The Breast 38 (2018) 150–153

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

Interval breast cancers in the 'screening with *t*omosynthesis *or* standard *m*ammography' (STORM) population-based trial

Nehmat Houssami [a, *], Daniela Bernardi [b], Francesca Caumo [c, d], Silvia Brunelli [c], Carmine Fantò [b], Marvi Valentini [b], Giovanna Romanucci [c], Maria A. Gentilini [e], Manuel Zorzi [f], Petra Macaskill [a]

[a] *Sydney School of Public Health, Sydney Medical School, University of Sydney, Sydney, Australia*
[b] *U.O. Senologia Clinica e Screening Mammografico, Department of Diagnostics, Azienda Provinciale Servizi Sanitari (APSS), Trento, Italy*
[c] *Breast Unit ULSS9, Ospedale di Marzana, Verona, Italy*
[d] *Veneto Institute of Oncology IOV–IRCC Comprehensive Cancer Centre, Padua, Italy*
[e] *Servizio di Epidemiologia Clinica Valutativa, Azienda Provinciale Servizi Sanitari (APSS), Trento, Italy*
[f] *Veneto Tumour Registry, Veneto Region, Padua, Italy*

ARTICLE INFO

Article history:
Received 23 November 2017
Received in revised form
20 December 2017
Accepted 4 January 2018

Keywords:
Digital breast tomosynthesis
Interval breast cancer
Mammography
Population screening
Sensitivity

ABSTRACT

Background & methods: The prospective 'screening with *t*omosynthesis *or* standard *m*ammography' (STORM) trial recruited women participating in biennial breast screening in Italy (2011–2012), and compared sequential screen-readings based on 2D-mammography alone or based on tomosynthesis (integrated 2D/3D-mammography). The STORM trial showed that tomosynthesis screen-reading significantly increased breast cancer detection compared to 2D-mammography alone. The present study completes reporting of the trial by examining interval breast cancers ascertained at two year follow-up.
Results: 9 interval breast cancers were identified; the estimated interval cancer rate was 1.23/1000 screens [9/7292] (95%CI 0.56 to 2.34) or 1.24/1000 negative screens [9/7235] (95%CI 0.57 to 2.36). In concurrently screened women who attended the same screening services and received 2D-mammography, interval cancer rate was 1.60/1000 screens [40/25,058] (95% CI 1.14 to 2.17) or 1.61/1000 negative screens [40/24,922] (95% CI 1.15 to 2.18). Estimated screening sensitivity for the STORM trial was 85.5% [59/69] (95%CI 75.0%–92.8%), and that for 2D-mammography screening was 77.3% [136/176] (95%CI 70.4%–83.2%).
Conclusion: Interval breast cancer rate amongst screening participants in the STORM trial was marginally lower (and screening sensitivity higher) than estimates amongst 2D-screened women; these findings should be interpreted with caution given the small number of interval cases and the sample size of the trial. Much larger screening studies, or pooled analyses, are required to examine interval cancer rates arising after breast tomosynthesis screening versus digital mammography screening.

1. Introduction

Findings from the first prospective population screening trial using digital breast tomosynthesis (3D-mammography technology) for breast cancer screening were reported in 2013 [1]. The 'screening with *t*omosynthesis *or* standard *m*ammography' (STORM) trial recruited women participating in biennial breast screening, and compared sequential screen-readings based on 2D-mammography alone or based on *integrated* 2D/3D-mammography [1,2]. The trial showed that integrating 3D with 2D-mammography significantly increased breast cancer detection compared with 2D-mammography screening [1,2]. Prospective trials [3–5] and several retrospective studies [6–9] have subsequently reported improved detection metrics using tomosynthesis (3D-mammography) screening relative to standard screening with 2D-mammography alone. Almost all published evidence on tomosynthesis technology for population screening has focused on initial detection metrics at screening with little to no evidence reported on interval breast cancers (cancers diagnosed after a negative mammographic screen

* Corresponding author. Sydney School of Public Health (A27), Sydney Medical School, University of Sydney, Sydney, 2006, Australia.
E-mail address: nehmat.houssami@sydney.edu.au (N. Houssami).

https://doi.org/10.1016/j.breast.2018.01.002
0960-9776/

and before the next routine screen [10]) at follow-up.

