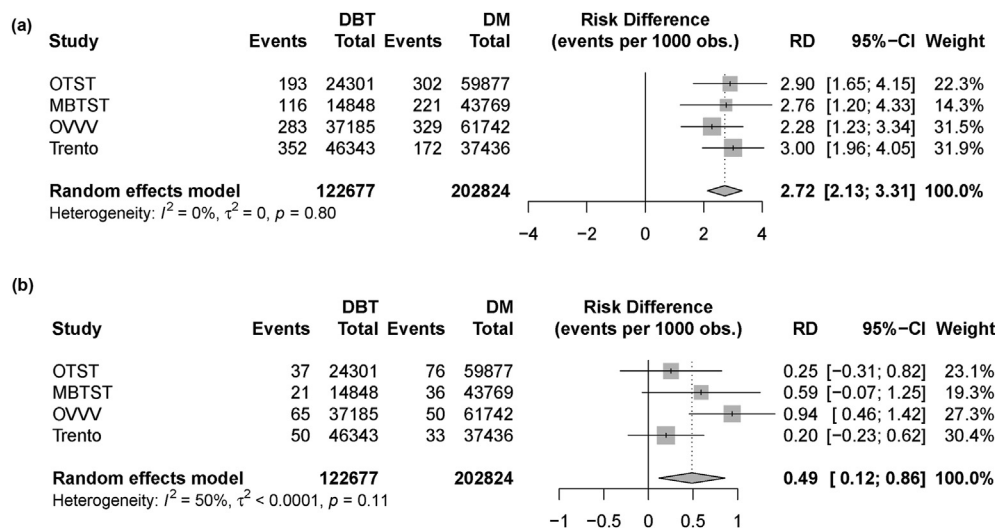

Fig. 5. (a) Difference in CDR (invasive breast cancer) between DBT and DM screening (unpaired data). (b) Difference in CDR (DCIS) between DBT and DM screening (unpaired data).
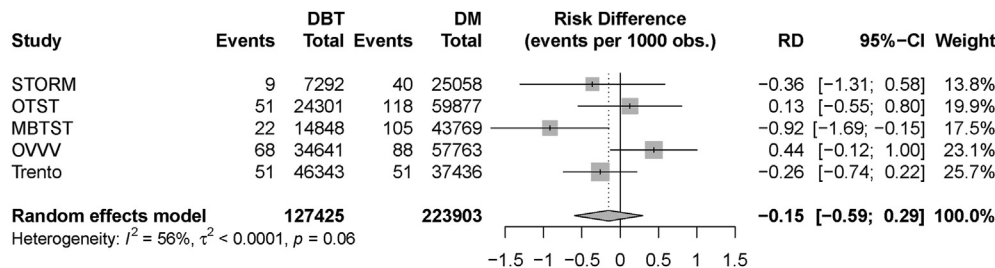

Fig. 6. Difference in *interval cancer rate* between DBT-screened and DM-screened cohorts.

this subsequently reduces ICR would give an early signal that increased CDR from DBT was enhancing screening effectiveness [8]. We therefore undertook a meta-analysis to assess the impact of DBT on *both* CDR and ICR, focusing on prospective DBT population screening studies. We found consistent evidence that DBT significantly increased CDR, which was evident across all studies in this meta-analysis, and in all pooled analyses including the main comparison of groups screened with either DBT or mammography (pooled CDR *difference* 3.15/1000 screens). It was also evident for the detection of invasive BC and DCIS. In contrast,

we found a pooled *difference* in ICR of −0.15/1000 comparing DBT-screened and mammography-screened cohorts from the same screening programs. The lack of effect on ICR was unchanged in several sensitivity analyses of the pooled difference in ICR. Although these are discouraging findings, we outline the relevance of our results, which address an evidence gap on the effect of DBT screening [1] and have implications for BC screening programs world-wide, while emphasising the limitations of study-level meta-analysis.

Our meta-analysis did *not* aim to examine performance measures (CDR, recall rates) for DBT screening,

this has been comprehensively done by Marinovich et al [6] in a meta-analysis of all studies of DBT screening (most of which did not report interval cancer data). Instead, we focused on DBT's effect on ICR, assuming this to be a surrogate for screening effectiveness, and meta-analysed CDR and subsequent interval cancer data from studies reporting *both* these outcomes. Our work examined a distinct issue from meta-analysis of performance measures and other DBT reviews [30–32]. We are also cognisant that we used pre-defined eligibility criteria when evaluating ICR as proposed in an IPD meta-analysis protocol [8]. By adapting those criteria, we included prospective DBT screening studies (to reduce selection bias) in the context of biennial screening practice. These criteria do not detract from our findings, they enhance quality through the inclusion of prospective DBT studies; we also outlined potential study limitations given that randomisation to DBT or DM was not used (online Appendix-2).

