

Proportional incidence and radiological review of large (T2+) breast cancers as surrogate indicators of screening programme performance

S. Ciatto · D. Bernardi · M. Pellegrini · G. Borsato ·
P. Peterlongo · M. A. Gentilini · F. Caumo · A. Frigerio ·
N. Houssami

Received: 15 September 2011 / Revised: 5 November 2011 / Accepted: 16 November 2011 / Published online: 27 December 2011
© European Society of Radiology 2011

Abstract

Objectives Surrogate measures of screening performance [e.g. interval cancer (IC) proportional incidence] allow timely monitoring of sensitivity and quality. This study explored measures using large (T2+) breast cancers as potential indicators of screening performance.

Methods The proportional incidence of T2+ cancers (observed/expected cases) in a population-based screening programme (Trento, 2001–2009) was estimated. A parallel review of ‘negative’ preceding mammograms for screen-

detected T2+ and for all ICs, using ‘blinded’ independent readings and case-mixes (54 T2+, 50 ICs, 170 controls) was also performed.

Results T2+ cancers were observed in 168 screening participants: 48 at first screen, 67 at repeat screening and 53 ICs. The T2+ estimated proportional incidence was 68% (observed/expected=168/247), corresponding to an estimated 32% reduction in the rate of T2+ cancers in screening participants relative to that expected without screening. Majority review classified 27.8% (15/54) of T2+ and 28% (14/50) of ICs as screening error ($P=0.84$), with variable recall rates amongst radiologists (8.8–15.2%).

Conclusions T2+ review could be integrated as part of quality monitoring and potentially prove more feasible than IC review for some screening services.

Key Points

- Interval breast cancers, assumed as screening failures, are monitored to estimate screening performance
- Large (T2+) cancers at screening may also represent failed prior screening detection
- Analysis of T2+ lesions may be more feasible than assessing interval cancers
- Analysis of T2+ cancers is a potential further measure of screening performance

Keywords Breast carcinoma · Screening · Quality control · Interval cancer · Proportional incidence

S. Ciatto · D. Bernardi · M. Pellegrini · G. Borsato · P. Peterlongo
U.O. Senologia Clinica e Screening Mammografico,
Dipartimento di Radiodiagnostica, APSS,
Trento, Italy

M. A. Gentilini
Servizio Osservatorio Epidemiologico,
Direzione promozione ed educazione alla salute, APSS,
Trento, Italy

F. Caumo
Centro di Prevenzione Senologica,
P.O. Marzana, ULSS 20,
Verona, Italy

A. Frigerio
CRR, Centro di Riferimento
Regionale per lo Screening Mammografico,
Torino, Italy

N. Houssami
Screening and Test Evaluation Program, School of Public Health,
Sydney Medical School, University of Sydney,
Sydney, Australia

S. Ciatto (✉)
Corte Cà Brusà 1G,
37067 Valeggio sul Mincio, VR, Italy
e-mail: stefano.ciatto@gmail.com

strategy [3] and implemented on a national basis in many countries [4]. Individual population-based screening programmes routinely use surrogate indicators for timely evaluation of screening efficacy because mortality reduction requires very large studies with long follow-up periods and is methodologically difficult to assess in a service screening setting. Several surrogate measures are therefore used for ongoing monitoring in population breast screening practices, including screening coverage and attendance rates, cancer detection rates and, importantly, screening sensitivity as an indicator of potential screening impact. Proportional incidence of interval cancers (IC) [5] is considered a reliable method to assess screening sensitivity, being less affected by bias from lead time and overdiagnosis [6]. However it is not easy to determine the IC proportional incidence: challenges exist particularly in defining the underlying breast cancer incidence in the absence of screening, and in identifying all ICs in the absence of a regional cancer registry. Radiological review of screening mammograms preceding an IC, preferably using blinded review [7], is another indirect method to estimate sensitivity, based on the proportion of ICs deemed to be within the screening error upon radiological review [7]. Standards for IC proportional incidence and for radiological review outcomes are provided in European (EC) guidelines [3].

