A randomised controlled trial of Digital Breast Tomosynthesis versus Digital Mammography as primary screening tests: screening results over subsequent episodes of the Proteus

Armaroli Paola (Orcid ID: 0000-0001-8965-4307)

A randomised controlled trial of Digital Breast Tomosynthesis versus Digital Mammography as primary screening tests: screening results over subsequent episodes of the Proteus Donna study

Brief title

Randomised trial of Digital Breast Tomosynthesis versus Digital Mammography as primary screening tests

Paola Armaroli[1], Alfonso Frigerio[2], Loredana Correale[1], Antonio Ponti[1], Franca Artuso[2], Denise Casella[1], Paolo Falco[3], Elisabetta Favettini[4], Paolo Fonio[5], Livia Giordano[1], Vincenzo Marra[6], Luisella Milanesio[2], Lia Morra[7], Pietro Presti[8], Emilia Riggi[1], Viviana Vergini[1], & Nereo Segnan[1]

Affiliation

[1] S.S.D. Epidemiologia Screening – CPO. AOU Città della Salute e della Scienza di Torino, Via Cavour 31, Turin, 10123, Italy. (Armaroli MD, Correale PhD, Casella MSc, Giordano MD, Ponti MD, Riggi PhD, Vergini BS, Segnan MD)

[2] S.S.D. Senologia di Screening. AOU Città della Salute e della Scienza di Torino, Via Cavour 31, Turin, 10123, Italy. (Frigerio MD, Milanesio MD, Artuso Radiographer)

[3] AizoOn Technology Consulting, Str. del Lionetto 6, Turin, 10146, Italy. (Falco MSc)

[4] ASL BI. Via dei Ponderanesi 2, Ponderano 13875, Italy (Favettini MD)

[5] Dipartimento di Diagnostica per Immagini e Radiologia Interventistica - AOU Città della Salute e della Scienza di Torino. Via Genova 3, Turin, 10126, Italy. (Prof. Fonio)

[6] S.C. Radiologia Sant'Anna, AOU Città della Salute e della Scienza di Torino, Corso Spezia 60, Turin, 10126, Italy. (Marra MD)

[7] Dipartimento di Automatica e Informatica, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, 10129, Italy. (Morra PhD)

[8] Fondazione Edo ed Elvo Tempia. Via Malta 3, Biella, 13900, Italy. (Presti PhD)

Corresponding author

Paola Armaroli, MD

S.S.D. Epidemiologia Screening – CPO. AOU Città della Salute e della Scienza di Torino, Via Cavour 31, 10123, Turin, Italy

paola.armaroli@cpo.it

+390116333876

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ijc.34161

Abstract

Proteus Donna is a randomised controlled trial aimed at prospectively evaluating screening with Digital Breast Tomosynthesis (DBT), including interval cancer detection (ICD) and cancer detection (CD) in the analysis as a cumulative measure over subsequent screening episodes.

Consenting women aged 46-68 attending the regional Breast Screening Service were randomly assigned to conventional digital mammography (DM, control arm), or DBT in addition to DM (DBT, study arm). At the subsequent round all participants underwent DM. 36-months follow-up allowed for the identification of cancers detected in the subsequent screening and inter-screening interval. Relative risk (RR) and 95% confidence interval (95%CI) were computed. Cumulative CD and Nelson-Aalen incidence were analysed over the follow-up period.

Between 31/12/2014 and 31/12/2017, 43022 women were randomised to DM and 30844 to DBT. At baseline, CD was significantly higher (RR 1·44; 95%CI 1·21-1·71) in the study arm. ICD did not differ significantly between the two arms. (RR 0·92; 95%CI 0·62-1·35). At subsequent screening with DM, the CD was lower (nearly significant) in the study arm (RR 0·83; 95%CI 0·65-1·06). Over the follow-up period, the cumulative CD (comprehensive of ICD) was slightly higher in the study arm (RR 1·15; 95%CI 1·01–1·31). The Nelson-Aalen cumulative incidence over time remained significantly higher in the study arm for approximately 24 months. Benign lesions detection was higher in the study arm at baseline and lower at subsequent tests.

Outcomes are consistent with a lead time gain of DBT compared to DM, with an increase in false positives and moderate overdiagnosis.

Introduction

Breast cancer (BC) screening using mammography reduces breast cancer mortality, as shown in randomised controlled trials (RCTs)[1,2]. However, beyond concerns about potential overdiagnosis[1], the application of mammography in screening has limitations related to its lower sensitivity in women with dense breast parenchyma[3]. Breast imagers are adapting to this challenge with the development of new technologies. Digital breast tomosynthesis (DBT) is an imaging technology that provides a three-dimensional reconstruction of the breast from a limited angle scan involving a series of low-dose mammographic exposures, hence the commonly used synonym of "3D mammography"[4]. Several studies have demonstrated a higher breast cancer detection (CD) for DBT either in adjunct to digital mammography (DM), or as a single test, integrated alongside "synthetic 2D" mammography images [5–13]. However, data relating to how this affects interval cancers detection (ICD) and CD at subsequent screening, and whether it produces overdiagnosis are very limited[5,14–17]. There are also challenges linked with using DBT in organised screening programmes, including increased radiation doses, longer reading times for radiologists, increased cost of the technology, and increased data storage requirements[18]. We designed the Proteus Donna RCT to prospectively compare DBT with DM in a population-based screening setting. While its primary outcome was to assess the impact of DBT on ICD, the trial design also allowed us to provide a relevant analysis of CD as a cumulative measure over subsequent screening episodes. In this paper, we reported the main trial results from two consecutive screening tests, and we attempted to elucidate whether the CD increase in the study arm was due to a gain in lead time, overdiagnosis, or both. Secondary outcomes of the study, such as diagnostic performance indicators and characteristics of BC detected, were also reported and briefly discussed. A further analysis of these and other outcomes, such as reading issues and cost-effectiveness will be the object of upcoming papers.

Materials and Methods

Study design and participants

Proteus Donna was a multicenter, multivendor prospective RCT on DBT implementation in BC screening. Recruitment began in December 2014 and was completed in December 2017, within the setting of the organised BC screening programme of the Piedmont Region Health Screening Service, Italy. In this Region, women aged 50-69 years are routinely invited to perform DM every 2 years, while women aged 45-49 may participate spontaneously and receive annual screening tests untill they reach the age of 50. All tests are double read, and further assessment is offered free of charge to all screen-positive women.

A total of six provincial screening centres, including 13 clinical hospitals, participated in the trial and 38 radiologists at these centres interpreted screening examinations. All radiologists fulfil the regional screening quality assurance criterion of at least 5000 mammograms read per year and participate in periodic audits of individual reading performance[19].

Eligible women were randomly assigned to either standard DM (control arm) or DBT, which was carried out in the so-called 'combo-mode' (i.e., combined tomosynthesis plus DM, study arm) (see Figure 1). Participants were followed for 36 months (18 months for women younger than 50 years of age) to identify cancers detected at the subsequent screening with DM alone, as well as during the inter-screening interval. The study design allowed us to compute: 1) CD at baseline and at subsequent screening; 2) ICD, i.e., BC occurred within the follow-up period after a negative study examination,; 3) overall cumulative detection (CD + ICD) after two screenings in the study and control arm.