Both the International Agency for Research on Cancer [1] and the European Commission Initiative on Breast Cancer [3] have reviewed the evidence on DBT, and concluded that there was limited evidence on DBT's longer term screening outcomes including ICR [1,3], so our findings have implications for BC screening programs navigating decision-making about transitioning to DBT. Such decisions would need to consider that despite increased CDR from DBT (versus DM) screening, there was little effect on ICR at follow-up, implying that some of DBT's additional BC detection could represent over-detection, including some BCs potentially detected at DM in subsequent screen-rounds had DBT not been used (therefore DBT may have extended lead time). However, over-detection cannot be quantified in this meta-analysis, and we lacked detailed tumour characteristics; so, our findings are not conclusive but suggest a potential for over-diagnosis based on discordance between increased CDR in DBT-screened populations and the lack of an effect on ICR compared to DM. Furthermore, it is quite possible that ICR might not be an appropriate surrogate for screening effectiveness, or it may require cumulative rates to be quantified over multiple screening rounds if there is a modest effect from DBT, or if screen-reading experience with DBT leads to a differential effect on ICR at repeat screening. We emphasize that the prospective studies in our meta-analysis reflect initial (prevalent) DBT screening, which could account for the high CDR, and that there was heterogeneity in the estimated study-specific and pooled *difference* in ICR. Also, the Malmo trial [13,14] showed a small but significant reduction in ICR in the DBT-screened population.

Although we have focused on cancer detection and ICR in this meta-analysis to fill existing evidence gaps regarding DBT [1,3], we concede that alternate outcome measures, such as cumulative effect on advanced BC rates as proposed in ongoing randomised trials of DBT screening [20,33], might better indicate the comparative effect of BC screening technologies on health outcomes.

There are limitations inherent in study-level meta-analysis, including limited scope to adjust for possible differences in participant characteristics (such as age and density) between groups being compared. However, study characteristics show generally similar screening settings and populations from which DBT-screened participants and comparison groups were drawn (Table 1), although a younger aggregate age is noted for the Malmo (MBTST) comparison group. The latter is unlikely to have affected our pooled comparisons, because study-specific estimates for CDR and ICR in our meta-analysis align with age-adjusted estimates reported for this study [14]. Importantly, our planned IPD meta-analysis will allow adjustment of pooled analyses for age and density, and will enable characterisation of interval cancers. Another limitation is that two included studies used historical (rather than concurrent) comparison groups from the same population [11,17]; we therefore used sensitivity analysis (excluding those studies) and found that there was no change in pooled estimates of the *difference* in CDR or ICR.

5. Conclusions

This study-level meta-analysis reports robust evidence that DBT significantly increased CDR compared with mammography population screening; however, this did not have an effect on ICR. Analyses using IPD, and additional comparative studies preferably randomised trials, including repeat DBT screen-rounds, and evaluations of longer-term outcomes including but not limited to ICR, are needed to inform future breast screening practice.

Role of funder

Funding bodies did not have any role in the design, analysis, or reporting of this work.

Conflict of interest statement

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: N. Houssami receives research support through a National Breast Cancer Foundation (NBCF Australia) *Breast Cancer Research Leadership Fellowship*; the other authors do not have conflict of interest to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2021.01.035.

References

[1] Lauby-Secretan B, Scoccianti C, Loomis D, Benbrahim-Tallaa L, Bouvard V, Bianchini F, et al. Breast-cancer screening–viewpoint of the IARC working group. N Engl J Med 2015;372:2353–8.

[2] Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Canc 2013;108:2205–40.

[3] European Commission Initiative on Breast Cancer (ECIBC). Use of digital breast tomosynthesis (DBT). 2020. Online source, https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines/screening-tests/DBT-or-DM. [Accessed December 2020].

[4] Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol 2013:583–9.

[5] Houssami N, Miglioretti DL. Digital breast tomosynthesis: a brave new world of mammography screening. JAMA Oncol 2016; 2:725–7.

[6] 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; 110:942–9.

[7] 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 Malmo Breast Tomosynthesis Screening Trial, a population-based study. Eur Radiol 2016;26:184–90.

[8] Houssami N, Lång K, Hofvind S, Zackrisson S, Bernardi D, Hunter K, et al. Effectiveness of digital breast tomosynthesis (3D-mammography) in population breast cancer screening: a protocol for a collaborative individual participant data (IPD) meta-analysis. Transl Cancer Res August 2017;6(No 4). Translational Cancer Research 2017.