While it is increasingly evident that both tumour biology and pathology affect BC outcomes [8], BC size remains an established strong prognostic indicator [9], with larger tumours (T2+) being associated with poorer outcomes. Reducing the rate of breast cancers that are T2+ is a necessary condition for maximising ‘early detection’ and for realising the anticipated reduction in BC mortality from screening. Thus, we hypothesised that measuring a reduction of T2+ cancers from screening (relative to pre-screening data) may be a valuable indirect indicator of screening sensitivity and hence its potential efficacy.

The aims of this study were to estimate the proportional incidence of T2+ breast cancers in a population screening programme (with documented quality evaluation including IC proportional incidence [10]), and to perform parallel radiological review of both screen-detected T2+ cancers and ICs in consideration of whether these measures may be comparable as surrogate indicators of screening efficacy.

Methods

Screening programme characteristics

A population-based mammography screening programme was started in the Trento province (Italy) in October 2000, following a preliminary 3-month pilot study in the district of Trento, with further extension to the whole Trento province

territory. According to EC and Italian Ministry of Health recommendations, screenings were based on an active mail invitation of resident women aged 50–69 years. Screening tests were performed in seven different sites in the province territory, with centralised reading at the Trento screening unit using double-reading. The current study is based on the years 2001 to 2009: from 2005, digital mammography [direct (full-field digital mammography) at the Trento centre, indirect (computer radiography) at other screening sites] replaced film-screen mammography.

Screening programme performance indicators are regularly provided and are available online (National Screening Observatory <http://win.osservatorionazionale.screening.it>). Detailed screening performance indicators from the Trento programme have recently been reported [10] and indicate good coverage (approximately 85% from a population of 26,000 invited women), and demonstrate cancer detection rates and recall rates within recommended standards [10]. In addition, IC proportional incidence analysis and review showed good performance for the programme, fitting within EC standards [10].

Proportional incidence of T2+ breast cancers

Evaluation of the proportional incidence of T2+ cancers in screening attendees (either screen-detected at prevalent or incident rounds, or detected as interval cancers) considered all T2+ cancers observed during 2001–2009, with incidence data being complete up to December 2010. The estimate of underlying incidence of T2+ cancers was based on local cancer registry data for the years 1999 and 2000, the time-frame before screening was started. Screen-detected T2+ were identified through screening archives. Interval cancers (histologically confirmed invasive breast cancers occurring within 2 years of a negative screening episode) were identified by linking the screening archives with those from the local cancer registry (<http://www.registri-tumori.it/cms/?q=RTTrento>), the local pathology department and hospital discharge records (HDR). Screening tests performed during the limited pilot study in the year 2000 were not included in this study. Proportional incidence was calculated as the ratio of observed T2+ cancers in screening attendees to the number of T2+ cancers expected in the absence of screening. The expected T2+ cancers estimate was obtained by multiplying the screening attendees-years by age-specific T2+ breast cancer incidence for the Trento province, from the local cancer registry data for women aged 50–69 years and relative to 2001–2009.

Radiological review methods

Films of ‘negative’ screening mammograms preceding screen-detected T2+ cancer (subset of 54 cases for which

mammograms were available in screening archives) and all ICs (subset of 50 cases occurring during 2009–2010) were considered for the study. The review of films prior to the 2-yearly screen was performed for T2+ cancers detected at repeated screening rounds. Mammographic review was performed using methods that reduce bias in classification [7]: we used blinded ('masked') review methods and integrated screening negative controls (at least one subsequent negative screening) randomly drawn for screening archives and matched to ICs on a 3:1 ratio (total of 170 controls). Independent review was carried out by a panel of three dedicated breast radiologists (external to the screening programme) with varying screening experience (F.C. >10 years, A.F. >20 years, S.C. >30 years). Mammograms (screen-film or digital prints for the study set) were randomly admixed and mounted on rotating viewers: the reviewer was requested to mark on a paper scheme of the four mammography views, whether recall is warranted, and if so, the exact site of the mammographic abnormality for further assessment. Classification as "occult", "minimal sign" or "screening error" review categories [7] was based on the majority report from the 3 reviewers. The proportion of T2+ and ICs classified as 'screening error' (based on at least 2 radiologists recalling the exact abnormality) was determined. Observed differences were checked by the chi-square test, statistical significance being set at $P < 0.05$.