The target population was women aged 46-68 years invited to the regional BC screening programme. Exclusion criteria were a personal history or symptoms of BC, oncological treatment for other malignancies, terminal or critical illness, inability to express informed consent, and breast implants. Eligible women received, with the standard letter of invitation to screening, a special leaflet outlining the potential benefits and risks of the study and specifying that their participation in the trial entailed the consent to receive additional DBT according to random allocation. The leaflet included a reference to the study website and a telephone contact number to make further information readily accessible. The information about screening arm assignments would be disclosed only after the woman's consent to participate had been signed. This occurred after registration at the screening facility and only a few minutes before the actual test was performed. Moreover, at the study clinical sites, the trial protocol was fully explained by trained volunteers. Women who agreed to participate signed the informed consent form. Before the exam, the radiographer reviewed the participants eligibility criteria. Those who did not meet all of the eligibility criteria were excluded from the study and offered DM screening according to the standard regional screening protocol.

The ethics committees of all participating hospitals approved the study protocol (Appendix 2)

Procedures

For DM and DBT examinations, mammography systems from Hologic, GE, Siemens, and Fujifilm were used at the study clinical sites[20]. All systems allowed for taking DM and DBT images with a single breast compression. All women underwent standard two-view (craniocaudal and mediolateral oblique views) DBT, or DM performed by trained radiographers. Radiation doses of a DBT acquisition were automatically determined by the mammography unit and set to a moderately higher dose than a standard mammographic view. Therefore, the complete dose entailed by the combination of a two-view DM and a two-view DBT, which was approximately 2.5 times higher than that of a standard DM alone. This increased dose was still within the limits of the radiation doses allowed by the Italian guidelines at the time of the study for a screening examination and was deemed acceptable by the ethical committees for a single round of screening within the experimental arm.

The radiologists participating in the trial received basic training in integrated DM and DBT mammography (details in Appendix 1). In both arms, examinations were interpreted

independently by two readers (independent double reading) as per regional screening protocol. In the DM arm, both readings for each examination were reported as usual. In the DBT arm, each reading was randomly allocated (on a per-case basis) to one of these three modes: DM plus DBT (DBT was only read after initial DM interpretation), DBT plus DM (DM was only read after DBT interpretation), and "concurrent mode" (DBT and DM were simultaneously available to the reader) (details in Appendix 1). In both arms of the study, breast imaging was reported based on an established European classification (1 – normal, 2 – benign, 3 – probably benign, 4 – suspicious for malignancy, and 5 – malignant). If either of the radiologists assigned a score of three or higher to one or both breasts, the case was considered as a positive test result and the woman was recalled for further investigations (without consensus or arbitration). Previous DM screening mammograms were available to readers.

For any suspected abnormalities, data related to the side, quadrant, and morphology were collected . Interpretation time (defined as the time taken to read the data once loaded onto the workstation) was automatically recorded. Reading issues, i.e., reading modes and times, will be the subject of further publications.

Participants recalled for further assessment underwent diagnostic work-up according to local routines, typically involving triple assessment (clinical assessment, further imaging [supplementary mammographic views, microfocus magnification, further 3D projections, and ultrasound], and needle biopsy when needed). All cancer diagnoses were histologically proven either after surgery or by a core biopsy.

Outcomes

Primary outcomes (recalls for further assessment, screen-detected cancers, and interval cancers) were ascertained by the population-based screening database, excision histology reports in patients who received surgery, and core biopsy reports and work-up imaging. Clinical information was extracted from a clinical database[21,22] and the population cancer registry, with both covering the entire study population. If any doubts persisted, the original clinical documents were retrieved. All screening participants were followed up to ascertain interval cancers, defined as cancers occurring after a baseline negative examination and before a subsequent screening test within the follow-up period. The follow-up was conducted by employing the clinical databases at the hospitals and the regional hospital discharge records (2014–2020). For hospital discharge records, the record linkage was based on two independent identifiers (social security number and an algorithm derived from name and birth date). International Classification of Diseases-9-CM diagnosis codes related to breast malignancy or codes for surgical, diagnostic, or medical procedures possibly related to breast cancer were selected. Histopathologic tumour characteristics were based on biopsy reports on the surgical specimen or, in the case of primary chemotherapy, on the core biopsy. The histologic type was defined as ductal carcinoma in situ (DCIS), invasive cancer (non-special type, lobular carcinoma, tubular carcinoma, or other invasive carcinoma). Prognostic tumour characteristics for invasive

tumours included pathologic tumour size (pT <=1c, >1c), lymph node status (pN negative or positive), tumour stage[23] (<=I, II+), and histologic grade (1–3). Lesions were classified according to maximum size, stage, grade.

Statistical analyses

For the present study, follow-up was available up to 31 December 2019. The sample size of the study was calculated by assuming an incidence of interval cancers of 15·2/10000[24] in the control arm. The study envisaged a variable randomisation ratio, to maximise the yield of DBT examinations according to changing local availability of DBT machines and time slots for DBT in the different sites of the trial. Assuming an average 1:1·3 DBT (study) to DM (control) allocation, a sample size of 70000 women (30000 in the experimental arm) would allow a power of 80% to observe a decrease in the ICD of at least 50%, accepting a 5% probability of alpha error with a chi-squared test. The sample size provides a power of 80% to observe an increase in breast cancer detection of at least 34% (6·7/1000 in the experimental arm versus 5·0/1000 in the control arm) with a 5% probability of alpha error.

For each study arm, we calculated the number of screen-detected cancers and CD per 1000 women performing the test, the number and percentage of screened women recalled to assessment (recall, %), the positive predictive value (PPV) for recall, and the benign to malignant (B/M) surgical biopsy ratio (i.e., the proportion of operated cases who had a benign histological diagnosis to those with a malignant outcome). The main characteristics of detected cancers (stage, size, histology, grade, and node status) were tabulated for DBT versus DM. All indicators were computed separately for initial and subsequent screening examinations. Data on interval cancers detection (ICD) were based on cancers identified over the follow-up period from the screening date (median, 25 months, range, 0-36) and were estimated accounting for all negative screens. The relative risk (RR), risk difference (RD), and 95% confidence interval (95% CI) were computed to compare the distribution of outcomes between those screened with DBT versus DM.

Cumulative incidence of breast events (malignant lesions, i.e., invasive and in situ BC, and benign lesions) was analysed for complete follow-up. A chi-square test of homogeneity was used to assess effect modification by age and screening history at baseline. The Nelson-Aalen estimator[24] was used to compare the cumulative incidence of breast events over time between women undergoing DBT versus DM screening. The outcome variables were time to first malignant lesions, or benign lesions.

We attempted to estimate the further gain, if any, in lead time of DBT over DM, assuming that the excess number of cases observed in the experimental arm was just due to advancement in the time of diagnosis. At index screening, we calculated this excess number of cases as observed to expected absolute difference (O-E).

Given the study design, where the experimental arm reverts to the control arm procedure at the subsequent screening, we assumed that the further lead time gained by DBT is exhausted when the cumulative incidence in the DBT and DM arms are equal.

For under and over age 50 we estimated the actual average interval (I) (in days) before the subsequent screening round, and we multiplied the excess number of cases (O-E) in each age group (i) by the interval, as the total amount of lead time that DBT excess cases could have experienced. We divided the total lead time by the number of cases detected by DBT at baseline (N), to obtain the average lead time for cases detected at the experimental arm (LTa) over the control arm.

The average gain in lead time "LTa" with DBT was calculated as LTa =[(O-E)*I]i/N, where "O-E" is the observed to expected absolute difference of screen-detected cancers in the DBT arm compared to DM arm, "I" is the time interval before the subsequent screening, and "N" is the total number of DBT cases at baseline[25]. Under the same assumption, we estimated the difference in the number of cancer cases per 1000 women who may have been detected one or two years earlier between the DBT arm and the DM arm at baseline screening.