We now complete the STORM trial results by reporting on interval breast cancers, based on ascertainment at two year follow-up from screening examinations. We do not re-report all screening detection measures for the study because these have been previously published in our primary and secondary analyses of the trial [1,2,11,12]. In the present work we focus on screening measures that require completed ascertainment of interval cancers, hence we estimate interval cancer rates and screening sensitivity for the trial. We additionally present interval cancer rates for concurrent cohorts screened with 2D-mammography alone to assist interpretation of trial findings.

2. Methods

The population-based STORM screening trial prospectively recruited asymptomatic women attending biennial screening services in Trento and Verona, Italy, August 2011 to June 2012 [1]. A detailed description of the study methods and population, and primary results, have been reported by Ciatto and colleagues [1], with secondary analyses presented in related publications [2,11,12]. Therefore here we briefly describe the study methods. Screening participants were invited to have integrated 2D/3D mammography screening, and those opting not to participate had (standard) 2D-mammography [1]. The trial was granted institutional ethics approval, and informed consent was obtained from participants [1].

Participants in STORM had digital mammography using a Selenia Dimensions Unit integrating 3D acquisitions (COMBO©; Hologic, Inc. Bedford MA, USA): each of the 2D and 3D acquisitions were obtained in cranio-caudal and mediolateral oblique views. Screen-reading was based on independent double-reading by radiologists experienced in mammography; screens were interpreted sequentially initially using standard 2D-mammography alone, and were then re-interpreted by the same radiologists using integrated 2D/3D mammography [1,2]. Screening examinations interpreted as positive by either screen-reader were recalled to assessment, which typically included further imaging (additional views, ultrasound) with needle biopsy where indicated by the imaging assessment.

2.1. Ascertainment of interval breast cancers

Population-based breast screening programs in Italy are routinely required to report interval cancer data as part of quality assurance processes. Interval cancers for the timeframe of the study were identified using a combination of (a) checking unique record numbers for screening episodes against local hospital and pathology databases; and (b) checking with the local cancer registry for cancer notifications at 24 months from screening episode date. In an earlier publication of the trial, we included results based on a minimum follow-up of 13 months for preliminary identification of interval cancers [2]. However that follow-up was incomplete because screening services in Trento and Verona routinely provide biennial screening therefore ascertainment of interval cases requires 24-month follow-up for all screens – complete ascertainment was undertaken for the present study. Because interval cancer ascertainment was performed over the trial timeframe for both Trento and Verona screening services, this also allowed identification of interval cases amongst the cohort of women screened in the same services using 2D-mammography (2011–2012), to provide contextual information that may assist interpretation of trial results.

2.2. Statistical considerations and methods

The STORM trial was powered for comparison of cancer detection as the primary end-point (described in Ciatto et al.) [1]. Both the design (within woman comparison) and sample size of the trial do not support comparative analysis of interval cancer rates – the latter requires very large datasets as outlined in a recent meta-analysis protocol [13]. For completion and transparency we report descriptive information on interval breast cancers (number and characteristics) and estimate the interval cancer rate for the STORM trial. We applied a standard definition of interval cancers, namely cancers identified after a negative mammographic screen result and before the next routine screen [10] (using biennial screening as the context). Because interval cancer rates may be estimated at the screened population level (per 1000 screens) or may be estimated based on negative screens (per 1000 negative screens), we report estimates for each of these definitions. We also estimated the interval cancer rate for the concurrent cohort of screening participants attending the same screening services that had 2D-mammography screening. Screening sensitivity was calculated as the number of cancers detected at screening from all cancers observed in study participants inclusive of interval cancers. For all estimated rates and proportions, exact (Clopper-Pearson) 95% confidence intervals were computed using StatsDirect v3.0.193 software [14].

