Summary of Outcomes from Consecutive Years of Tomosynthesis Screening at an American Academic Institution

Emily F. Conant[1(✉)], Andrew Oustimov[1], Samantha P. Zuckerman[1], Elizabeth S. McDonald[1], Susan P. Weinstein[1], Andrew D.A. Maidment[1], Bruno Barufaldi[1], Marie Synnestvedt[2], and Mitchell Schnall[1]

[1] Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
emily.conant@uphs.upenn.edu
[2] Division of General Internal Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Abstract. Digital breast tomosynthesis (DBT) screening outcomes are sustainable over consecutive years with significant reductions in recall and increasing cancers per recalled patients compared to screening with digital mammography alone (DM). There is a prevalence effect with a reduction in cancer detection at the second round of screening that is no longer present at the third round. There is a non-statistically significant trend of decreased interval cancers with DBT compared to DM alone screening. Early data on the implementation of synthetic 2D (s2D) imaging coupled with DBT shows maintenance of screening outcomes with reduction in radiation dose compared to DM/DBT screening .

Keywords: Breast cancer · Screening · Mammography · Digital breast tomosynthesis

1 Introduction

Screening with a combination of digital mammography and digital breast tomosynthesis (DBT) has been shown to reduce false positives and improve cancer detection compared to screening with digital mammography (DM) alone. The bulk of the evidence on the improved outcomes achievable with DBT has come from first round screening in either limited prospective trials [1–3] or retrospective studies, mostly from single sites [4–7]. Questions remain whether the improved outcomes are sustainable over multiple rounds of screening and whether the implementation of synthetic 2D imaging as a method to reduce the dose of DBT screening will have similarly improved outcomes. There are also few reports on the rate of false negative DBT screens.

At the University of Pennsylvania, we are entering our 5th year of screening all patients with DBT and therefore have the data to analyze the "prevalence effect" of screening with this new modality. We also have tumor registry follow-up from the first few years of screening to begin to evaluate false negative rates. Finally, over the last year of screening, we have implemented synthetic 2D images and removed the DM

A. Tingberg et al. (Eds.): IWDM 2016, LNCS 9699, pp. 20–24, 2016.
DOI: 10.1007/978-3-319-41546-8_3

acquisition of the combination DM/DBT mode of screening. A summary of our DBT screening outcomes will be reviewed [8].

2 Method

An IRB approved, retrospective analysis of DBT screening metrics is ongoing at our institution. Screening outcomes were evaluated for a total of 44468 examinations attributable to 23,958 women over a 4 year period (year 0 cohort (DM0) n = 10728, year 1 cohort (DBT1) n = 11007, year 2 cohort (DBT2) n = 11157, year 3 cohort (DBT3) n = 11576. For women screened with DBT, 21,395 women were screened at least once, 9316 women were screened at least twice and 3023 women underwent 3 rounds of screening. Outcome metrics (recall and cancer detection rates and PPV1) were calculated for all patients presenting for screening over 4 consecutive years (DM, year 0; DBT, years 1–3). Interval cancer rates for DM0 and DBT1, defined as cancer presenting as symptomatic within one year of a negative screen, were calculated after linkage with the state cancer registry thru June 2014 [8].

In addition, preliminary data on recall and cancer detection rates and average glandular dose from the first 6 months (January 1 – June 30th, 2015) of screening with synthetic 2D combined with DBT (s2D/DBT) were compared with similar published metrics from DBT year 1 [9].

3 Results

There was no practical difference in the patient characteristics including age, breast density, race and screening volumes race across the years 0–3. Additionally, we have previously demonstrated no statistically significant difference in calculated breast cancer risk between DM year 0 and the first 18 months of screening with DBT [5].

Table 1 demonstrates that at the site level, recall rates rose slightly for DBT1–3 years (88, 90, 92/1000 screened, respectively) but remained significant reduced compared to DM0 rate of 104/1000 screened. There was also no statistically significant difference in the recall rate across the three DBT years (p = 0.549). Cancer detection rate continued to increase at 4.6, 5.5, 5.8, and 6.1/1000 screened for years 0, 1, 2, and 3 respectively, but was not significantly different from that of DM (DM vs DBT1 p = 0.370, DM vs DBT2 p = 0.196, DM vs DBT3 p = 0.110), and was not significantly different across three DBT years (p = 0.796). Although, the rate of invasive cancers detected per 1000 screened increased slightly over time (DM = 3.2, DBT1 = 3.8, DBT2 = 4.1, DBT3 = 4.1), the increase, in any DBT year compared with DM, or across the DBT years, was not significant. The positive predictive value for recalls (PPV1) continued to rise from DM0 rate of 4.4 % to 6.2 % (p = 0.06), 6.5 % (p = 0.03), 6.7 % (p = 0.02) for DBT1–3 (Table 1). The biopsy rate in each DBT year did not significantly differ from that of DM (DM vs DBT1 p = 0.167, DM vs DBT2 p = 0.606, DM vs DBT3 p = 0.597) and did not differ significantly over three DBT years (p = 0.600). Complete cancer registry data were only available to assess DM0 and DBT 1 years. The change in interval

cancer rates/1000 screened across these years showed a downward trend (DM = 0.7, DBT1 = 0.5) but was not statistically significant (p = 0.603).