Categorical variables were reported as a frequency, while continuous variables as a mean and standard deviation (SD).

Statistical analyses were performed using STATA version 15[26]. All significance tests were two-sided and considered statistically significant at $p<0.05$.

The trial was registered with ClinicalTrials.gov (number NCT02590315)[27].

Results

The flow diagram in Figure 1 summarises the main results of the trial. Of the 212794 women invited, 129778 attended the screening programme during the study period. Trial information was given to 108842 (84·0%) attendees, of whom 77057 (70·8%) accepted to participate. After application of the exclusion criteria (n=3190 exclusions), a total of 73866 participants were included in the final analysis: 30844 randomised to the DBT (experimental) arm and 43022 randomised to the DBT (control) arm, with an average study to control randomisation ratio of 1:1·39.

Baseline screening

Age was balanced in the two arms: mean 56·99 (SD 6) vs. 56·83 (SD 6) years in the experimental and control arms, respectively. As for screening history, first tests were 3316 (10·75%) in the DBT arm and 3842 (8·93%) in the DM arms ($p<0·05$).

Table 1 reports the results of the baseline tests. In the DBT arm significantly more women were recalled for further assessment (RR: 1·24; 95%CI 1·17-1·32), underwent needle biopsy (RR: 1·60; 95%CI 1·35-1·90), and were referred for surgery (RR: 1·54; 95%CI 1·31-1·83).

Also, a higher proportion of women in the DBT arm (443, 1·44%) were recalled for a further test within six months or one year, compared with the DM arm (RR: 1·36; 95%CI 1·19-1·56).

CD was significantly higher in the DBT arm (8·30 per 1000 women vs 5·76 per 1000 women; RR: 1·44; 95%CI 1·21-1·71, RD: 2·54; 95%CI 1·30-3·78). When considering only invasive cancers, 224 lesions were detected out of 30844 women in the DBT arm (7·26 per 1000 women) vs 214 lesions out of 43022 women (4·97 per 1000 women) in the DM arm (RR: 1·46; 95%CI 1·21-1·77). DCIS detection was 39% higher in the experimental arm (1·04 per 1000 women vs 0·74 per 1000 women), although not significantly (RR: 1·39; 95%CI 0·83-2·35). Significantly more benign lesions were detected in the experimental arm (52 out of 30844, 1·69 per 1000 women) as compared to the control arm (29 out of 43022, 0·67 per 1000 women; RR: 2·50; 95%CI 1·59-3·94, RD: 1·01 95%CI 0·49-1·53). The ratio between a benign and a malignant outcome for lesions referred for surgery was increased to 0·20 in the DBT arm vs 0·12 in the control arm ($p<0·05$). A higher PPV of surgical referral was observed with DBT, however, false positive recalls were significantly higher by 22% in the DBT arm. In the DBT arm, the increase in CD was significant for smaller (pT<=1c; RR: 1·54; 95%CI 1·24-1·92), node negative invasive lesions (RR: 1·54; 95%CI 1·23-1·97). The CD gain was especially relevant for stage <=I (RR: 1·67; 95%CI 1·30-2·15), and grade 1 and 2 invasive cancers (RR: 1·73; 95%CI 1·25-2·40; 1·38; 95%CI 1·01-1·89; respectively). The observed increase in CD was not significant for larger lesions, above 2 cm in size (RR: 1·12; 95%CI 0·64-1·97) (Table 1).

Subsequent screening

After the baseline testing, 30039 women in the DBT arm and 42216 in the control arm, were due to be recalled for their subsequent screening test after one year if younger than 50 years of age and after two years if above this cut-off.

Within the follow-up period, 1414 women in the DBT and 1966 in the DM arm did not receive an invitation, 4860 and 6705, respectively, received an invitation but did not participate, while five and 11 women could not perform the screening test for technical reasons. As a result, 23760 women in the DBT arm and 33534 in the DM arm performed subsequent screening test. Results are reported in Table 2.

Age was balanced in the two arms: mean age was 59·06 (SD 6) vs 59·18 (SD 6) years in the DBT and DM arm, respectively. Cases were also balanced when considering the interval between the baseline and the subsequent test: the mean was 718·47 (SD 144) and 715·35 (SD 138) days in the study and control arms.

For consecutively screened women, a reduction in all the considered outcomes was observed with respect to baseline tests. Recall to further assessment was 4·21 per 100 women (1000 of 23760) in the DBT arm and 4·34 per 100 women (1456 of 33534) in the DM arm (RR 0·97; 95%CI 0·89-1·05). Surgery referral was markedly reduced compared to baseline and significantly lower in the DBT arm compared to the DM arm (RR: 0·76; 0·59-0·97). A similar pattern was observed for CD, with a marked reduction vs baseline values for DBT and a non-significant, lower CD in the DBT arm (4·21 per 1000 women) as compared with the DM arm (5·07 per 1000 women) (RR 0·83; 95%CI 0·65-1·06, RD -0·86; 95%CI -1·98-0·26). When considering only invasive cancers, 81

lesions were detected out of 23760 women in the DBT arm (3·41 per 1000 women) vs 135 lesions out of 33534 women (4·03 per 1000 women) in the DM arm (RR: 0·85; 95%CI 0·63-1·12). DCIS detection was 25% lower in the experimental arm, although not significantly different (RR: 0·75; 95%CI 0·39-1·39). Surgery with a benign outcome was markedly reduced compared to baseline and, non-significantly, lower in the DBT arm compared to the DM arm (0·25 per 1000 women in the DBT arm, vs 0·57 in the DM arm, RR 0·45; 95% CI 0·18-1·12, RD -0·31; 95% CI -0·64-0·01). Among invasive BCs detected in the DBT arm, a reduction of the more advanced cases was observed compared to the baseline test and in comparison with the DM arm, although non-significant, in pT>1c cases (RR: 0·74; 95%CI 0·31-1·68), positive nodes BC (RR: 0·71; 95%CI 0·34-1·39), and stage II+ (0·72 per 1000 women in the DBT arm vs 1·10 in the control arm; RR: 0·65; 95%CI 0·34-1·18). No difference was observed between the study and control arms with respect to histologic grade.

Interval cancers

The mean follow-up period was 25 months (range 0-36 months). No statistically significant difference in ICD was observed between the two arms (RR: 0·92; 95%CI 0·59-1·40) (Table 3). In the DBT arm 42 interval cancers were observed among 30588 women (1·37 per 1000 women). In the control arm the ICD was 1·50 per 1000 women (64 of 42776). As reported in Table 3, no difference in stage was observed between the two arms. The ICD was similar for smaller lesions, while a reduction of the more advanced cases was observed (pT>1c RR: 0·52; 95%CI 0·09-2·18), although non-significant. When considering grade, ICD was similar comparing DBT to DM, except for grade 3, even if not significantly (RR: 0·54; 95%CI 0·15-1·61).

Cumulative cancers

In the complete period of follow-up (Table 4), there were 880/73866 BCs (1·19 per 1000 women) including invasive and in situ lesions. The cumulative detection of BCs (including events at the baseline screening) was significantly higher in the DBT arm than in the DM arm, with 398/30844 (12·90; 95% CI:11·67-14·23 per 1000) cancers in the DBT arm versus 482/43022 (11·20; 95% CI:10·23-12·24) cancers in the DM arm (RR 1·15; 95% CI 1·01–1·31; RD 1·70; 95% CI 0·10-3·30). The cumulative detection of a benign diagnosis was also significantly higher in the DBT than in DM arm (RR: 1·69; 95% CI: 1·15-2·47, RD 0·76; 95% CI 0·19-1·34).