[9] Houssami N, Bernardi D, Caumo F, Brunelli S, Fantò C, Valentini M, et al. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. Breast 2018;38:150–3.

[10] Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, 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;23:2061–71.

[11] Skaane P, Sebuødegård S, Bandos AI, Gur D, Østerås BH, Gullien R, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial. Breast Canc Res Treat 2018;169:489–96.

[12] Skaane P, Bandos AI, Niklason LT, Sebuødegård S, Østerås BH, Gullien R, et al. Digital mammography versus digital mammography plus tomosynthesis in breast cancer screening: the Oslo tomosynthesis screening trial. Radiology 2019;291:23–30.

[13] Zackrisson S, Lång K, Rosso A, Johnson K, Dustler M, Förnvik D, et al. One-view breast tomosynthesis versus two-view mammography in the Malmo Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. Lancet Oncol 2018;19:1493–503.

[14] Johnson K, Lång K, Ikeda D, Åkesson A, Andersson I, Zackrisson S. Interval breast cancer rates and tumor character-

istics in the prospective population-based Malmö Breast Tomosynthesis Screening Trial (MBTST). Radiology 2021 [in press].

[15] Hofvind S, Hovda T, Holen ÅS, Lee CI, Albertsen J, Bjørndal H, et al. Digital breast tomosynthesis and synthetic 2D mammography versus digital mammography: evaluation in a population-based screening program. Radiology 2018;287:787–94.

[16] Hovda T, Holen ÅS, Lång K, Albertsen JL, Bjørndal H, Brandal SH, et al. Interval and consecutive round breast cancer after digital breast tomosynthesis and synthetic 2D mammography versus standard 2D digital mammography in BreastScreen Norway. Radiology 2020;294:256–64.

[17] Bernardi D, Gentilini MA, De NM, Pellegrini M, Fantò C, Valentini 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;50:135–40.

[18] Houssami N, Lockie D, Clemson M, Pridmore V, Taylor D, Marr G, et al. Pilot trial of digital breast tomosynthesis (3D mammography) for population-based screening in BreastScreen Victoria. Med J Aust 2019;211:357–62.

[19] Aase HS, Holen ÅS, Pedersen K, Houssami N, Haldorsen IS, Sebuødegård S, et al. A randomized controlled trial of digital breast tomosynthesis versus digital mammography in population-based screening in Bergen: interim analysis of performance indicators from the To-Be trial. Eur Radiol 2019;29:1175–86.

[20] Pattacini P, Nitrosi A, Giorgi RP, Iotti V, Ginocchi V, Ravaioli S, et al. Digital mammography versus digital mammography plus tomosynthesis for breast cancer screening: the reggio emilia tomosynthesis randomized trial. Radiology 2018;288:375–85.

[21] Caumo F, Zorzi M, Brunelli S, Romanucci G, Rella R, Cugola L, 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;287:37–46.

[22] Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, 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;17:1105–13.

[23] 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;287:49–57.

[24] 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;2:737–43.

[25] Winkler NS, Freer P, Anzai Y, Hu N, Stein M. Impact of immediate interpretation of screening tomosynthesis mammography on performance metrics. Acad Radiol 2019;26:210–4.

[26] Sharpe Jr RE, Venkataraman S, Phillips J, Dialani V, Fein-Zachary VJ, Prakash S, et al. Increased cancer detection rate and variations in the recall rate resulting from implementation of 3D digital breast tomosynthesis into a population-based screening program. Radiology 2016;278:698–706.

[27] Maxwell AJ, Michell M, Lim YY, Astley SM, Wilson M, Hurley E, et al. A randomised trial of screening with digital breast tomosynthesis plus conventional digital 2D mammography versus 2D mammography alone in younger higher risk women. Eur J Radiol 2017;94:133–9.

[28] Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Base Ment Health 2019;22:153–60.

[29] R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. 2020.

[30] Gilbert FJ, Tucker L, Young KC. Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool. Clin Radiol 2016;71:141–50.

[31] Yun SJ, Ryu CW, Rhee SJ, Ryu JK, Oh JY. Benefit of adding digital breast tomosynthesis to digital mammography for breast cancer screening focused on cancer characteristics: a meta-analysis. Breast Canc Res Treat 2017;164:557–69.

[32] Houssami N, Skaane P. Overview of the evidence on digital breast tomosynthesis in breast cancer detection. Breast 2013;22:101–8.

[33] The tomosynthesis mammographic imaging screening trial (TMIST). Online source, www.cancer.gov/about-cancer/treatment/clinical-trials/nci-supported/tmist. [Accessed December 2020].