Table 1. Site level analysis of screening outcomes comparing each digital breast tomosynthesis year (DBT1–3) with the digital mammography year (DM0), as well as across the three DBT years (modified from Reference [8])

Characteristic	DM 0	DBT 1	*P*	DBT 2	*P*	DBT 3	*P*	*P* (DBT)
Recall, n/1000	104	88	<0.0001	90	0.0005	92	0.0025	0.549
Biopsy, n/1000	18	20	0.167	19	0.606	19	0.597	0.600
Cancer, n/1000	4.6	5.5	0.370	5.8	0.196	6.1	0.110	0.796
PPV 1, %	4.4	6.2	0.063	6.5	0.034	6.7	0.020	0.915
Invasive cancer, n/1000	3.2	3.8	0.420	4.1	0.243	4.1	0.269	0.929
In-situ cancer, n/1000	1.4	1.5	0.779	1.3	0.914	1.8	0.440	0.668
Interval cancer, n/1000	0.7	0.5	0.603	NA	NA	NA	NA	0.069

Note— *P*s represent Pearson chi-square and Fisher's exact test p-values, comparing proportions across each of the DBT years (1–3) with the baseline DM0 year. The *P* (DBT) column contains the p-values for testing null hypothesis that values are not significantly different across DBT 1–3 years. The NAs in the interval cancer line (bottom row) indicate that adequate cancer registry follow up was not available for DBT years 2 and 3. PPV1 is defined as cancers/recalled cases.

Table 2 demonstrates at the individual level, the outcomes for women screened with DBT at a first, second and third round. Recall rate decreased at each consecutive round of screening from 103 to 69 to 59/1000 screened, from the first round to third, respectively. Cancer detection rates were 6.3, 4.2 and 7.3/1000 screened at the first, second and third rounds of screening, respectively. The decrease in the cancer detection rate in the second round suggests a prevalence or, first round screening effect that was no longer evident at the third round. PPV1 showed a similar reduction at the second round and an increase at the third round.

Table 2. Recall, cancer detection, and PPV 1 for first, second, and third round of digital breast tomosynthesis (DBT) screening **(modified from Reference [8])**

	One screen	Two screens	Three screens
	21395 (100 %)	9316 (44 %)	3023 (14 %)
Recall, n/1000	103	69.2	59
Cancer, n/1000	6.3	4.2	7.3
PPV 1, %	6.1	6.1	12.4
Recall, n	2201	645	177
Cancer, n	135	39	22

When comparing the screening metrics from the first six months of implementing s2D with DBT compared to DBT year 1, there was a reduction in recall from 88 to 71/1000 screened (p < 0.001). The cancer detection rate for s2D/DBT was not significantly different from DBT1: 5.03 vs 5.5/1000 women screened, respectively (p = 0.72). The reduction in recalls and maintenance of the overall cancer detection rate was

accompanied by a 39 % reduction in average glandular radiation dose (4.88 mGy vs 7.99 mGy, $p < 0.001$).

4 Discussion

The continued controversy surrounding mammographic screening revolves mainly around the so-called "harms" of a false positive screen. Incorporating DBT in screening has clearly shown that the new modality can reduce false positives while increasing cancer detection, specifically invasive cancers [1, 4]. Our study demonstrates that while there may be a small prevalence effect with the first round of screening, the improved outcomes achieve with DBT screening are sustainable and associated with a trend of decreasing interval cancers. However, the improved outcomes achieved with DM/DBT screening are obtained with an increase in radiation dose. Reconstructed or, "synthetic 2D DM" images offer the opportunity to significantly reduce patient dose over a lifetime of screening. In addition, our preliminary results with synthetic 2D imaging suggest that s2D is a non-inferior replacement of the DM portion of combination DM/DBT imaging. Screening with s2D/DBT allows for the benefits of DBT with a significant decrease in increased radiation dose compared to DM/DBT.

References

1. Skaane, P., Bandos, A.I., Eben, E.B., Jebsen, I.N., Krager, M., Haakenaasen, U., Ekseth, U., Izadi, M., Hofvind, S., Gullien, R.: Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. Radiology **271**(3), 655–663 (2014)
2. Ciatto, S., Houssami, N., Bernardi, D., Caumo, F., Pellegrini, M., Brunelli, S., Tuttobene, P., Bricolo, P., Fanto, C., Valentini, M., Montemezzi, S., Macaskill, P.: Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol. **14**(7), 583–589 (2013)
3. Gilbert, F.J., Tucker, L., Gillan, M.G., et al.: The TOMMY trial: a comparison of TOMosynthesis with digital MammographY in the UK NHS breast screening programme–a multicentre retrospective reading study comparing the diagnostic performance of digital breast tomosynthesis and digital mammography with digital mammography alone. Health Technol. Assess. **19**(i–xxv), 1–136 (2015)
4. Friedewald, S.M., Rafferty, E.A., Conant, E.F.: Breast cancer screening with tomosynthesis and digital mammography-reply. JAMA **312**, 1695–1696 (2014)
5. McCarthy, A.M., Kontos, D., Synnestvedt, M. et al.: Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. J. Natl. Cancer Inst. **106**, 1–7 (2014)
6. Rose, S.L., Tidwell, A.L., Bujnoch, L.J., et al.: Implementation of breast tomosynthesis in a routine screening practice: an observational study. AJR Am. J. Roentgenol. **200**, 1401–1408 (2013)
7. Haas, B.M., Kalra, V., Geisel, J., et al.: Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. Radiology **269**, 694–700 (2013)

8. McDonald, E.S., Oustimov, A., Weinstein, S.P., Synnestvedt, M., Schnall, M., Conant, E.C.: Effectiveness of tomosynthesis in combination with digital mammography: outcomes analysis from three years of breast cancer screening. JAMA Oncol. (2016, in press)
9. Zuckerman, S.P., Conant, E.F., Weinstein, S.P., Synnestvedt, M., Korhonen, K.E., McDonald, E.S.: Early implementation of synthesized 2D in screening with digital breast tomosynthesis: a pictorial essay of early outcomes. Presented at the 101st RSNA Annual Meeting, Chicago, IL, 29 November – 4 December 2015