For malignant lesions, age (i.e., 46-49 versus 50-69) and screening history at baseline were not effect modifiers (p=0·47 for age; p= 0·86 for screening history); the same applies to benign lesions (p=0·11 for age; p= 0·05 for screening history).

When considering the Nelson-Aalen cumulative incidence over time (Figures 2), a significant excess of cases was observed in the DBT arm with respect to the DM arm up to 24 months. After that, the cumulative incidence curves for malignant lesions crossed (Figure 2A), showing a non-significant lower cumulative incidence in the DBT arm after two subsequent tests within 36 months of follow-up. For benign lesions (Figure 2B), the Nelson-Aalen cumulative incidence remained higher throughout the follow-up period, although not significantly different.

Lead time gain

In the DBT arm at baseline screening, we observed an excess of 78 malignancies over those expected based on the observed DM detection. Weighing the results by the interval between screening tests (742 median days, accounting for both women under the age of 50 and those above this threshold), these cases account for 158 years of further advancement in the time of diagnosis, corresponding, on average, to 226 days for each of the 256 malignancies in the DBT arm. Given the difference in detection between the two arms at baseline screening (Table 4), 2.54 per 1000 women experienced due to DBT further advancement in the time of diagnosis of one or two years.

Discussion

Significant increases in CD have been consistently reported in recent years for screening with DBT over standard DM screening. However, data relating to the effect of DBT screening on ICD and overdiagnosis related to this enhanced detection are still limited. The *Proteus Donna* randomised, controlled trial was designed to provide relevant information on such effects in a population-based screening setting. In this paper, we present the results of the primary outcomes of our study, i.e., CD and ICD at baseline, first subsequent screening test, and in the first interval. Secondary outcome measures, such as diagnostic performance indicators for organised screening, cost-outcome, and cost-effectiveness analyses will be the subject of further publications.

At baseline, a significantly higher number of women was recalled for further assessment, and underwent biopsy and surgery, with a 44%, significant increase in CD in the DBT arm. Also, a significantly higher detection of benign lesions was observed. Among invasive cancers, a gain in detection was observed for smaller lesions. No statistically significant difference in ICD was observed between the two arms. ICD was still similar in the two arms when considering lesion size, stage and grade.

A reduction in all outcomes was observed at the first subsequent round of screening. A non-significant decrease in screen-detected BCs was observed in the DBT arm. Among invasive cancers, a non-significant greater reduction was observed pT>1c lesions, in those with positive lymph nodes, and in lesions of stage II+.

These results are in line with results from prospective studies, as for CD at baseline[5,9,28], interval cancer[5,28], and subsequent screening[16,17].

This design allows to evaluate whether a net increase in CD of DBT vs DM exists in the screened population, overcoming the potential bias of the learning curve linked to successive DBT rounds, and to estimate the lead time gain of DBT vs DM.

For asymptomatic women with an average risk of BC, the ECIBC's Guidelines Development Group[29] suggests using either DBT or DM in the context of an organised screening programme. This is described as a 'conditional recommendation, very low certainty of the evidence'. A

similar recommendation was issued by the International Agency for Research on Cancer[2] (IARC), because studies have consistently showed higher CD compared to DM, with no significant decrease in the ICD.

This raised the dilemma about the DBT's ability to detect lesions that are clinically malignant rather than indolent (overdiagnosis). However, all but one observational study[17], reported that the results of the first screening test alone, without a long enough follow-up period, required the measurements of cumulative CD over more than one screen. If the lead time of DBT was indeed longer than that of standard DM, then the CD would decrease when the DBT is stopped after one or more screens and DM is performed again. We observed that CD was 44% significantly higher at baseline and decreased by 17% (RR 0·83; 95% CI 0·65-1·06) at the subsequent screen, in the DBT arm compared to the DM arm. This finding suggests a trend in the reduction of CD at subsequent rounds after DBT, and thus, that some advancement in the time of diagnosis was gained by DBT over DM at the first screen. The cumulative CD, inclusive of interval cancers, in the DBT arm was significantly higher than the cumulative CD in the DM arm. However, according to the Nelson-Aalen cumulative incidence results (including interval cancers), the estimated CD in the DBT arm is equal to the estimated CD in the DM arm at around 30 months of follow-up. Detection of malignancies pT>1c, with positive nodes, or stage II+ was lower at the subsequent screening in the DBT arm (Table 2); however, none of these differences were significant. A similar trend was observed for benign lesions, even if the cumulative detection was still significantly higher in the DBT arm than in the DM arm.

These results are compatible either with greater lead time gain with DBT screening in comparison with DM screening and/or with an excess of detection of indolent cancers (overdiagnosis) in the DBT arm. In fact, although a higher cumulative CD in the DBT arm was observed, the Nelson-Aalen cumulative incidence analysis (including interval cancers) showed the excess CD tends to disappear over time. We estimated an average lead time gain of 226 days attributable to DBT over DM. This estimated lead time gain applies to the screening interval observed in our study, being a component of the total lead time (DM, DBT, and the interval cases). In the absence of regular screening or for longer intervals between tests, the advancement in the time of diagnosis should be measured when the cumulative incidence is equal in the two arms.

We can alternatively speculate that: a) the non-significant decrease of CD and cumulative incidence of BC at the second screen in the DBT arm, plus the favorable distribution of the prognostic factors of BC in the study arm (size, node, and stage) are chance effects; b) the DBT diagnostic advancement did not fully emerge before the subsequent screen; c) longer advancement and higher overdiagnosis of DBT versus DM coexist. From a clinical significance point of view, the extent to which the advancement in the time of diagnosis of asymptomatic cancers detected by DBT would translate to clinical benefit is uncertain. It is possible that it would just extend the amount of time a woman lives with a breast cancer diagnosis, without

improving her quality or quantity of life, or providing any other benefits. However, if confirmed, advancement in the time of diagnosis would permit the downstaging of screen-detected cancers, less harmful treatment, and/or longer screening intervals.

The strength of our study is that it is a randomised trial of DBT screening reporting cumulative incidence where the study arm reverts to the same procedure as the control arm during the follow-up period. The alternative approach of repeating the DBT at subsequent screening would not help to disentangle the occurrence of further diagnostic advancement versus overdiagnosis with DBT[30]. Overdiagnosis with either DBT or digital mammography is likely to be present[1,31].

Our results suggest a gain in lead time with DBT screening, but statistical significance is not acquired. We are now monitoring the results of the third screening episode in the trial, including interval cancers in the first and second intervals for future analyses with a longer follow-up period.

An important limitation of our study for accurately estimating comparative overdiagnosis is due to the need for a longer follow-up than is currently available. Another limitation is that our study eventually proved underpowered since we assumed a 50% reduction in ICD. Additionally, we could not collect breast density information and stratify for this important condition.

In conclusion, the *Proteus Donna* randomised controlled trial, a comparative effectiveness study, adopted a longitudinal design for comparing CD, ICD, and cumulative incidence within a population-based screening programme, by reverting all participants to standard DM screening after one round of DBT in the study arm. Our results are not conclusive, although most of our outcomes are consistent with a probable lead time gain with DBT as compared to DM, while some concurrent increase of overdiagnosis cannot be excluded. We are prolonging the follow-up of the trial population at the following DM screens in order to estimate whether and when the detection in the two trial arms will be equal.