Results

T2+ proportional incidence

Underlying T2+ incidence for the years 1999 and 2000 (immediately preceding implementation of screening) in

the 50–69 age group was 0.00091. Screening attendees during 2001–2009 accounted for 271,385 women-years (89,219 and 182,167 women-years for first and repeat screens, respectively). Based on these data, the estimated number of T2+ cancers expected during 2001–2009 in this screened population was 247 cancers (year-specific data shown in Table 1).

During 2001–2009, there were 48 T2+ detected at first screen, and 67 T2+ detected at repeat screen, while 53 T2+ cancers were observed as IC—a total of 168 T2+ occurred in the screened cohort. The proportional incidence of T2+ cancers was therefore 68% (168 observed to 247 expected), corresponding to an estimated 32% reduction in T2+ rate in the screened population relative to expected incidence in the absence of screening.

The number of screens, and the number of observed and expected T2+ breast cancers, by year and screening round, are provided in Table 1

Radiological review of T2+ and ICs

Table 2 shows the results of the majority review from three radiologists. Fourteen of 50 (28%) ICs, and 15 of 54 (27.8%) T2+ cancers, were reviewed as screening errors (identified by at least two radiologists) ($P = 0.84$). The proportion of T2+ and ICs correctly identified by each radiologist at blinded review slightly differed: Radiologist A (T2+ 27.7% vs IC 36.1%, $P = 0.49$); Radiologist B (T2+ 29.6% vs IC 20.0%, $P = 0.36$); and Radiologist C (T2+ 35.1% vs IC 40.0%, $P = 0.76$). Recall rate (determined on negative controls) was different between radiologists (A=15.2%, B=8.8% and C=8.8%; $P = 0.08$) though not reaching significant levels. The proportion of cases with abnormalities indicated by at least two reviewers was significantly higher for T2+ ($P = 0.004$) or IC ($P = 0.005$), as compared to controls.

Table 1 Observed and expected numbers of T2+ cancers in the breast screening programme (2001–2009) for prevalent and incident screening

Prevalent screen	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
Number of attenders	9852	12787	3714	3133	3409	2963	2772	3931	4097	
T2+ detected	9	13	5	4	4	6	1	4	2	48
T2+ expected 1st year	8.96	11.64	3.38	2.85	3.10	2.70	2.52	3.58	3.73	42.46
T2 expected 2nd year		8.96	11.64	3.36	2.85	3.10	2.70	2.52	3.58	38.71
Repeat screens										
Number of attenders			7762	15651	14879	14176	14916	15493	16413	
T2+ detected			4	20	9	9	5	5	15	67
T2+ expected 1st year			7.06	14.24	13.54	12.90	13.57	14.10	14.94	90.35
T2 expected 2nd year				7.06	14.24	13.54	12.90	13.57	14.10	75.41
T2+ interval cancers		2	4	3	11	4	8	9	12	53
Total observed T2+	9	15	13	27	24	19	14	18	29	168
Total expected T2+	8.96	20.60	22.08	27.51	33.73	32.24	31.69	33.77	36.35	246.93

Table 2 Results of blinded radiological review of previous screening mammograms of interval cancers and T2+ cancers from three radiologists [(-)=negative; (+)=positive]

Review of interval cancers	Cancer (n=50)	Control (n=170)
(-)(-)(-)	25	130
(+)(-)(-)	11	27
(+)(+)(-)	6	12
(+)(+)(+)	8	1
Majority review ^a [(+)(+)(-) or (+)(+)(+)]	14 (28%)	–
Review of T2+ cancers	Cancer (n=54)	Control (n=170)
(-)(-)(-)	31	130
(+)(-)(-)	8	27
(+)(+)(-)	4	12
(+)(+)(+)	11	1
Majority review ^a [(+)(+)(-) or (+)(+)(+)]	15 (27.8%)	–

^a Majority review: at least 2 radiologists recalling exact abnormality as cancer on subsequent screen

Discussion

Surrogate measures of screening performance allow the timely monitoring of screening quality and increases the benefits from screening, and have been used in screening programmes throughout the world. Our study explored whether estimates of the proportional incidence of large (T2+) breast cancers, and radiological reviews of preceding mammograms for T2+ cancers in screening participants (measures not usually adopted when monitoring screening), may provide practical surrogate measures of screening sensitivity and potential effect.