3. Results

There were 7292 screening participants in the STORM trial: we reported that 59 breast cancers were detected at screening in 57 subjects and described the tumour characteristics of these cancers in our initial report [1]. In the present report, based on ascertainment at 24-month follow-up for all screening examinations, there were 9 interval breast cancers: 8 presented in participants who developed breast symptoms, and one was identified at spontaneous screening at 22 months from the index screening examination. We did not classify as an interval case one false-negative assessment because that case was positive at the index screen obtained as part of the trial [2] (however this was included in estimation of screening sensitivity). The estimated interval cancer rate for the STORM trial was 1.23/1000 screens [9/7292] (95%CI 0.56 to 2.34) or 1.24/1000 negative screens [9/7235] (95%CI 0.57 to 2.36). Descriptive data on the 9 interval breast cancers identified amongst participants are shown in Table 1, including tumour characteristics and biomarker profile. Screening sensitivity, based on integrated 2D/3D-mammography in the STORM trial, was 85.5% [59/69] (95% CI 75.0%–92.8%).

In the concurrent group of women who had attended the same screening services and received 25,058 2D (digital) mammography screens, 40 interval breast cancers were identified: the estimated interval cancer rate was 1.60/1000 screens [40/25,058] (95% CI 1.14 to 2.17) or 1.61/1000 negative screens [40/24,922] (95% CI 1.15 to 2.18). Screening sensitivity for 2D-mammography screening was 77.3% [136/176] (95%CI 70.4%–83.2%).

4. Discussion

The STORM trial was the world's first prospective population-based screening trial that used 3D mammography technology with 2D mammography for breast cancer screening and that reported screen-detection metrics for the completed study [1]. Those initial results focused on breast cancer detection as the primary end-point of the trial; we now complement those results by reporting on interval breast cancers as the final outcome from the trial, based on ascertainment of all screens at 24 months follow-up. The STORM trial was not designed to compare interval cancer data as an end-point – the data we report are primarily for estimating interval cancer rate and screening sensitivity, hence declaring all

Table 1
Characteristics of interval breast cancers observed in the STORM trial.

Case	Age at screen (years)	Density[a]	Interval[b]	Cancer histology	Tumour size mm	Tumour grade	Lymph node (N) status	ER/PR/HER2	MIB-1/Ki-67%
1	60	3	Year 1	Invasive ductal	22	3	N1*micro*	+/−/+	60
2	58	2	Year 1	Invasive ductal	30	3	N0	+/−/+	30
3	69	3	Year 1	DCIS (Paget's disease)	20	High-grade DCIS	N0	−/−/NR	8
4	60	3	Year 2	Invasive ductal	30	3	N3	−/−/+	36
5	60	2	Year 2	Invasive ductal	24	3	N0	+/+/−	38
6	50	4	Year 2	Invasive ductal	6	1	N0	+/+/−	10
7	65	2	Year 2	Invasive ductal	35	3	N0	−/−/−	70
8	51	2	Year 2	Invasive lobular	19	2	N0	+/+/−	9
9	51	3	Year 2	Invasive ductal, multifocal	12 (index)	2	N0	+/+/−	12

Key: (+) positive; (−) negative; (ER) oestrogen receptor; (PR) progesterone receptor; (HER2) human epidermal growth factor receptor 2; (NR) not reported; (DCIS) ductal carcinoma in-situ.
[a] Based on BI-RADS classification using mammographic breast density at digital mammography [1].
[b] Whether diagnosed in year 1 or year 2 (from negative screen) in biennial screening.

potentially informative findings from the study. Our estimated interval cancer rate (1.23 or 1.24/1000) may also help inform future screening studies or could potentially contribute to collective analyses [13]. The estimated interval breast cancer rate for concurrent 2D-mammography screened women (1.60/1000 or 1.61/1000) allows us to contextualise the interval cancer rate from the trial however we have not attempted to compare these interval cancer rates because the trial sample size does not support valid comparisons for this outcome. Comparison of interval cancer rates requires much larger screening datasets than our study, as pointed out in a published protocol aimed at future collective analyses from breast screening trials [13].

We also cannot compare our screening sensitivity of 85.5% to other prospective population-based trials of tomosynthesis screening that were embedded in biennial screening programs, because none of those trials have yet reported their final study results inclusive of interval cancers [3–5]. However, our estimated screening sensitivity compares favourably to recently reported estimates from biennial screening programs in Europe and Australia (75.5%–79%) [15,16] in the setting of digital (2D) mammography screening, and also to the sensitivity of 2D-mammography screening (77.3%) in our concurrent cohort (see Results). We avoid direct comparison of our interval cancer rates with interval cancer rates recently reported from other countries and programs because these are influenced by underlying breast cancer risk in the screened population and also by the methods used to ascertain interval cases – we merely point out that our estimates are similar to, or lower than, those reported from biennial screening practice based on digital mammography screening [15,16].