Acknowledgments

The authors wish to acknowledge Dettori Davide and Delsanto Silvia for their passionate support to the project. Thanks to Ferraris Raffaella, Vecchia Renata, Zandona Emanuela, Gavelio Carla (Piedmont Region); Petruzza Francesco, Lizza Tarcisio (CSI); Associazione Volontari Ospedalieri S. Anna; Fondazione Edo ed Elvo Tempia; Associazione Volontari Ospedalieri Chieri; Associazione Volontari Ospedalieri Borgomanero; Associazione VITA (Vivere Il Tumore Attivamente); ANDOS; ANAPACA; MIMOSA and Forti e Serene. Collaborative group: Di Leo Aurora, Pitarella Sabina, Aiello Adriana, Selvestrel Valeria, Rosi Sabrina, Tota Donatella, Dore Daniela, Giordano Francesca, Suteu Liana, Anrò Patrizia, Amoruso Angela, Vaudagna Carlo Eugenio, Guasti Elisa, Albertinazzi Stefania, De Rocco Leonardo, Gulminelli Giorgio, Festari Sara, Chiarelli Giusi, Marini Federica, Ferrari Francesca, Mazzocchi Paolo, Tranchini Roberto, Rossi

Alessandro, Agostino Nadia, Germinetti Fiorella, Francese Benedetto, Magnani Corrado, Bestagini Piero, Polizzi Salvatore, Panarisi Pierino, Cavalot Gabriella.

Author Contributions

NS,AF,AP,LMO,LC,PP,LG did the design and supervised the study. AF,VM,EF,PFO,LMI,PFA,FA contributed to screen-reading and data collection. NS,PA,LC,DC,VV,ER contributed to statistical planning and data analysis. PA,NS,AP,LC,AF wrote the paper. PA and NS verified the data. All authors contributed to the revision of the manuscript and data interpretation. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. The work reported in the paper has been performed by the authors, unless clearly specified in the text.

Conflict of Interest

We declare no competing interests.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics statement

The ethics committees of all participating hospitals approved the study protocol. The study conformed to the Declaration of Helsinki, and all participants provided written informed consent. The study was registered at ClinicalTrials.gov (number NCT02590315).

Funding

The Piedmont Region (Department of Health), the University of Turin, the foundation Edo Tempia (a non-profit organization), and the regional Consortium for Informative Systems (CSI) sponsored the research. im3D S.p.A Torino (Italy) and Maita-Italian Ministry of Health (grant number RF-2016-02363686) funded the study by providing tomosynthesis equipment and financial support for patient recruitment and additional examinations. Patient recruitment was conducted within the population screening program funded by the Italian NHS.

References

1. Indipendent UK Panel on Breast Cancer. The Benefits and harms of breast cancer screening: an indipendent review. *Lancet*. 2012;380(9855):1778-1786.
2. IARC. International Agency for Research on Cancer. *Handbook in Cancer Prevention*.; 2016. http://publications.iarc.fr/Book-And-Report-Series/Iarc-%0AHandbooks-Of-Cancer-Prevention/Breast-Cancer-Screening-2016%0A
3. Dibden A, Offman J, Parmar D, et al. Reduction in interval cancer rates following the introduction of two-view mammography in the UK breast screening programme. *Br J Cancer*. 2014;110(3):560-564. doi:10.1038/bjc.2013.778
4. Vedantham S, Karellas A, Vijayaraghavan GR, Kopans DB. Digital breast tomosynthesis: State of the art1. *Radiology*. 2015;277(3):663-684. doi:10.1148/radiol.2015141303
5. Houssami N, Zackrisson S, Blazek K, et al. Meta-analysis of prospective studies evaluating breast cancer detection and interval cancer rates for digital breast tomosynthesis versus

mammography population screening. *Eur J Cancer*. 2021;148:14-23. doi:10.1016/j.ejca.2021.01.035

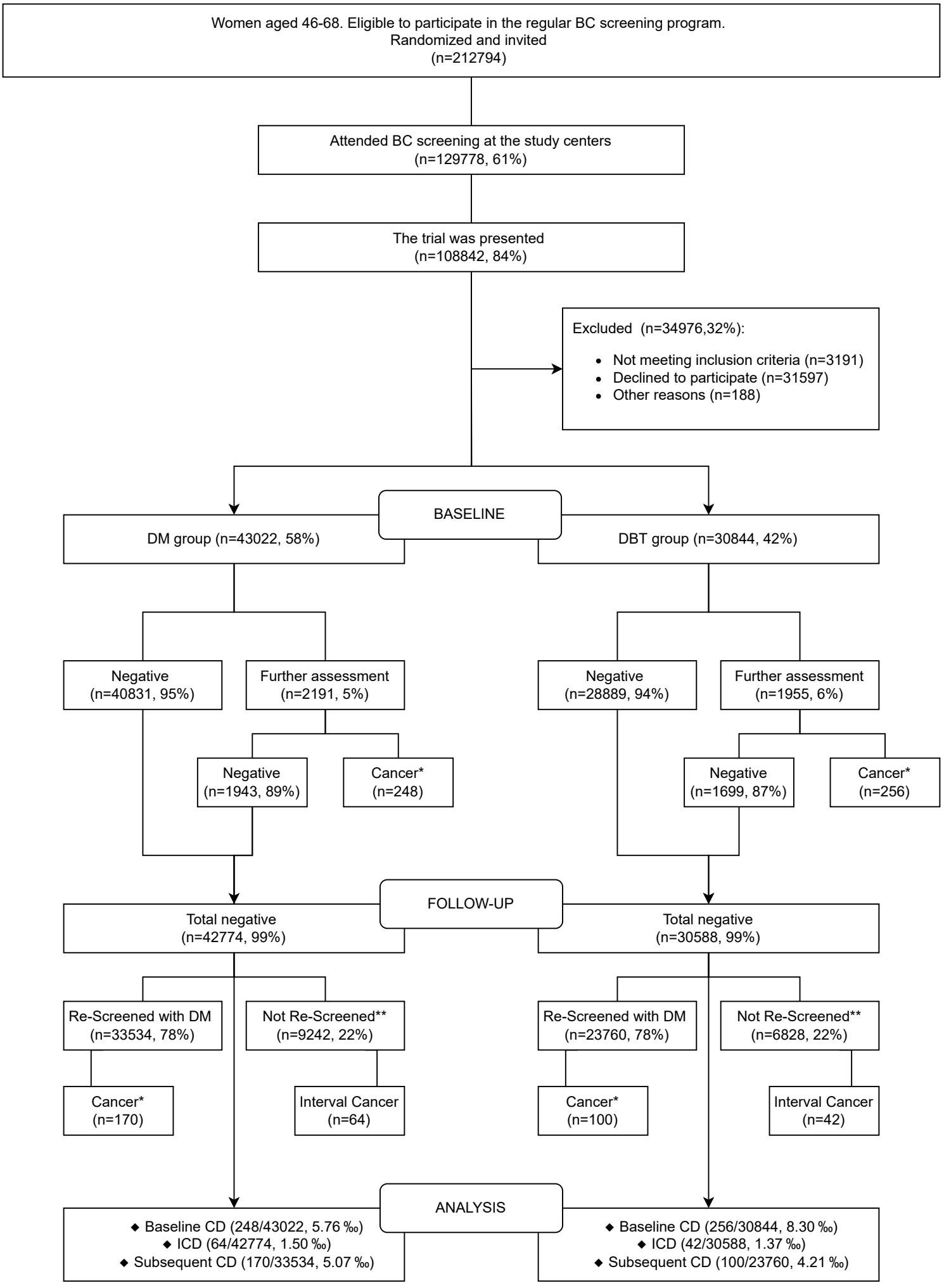
6. Lång K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. *Eur Radiol*. 2016;26(1):184-190. doi:10.1007/s00330-015-3803-3
7. 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. doi:10.1093/jnci/djy121
8. Gilbert FJ, Tucker L, Gillan MG, et al. TOMMY trial: A comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme. *Health Technol Assess (Rockv)*. 2015;19(4):1-136. doi:10.3310/HTA19040
9. Pattacini P, Nitrosi A, Rossi PG, et al. Digital mammography versus digital mammography plus tomosynthesis for breast cancer screening: The reggio emilia tomosynthesis randomized trial. *Radiology*. 2018;288(2):375-385. doi:10.1148/radiol.2018172119
10. Conant EF, Barlow WE, Herschorn SD, et al. Association of Digital Breast Tomosynthesis vs Digital Mammography with Cancer Detection and Recall Rates by Age and Breast Density. *JAMA Oncol*. 2019;5(5):635-642. doi:10.1001/jamaoncol.2018.7078
11. Alabousi M, Zha N, Salameh JP, et al. Digital breast tomosynthesis for breast cancer detection: a diagnostic test accuracy systematic review and meta-analysis. *Eur Radiol*. 2020;30(4):2058-2071. doi:10.1007/s00330-019-06549-2
12. Zackrisson S, Lång K, Rosso A, et al. One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. *Lancet Oncol*. 2018;19(11):1493-1503. doi:10.1016/S1470-2045(18)30521-7
13. Lowry KP, Coley RY, Miglioretti DL, et al. Screening Performance of Digital Breast Tomosynthesis vs Digital Mammography in Community Practice by Patient Age, Screening Round, and Breast Density. *JAMA Netw Open*. 2020;3(7). doi:10.1001/jamanetworkopen.2020.11792
14. 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. doi:10.1016/j.breast.2018.01.002
15. 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. doi:10.1016/j.breast.2019.09.012
16. Hofvind S, Moshina N, Holen ÅS, et al. Interval and Subsequent Round Breast Cancer in a Randomized Controlled Trial Comparing Digital Breast Tomosynthesis and Digital Mammography Screening. *Radiology*. 2021;300(1):66-76. doi:10.1148/radiol.2021203936
17. Hovda T, Holen ÅS, Lång K, 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(2):256-264. doi:10.1148/radiol.2019191337

18. Lee CI, Zhu W, Onega TL, et al. The effect of digital breast tomosynthesis adoption on facility-level breast cancer screening volume. *Am J Roentgenol*. 2018;211(5):957-963. doi:10.2214/AJR.17.19350
19. Perry N, Broeders M, DeWolf C, Al. E. *European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis. Fourth Edition.* European Commission Office for Official Publications of the European Communities, Luxembourg; 2006.
20. Ortenzia O, Rossi R, Bertolini M, Nitrosi A, Ghetti C. Physical characterisation of four different commercial digital breast tomosynthesis systems. *Radiat Prot Dosimetry*. 2018;181(3):277-289. doi:10.1093/RPD/NCY024
21. Distante V, Mano MP, Ponti A, et al. Monitoring surgical treatment of screen-detected breast lesions in Italy. *Eur J Cancer*. 2004;40(7):1006-1012. doi:10.1016/j.ejca.2004.01.015
22. Ponti A, Mano MP, Tomatis M, et al. Audit system on Quality of breast cancer diagnosis and Treatment (QT): Results of quality indicators on screen-detected lesions in Italy, 2011-2012. *Epidemiol Prev*. 2015;39(3):40-47.
23. American Joint Committee on cancer (AJCC). *Cancer Staging Manual*. Eigth edit. (Amin B, Edge S, Greene F, eds.).; 2017.
24. Lin DY, Shen L, Ying Z, Breslow NE. Group Sequential Designs for Monitoring Survival Probabilities. *Biometrics*. 1996;52(3):1033. doi:10.2307/2533064
25. Morrison AS. *Screening in Chronic Disease*. Second edi. Oxford University Press; 1992.
26. StataCorp.2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp. LLC.
27. Trial's registration link. https://clinicaltrials.gov/ct2/show/record/NCT02590315
28. Houssami N, Hofvind S, Soerensen AL, et al. Interval breast cancer rates for digital breast tomosynthesis versus digital mammography population screening: An individual participant data meta-analysis. *EClinicalMedicine*. 2021;34:100804. doi:10.1016/j.eclinm.2021.100804
29. European Commission Initiative on Breast Cancer (ECIBC). European guidelines on breast cancer screening and diagnosis: Use of tomosynthesis. Updated March 12, 2021. Published 2021. https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines/screening-tests
30. Gulati R, Feuer EJ, Etzioni R. Conditions for Valid Empirical Estimates of Cancer Overdiagnosis in Randomized Trials and Population Studies Am J Epidemiol 2016 Jul 15;184(2):140-7.
31. Canelo-Aybar C, Ferreira DS, Ballesteros M, et al. Benefits and harms of breast cancer mammography screening for women at average risk of breast cancer: A systematic review for the European Commission Initiative on Breast Cancer J Med Screen 2021 Dec;28(4):389-404. doi: 10.1177/0969141321993866.

Figure legend

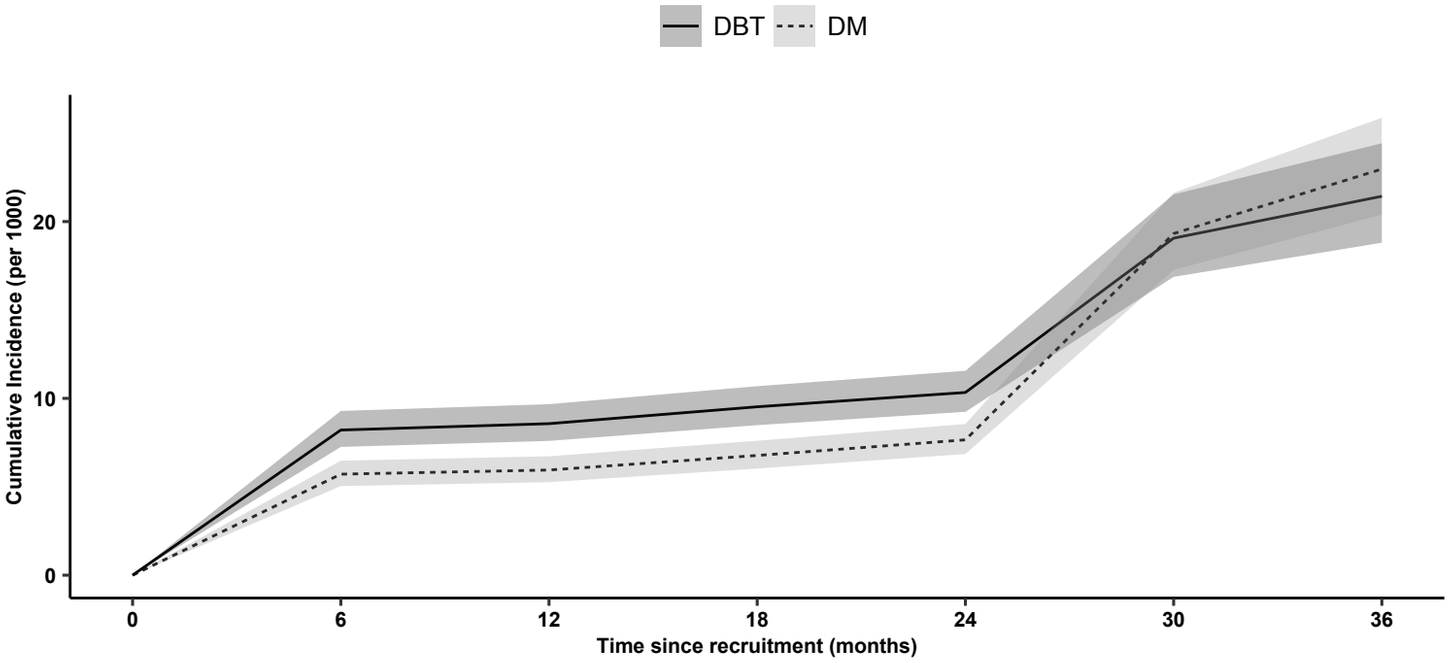
Figure 1. Flowchart of recruitment phase and study design.
DBT = digital breast tomosynthesis, DM = digital mammography, CD = cancer detection , ICD = interval cancer detection, including all malignant cases detected during the inter-screening interval, within the study follow-up period (range, 0-36 months) * Cancer - inclusive of all invasive and in situ malignant lesions