A key limitation of this report is that the small number of interval cancers precludes analytic evaluation; however, we note the predominance of high grade and large tumours, the presence of nodal macro-metastases in only one case, and the heterogeneity in biomarker profile inherent in breast tumours. These findings are generally in keeping with those from a recent overview of interval breast cancers observed in population breast screening programs based on 2D-mammography screening [10]. The finding that interval cases more frequently emerged in year 2 than year 1of the inter-screen time interval (Table 1) is also in line with reported data from biennial screening programs [10].

We complete our reporting of the STORM trial concluding that the interval breast cancer rate amongst screening participants in the trial (1.23 or 1.24/1000) was marginally lower than that estimated for concurrent 2D-screened cohorts; and that screening sensitivity using 2D/3D-mammography in the trial was higher than that observed for 2D-screened women. These results should be interpreted with caution given the small number of interval cases and the sample size of the trial, and noting that the trial was not planned for comparison of interval cancer rates. Interval breast cancers in the trial had characteristics that were generally similar to those described for interval cancers emerging following conventional mammography screening. To determine the effect on interval cancer rates from tomosynthesis population screening, much larger screening studies or pooled analyses are required to compare interval cancer rates amongst tomosynthesis-screened women with those estimated amongst women screened only with digital mammography.

Conflict of interest statement

None to disclose.

Acknowledgements

N. Houssami receives research support through a National Breast Cancer Foundation (NBCF, Australia) *Breast Cancer Research Leadership* Fellowship.

References

[1] 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.
[2] Houssami N, Macaskill P, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Breast screening using 2D-mammography or integrating digital breast tomosynthesis (3D-mammography) for single-reading or double-reading–evidence to guide future screening strategies. Eur J Canc 2014;50:1799–807.
[3] 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.
[4] 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.
[5] Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fanto 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.
[6] Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. J Am Med Assoc 2014;311:2499–507.
[7] McCarthy AM, Kontos D, Synnestvedt M, Tan KS, Heitjan DF, Schnall M, et al. Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. J Natl Cancer Inst 2014:106.
[8] Conant EF, Beaber EF, Sprague BL, Herschorn SD, Weaver DL, Onega T, et al. Breast cancer screening using tomosynthesis in combination with digital mammography compared to digital mammography alone: a cohort study within the PROSPR consortium. Breast Canc Res Treat 2016;156:109–16.
[9] Zuckerman SP, Conant EF, Keller BM, Maidment AD, Barufaldi B, Weinstein SP,

et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. Radiology 2016;281:730–6.
[10] Houssami N, Hunter K. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. Nature PJ Breast Cancer 2017;3:12.
[11] Bernardi D, Caumo F, Macaskill P, Ciatto S, Pellegrini M, Brunelli S, et al. Effect of integrating 3D-mammography (digital breast tomosynthesis) with 2D-mammography on radiologists' true-positive and false-positive detection in a population breast screening trial. Eur J Canc 2014;50:1232–8.
[12] Caumo F, Bernardi D, Ciatto S, Macaskill P, Pellegrini M, Brunelli S, et al. Incremental effect from integrating 3D-mammography (tomosynthesis) with 2D-mammography: increased breast cancer detection evident for screening centres in a population-based trial. Breast 2014;23:76–80.
[13] 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 2017;6(4):869–77.
[14] StatsDirect Ltd. StatsDirect statistical software (v3.0.167). England: StatsDirect Ltd; 2013. Available from: http://www.statsdirect.com.
[15] Domingo L, Hofvind S, Hubbard RA, Roman M, Benkeser D, Sala M, et al. Cross-national comparison of screening mammography accuracy measures in U.S., Norway, and Spain. Eur Radiol 2016;26:2520–8.
[16] Australian Institute of Health and Welfare. BreastScreen Australia monitoring report 2014-2015. Canberra: Australian Institute of Health and Welfare; 2017. Cancer Series no.6. Cat. no. CAN 105.