Figure 2. Nelson-Aalen Cumulative Incidence Estimates for malignant (A) and benign (B) breast lesions according to screening arm.



* Cancer - inclusive of all invasive and in situ malignant lesions
** Assessed for interval cancer

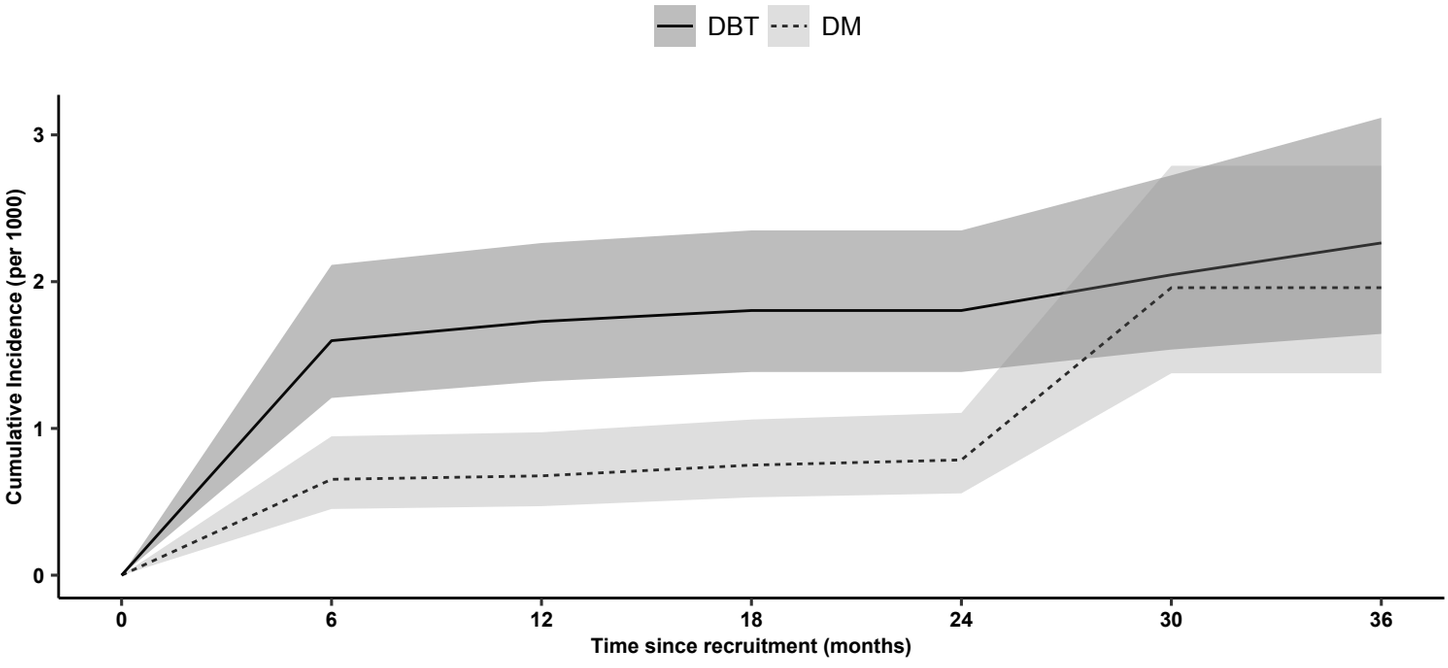
(A)



Number at risk

DM	43022 (245)	42696 (10)	42484 (32)	37223 (30)	27753 (145)	6589 (19)	3663
DBT	30844 (252)	30511 (11)	30359 (26)	26427 (18)	19456 (80)	4787 (9)	2313

(B)



Number at risk

DM	43022 (28)	42696 (1)	42484 (3)	37223 (1)	27753 (14)	6589 (0)	3663
DBT	30844 (49)	30511 (4)	30359 (2)	26427 (0)	19456 (2)	4787 (1)	2313

Table 1. Baseline results according to arm

	DBT (Study arm) N=30844	DM (Control arm) N=43022	
Parameter	**n (%)**	**n (%)**	**RR (95%CI)**
Recalls	1955 (6·34)	2191 (5·09)	1·24 (1·17-1·32)
Surgery referrals	305 (0·99)	276 (0·64)	1·54 (1·31-1·82)
Invasive cancers	224 (0·73)	214 (0·50)	1·46 (1·21-1·77)
DCIS	32 (0·10)	32 (0·07)	1·39 (0·83-2·35)
B5	0 (0·00)	2 (0·00)	NA
Unknown	5 (0·02)	6 (0·02)	1·16 (0·28-4·57)
Cancers detected (CD)	256 (0·83)	248 (0·58)	1·44 (1·21-1·71)
PPV	256/1955 (13·09)	248/2191 (11·32)	p>0·05
Benign	52 (0·17)	29 (0·07)	2·50 (1·59-3·94)
Benign/malignant ratio	52/256 (0·20)	29/248 (0·12)	p<0·05
False positives	1699 (5·51)	1943 (4·52)	1·22 (1·14-1·30)
Characteristic of the invasive screen detected lesions	**n (‰)**	**n (‰)**	**RR (95%CI)**
Invasive	224 (7·26)	214 (4·97)	1·46 (1·21-1·77)
pT			
<=1c	181 (5·87)	164 (3·81)	1·54 (1·24-1·92)
>1c	25 (0·81)	31 (0·72)	1·12 (0·64-1·97)
Unknown	18 (0·58)	19 (0·44)	1·32 (0·65-2·66)
Lymph nodes			
Negative	155 (5·03)	140 (3·25)	1·54 (1·23-1·97)
Positive	35 (1·13)	42 (0·98)	1·16 (0·72-1·86)
Unknown	34 (1·10)	32 (0·74)	1·48 (0·89-2·48)
Stage			
<=I	141 (4·57)	118 (2·74)	1·67 (1·30-2·15)
II+	38 (1·23)	53 (1·23)	1·00 (0·64-1·55)
Unknown	45 (1·46)	43 (1·00)	1·46 (0·94-2·27)
Histologic Grade			
1	88 (2·85)	71 (1·65)	1·73 (1·25-2·40)
2	84 (2·72)	85 (1·98)	1·38 (1·01-1·89)
3	17 (0·55)	22 (0·51)	1·08 (0·54-2·13)
Unknown	35 (1·13)	36 (0·84)	1·36 (0·83-2·22)

Subjects recalled to further assessment but without a final round result (17 in the study arm and 14 in the control arm) are not included.

CD = cancer detection, DBT = digital breast tomosynthesis, DCIS = ductal carcinoma in situ, DM = digital mammography, RR = relative risk, 95%CI = 95% confidence interval, PPV = positive predictive value, NA = not applicable. PPV was calculated as all cancers found in the total number of recalled women. A false positive result occurred when a woman was recalled but did not have cancer.

Table 2. Subsequent screening results according to arm

	DBT (Study arm) N=23760	DM (Control arm) N=33534	
Parameter	**n (%)**	**n (%)**	**RR (95%CI)**
Recalls	1000 (4·21)	1456 (4·34)	0·97 (0·89-1·05)
Surgery referrals	103 (0·43)	191 (0·57)	0·76 (0·59-0·97)
Invasive cancers	81 (0·34)	135 (0·40)	0·85 (0·63-1·12)
DCIS	17 (0·07)	32 (0·09)	0·75 (0·39-1·39)
B5	2 (0·01)	3 (0·01)	0·94 (0·08-8·21)
Unknown	3 (0·01)	8 (0·02)	0·53 (0·09-2·20)
Cancers detected (CD)	100 (0·42)	170 (0·50)	0·83 (0·65-1·06)
PPV	100/1000 (10·00)	170/1456 (11·68)	p>0·05
Benign	6 (0·25)	19 (0·57)	0·45 (0·18-1·12)
Benign/malignant ratio	6/100 (0·06)	19/170 (0·11)	p>0·05
No. of false-positive results	900 (3·79)	1286 (3·83)	0·99 (0·91-1·08)
Characteristic of the invasive screen detected lesions	**n (‰)**	**n (‰)**	**RR (95%CI)**
Invasive	81 (3·41)	135 (4·03)	0·85 (0·63-1·12)
pT			
<=1c	64 (2·69)	105 (3·13)	0·86 (0·62-1·18)
>1c	10 (0·42)	19 (0·57)	0·74 (0·31-1·68)
Unknown	7 (0·29)	11 (0·33)	0·90 (0·30-2·54)
Lymph nodes			
Negative	55 (2·31)	86 (2·56)	0·90 (0·63-1·28)
Positive	14 (0·59)	28 (0·83)	0·71 (0·34-1·39)
Unknown	12 (0·51)	21 (0·63)	0·81 (0·36-1·72)
Stage			
<=I	50 (2·10)	78 (2·33)	0·90 (0·62-1·31)
II+	17 (0·72)	37 (1·10)	0·65 (0·34-1·18)
Unknown	14 (0·59)	20 (0·60)	0·99 (0·46-2·06)
Histologic Grade			
1	28 (1·18)	39 (1·16)	1·01 (0·60-1·69)
2	31 (1·30)	59 (1·76)	0·74 (0·46-1·16)
3	12 (0·51)	14 (0·42)	1·21 (0·51-2·82)
Unknown	10 (0·42)	23 (0·69)	0·61 (0·26-1·34)

Subjects recalled to further assessment but without a final round result (8 in the study arm and 16 in the control arm) are not included.

CD = cancer detection, DBT = digital breast tomosynthesis, DCIS = ductal carcinoma in situ, DM = digital mammography, RR = relative risk, 95%CI = 95% confidence interval, PPV = positive predictive value. PPV was calculated as all cancers found in the total number of recalled women. A false positive result occurred when a woman was recalled but did not have cancer.

Table 3. Interval cancers according to arm

	DBT (Study arm) N=30588	DM (Control arm) N=42774	
Parameter	**n (‰)**	**n (‰)**	**RR (95%CI)**
Invasive cancers	38 (1·24)	58 (1·36)	0·92 (0·59-1·40)
DCIS	4 (0·13)	5 (0·12)	1·12 (0·22-5·20)
B5	0 (0·00)	0 (0·00)	NA
Unknown	0 (0·00)	1 (0·02)	NA
Total	42 (1·37)	64 (1·50)	0·92 (0·62-1·35)
Characteristic of the invasive interval cancers	**n (‰)**	**n (‰)**	**RR (95%CI)**
Invasive	38 (1·24)	58 (1·36)	0·92 (0·59-1·40)
pT			
<=1c	19 (0·62)	26 (0·61)	1·02 (0·53-1·92)
>1c	3 (0·10)	8 (0·19)	0·52 (0·09-2·18)
Unknown	16 (0·52)	24 (0·56)	0·93 (0·46-1·83)
Lymph nodes			
Negative	15 (0·49)	23 (0·54)	0·91 (0·44-1·82)
Positive	6 (0·20)	7 (0·16)	1·20 (0·33-4·17)
Unknown	17 (0·56)	28 (0·65)	0·85 (0·44-1·61)
Stage			
<=I	14 (0·46)	20 (0·47)	0·98 (0·46-2·04)
II+	8 (0·26)	12 (0·28)	0·93 (0·33-2·48)
Unknown	16 (0·52)	26 (0·61)	0·86 (0·43-1·67)
Histologic Grade			
1	7 (0·23)	9 (0·21)	1·09 (0·34-3·28)
2	7 (0·23)	9 (0·21)	1·09 (0·34-3·28)
3	5 (0·16)	13 (0·30)	0·54 (0·15-1·61)
Unknown	19 (0·62)	27 (0·63)	0·98 (0·52-1·84)

DBT = digital breast tomosynthesis, DCIS = ductal carcinoma in situ, DM = digital mammography, RR=relative risk, 95%CI = 95% confidence interval, NA = not applicable.

Table 4. Cumulative (baseline + interval cancers + subsequent screening) detection

	DBT (Study arm)			DM (Control arm)				
	n	***N***	***Detection* (‰)**	***n***	***N***	***Detection* (‰)**	**RR (95%CI)**	**RD (95%CI)**
Malignant lesions								
Baseline screening	256	30844	8·30 (7·30-9·40)	248	43022	5·76 (5·07-6·52)	1·44 (1·21-1·71)	2·54 (1·30-3·78)
Interval cancers	42	30588	1·37 (0·99-1·90)	64	42774	1·50 (5·07-6·52)	0·92 (0·62-1·35)	0·12 (-0·68-0·43)
Subsequent screening	100	23760	4·21 (3·40-5·10)	170	33534	5·07 (4·34-5·89)	0·83 (0·65-1·06)	-0·86 (-1·98-0·26)
Cumulative detection	398	30844	12·90 (11·67-14·23)	482	43022	11·20 (10·23-12·24)	1·15 (1·01-1·31)	1·70 (0·10-3·30)
Benign lesions								
Baseline screening	52	30844	1·69 (1·26-2·21)	29	43022	0·67 (0·45-0·97)	2·50 (1·59-3·94)	1·01 (0·49-1·53)
Subsequent screening	6	23760	0·25 (0·09-0·55)	19	33534	0·57 (0·34-0·88)	0·45 (0·18-1·12)	-0·31 (-0·64-0·01)
Cumulative detection	58	30844	1·88 (1·43-2·43)	48	43022	1·12 (0·82-1·48)	1·69 (1·15-2·47)	0·76 (0·19-1·34)

DBT = digital breast tomosynthesis, DM = digital mammography, RR = relative risk, RD = risk difference, 95%CI = 95% confidence interval.