We derived the numbers of T2+ cancers expected in the absence of screening from age-specific data from the local cancer registry for the period immediately preceding the screening programme implementation—this allowed estimation of the cancer incidence in the absence of screening because it excluded major long-term incidence temporal shifts and took into account the presumed spontaneous access to mammography available immediately before initiation of the screening programme. Nonetheless, our estimates required an assumption of a relatively stable underlying risk of T2+ tumours and assumed minimal effect from migration for the region. We estimated a proportional incidence of T2+ cancers of 68% in the screened cohort, suggesting that screening may have reduced up to one-third of the T2+ cancers that would have occurred in the absence of screening. This appears to be a plausible estimate given expected screening effects from randomised trials [1], however currently there are no standards for the proportional incidence of large cancers, so we cannot compare our

estimates with other studies or screening performance standards. It is unlikely that current standards for IC proportional incidence (EC standard <30% first year, <50% second year and 40% for biennial screening [3]) translate as a standard for T2+ proportional incidence, so we strongly encourage other investigators to extend the proposed concept of T2+ proportional incidence through screening research.

Radiological review of the screening mammogram preceding an IC is not a new concept [7], but our study is the first to investigate preceding mammograms for both ICs and screen-detected T2+ cancers occurring in the same time-frame, with consistent review methods performed by the same radiologists. A strength of the review strategy used in this work is the blinded classification and the case-mix (cancers and controls) which are recommended methods for radiological reviews to help reduce the bias in the classification of mammograms [7, 11] and to approximate the scenario of screening practice [1, 2]. Nonetheless, in a research context, reviewers might be expected to achieve higher sensitivity and lower specificity by adopting a lower threshold for suspicion, as suggested by the recall rates (see Table 2) which were higher than in routine practice, but comparable to other studies using radiological reviews [11–13]. The applied comparison of reviews of T2+ and ICs is reliable (performed by the same radiologist panel in the same setting using a standardised set of negative controls) and was performed to interrogate the extent to which preceding mammograms for these cancers would be classified as screening errors. We found that the majority of radiological reviews classified approximately 28% of these cancers as screening errors [3, 7]. Based on recently published data on proportional IC incidence and review, the performance of the Trento screening programme has been shown to be good within the limits recommended by the EC [10]. So, although the 28% screening error for IC review is slightly high, the more relevant issue here is that the distribution of radiological review results was similar for IC and T2+ cancers (see Table 1), raising the possibility that review of T2+ cancers may be integrated into a review of ICs as a complementary (and potentially equivalent) surrogate measure of screening sensitivity.

The results from the radiological review of T2+ cancers in this study suggest that this may have potential utility for monitoring screening performances, however, additional research is needed to further examine this, and to define whether standards for the proportion of ‘acceptable’ screening error [3] for radiological review of ICs could be adopted for review of T2+ cancers. Because of the logistical challenges inherent in identification and verification of ICs, our study has taken a pragmatic approach by exploring the review of T2+ cancers, as this may be more feasible to perform in routine screening clinical quality assurance, particularly where services encounter difficulties in identifying

IC cases. Furthermore, there is an inherent educational value in conducting reviews of T2+ cancers and potentially characterising the features of ‘missed’ mammographic lesions that are subsequently detected as large cancers.

Using the combination of epidemiological surveillance (proportional incidence) and radiological review of T2+ cancers, as shown in this study, potentially provides practical indicators of screening performance and warrants further evaluation by breast screening programmes. In particular, radiological review of T2+ cancers could be integrated with review of ICs (as part of quality monitoring) and may potentially prove more feasible than (and hence alternate to) review of ICs for some screening services.

References

1. Glasziou P, Houssami N (2011) The evidence base for breast cancer screening. *Preventive Medicine* doi:10.1016/j.ypmed.2011.05.011
2. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL (1995) Efficacy of screening mammography. A meta-analysis. *J Am Med Ass* 273:149–154
3. Perry NM, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L (2006) European guidelines for quality assurance in breast cancer screening and diagnosis. European Commission. Fourth Edition, Luxembourg
4. Osservatorio Nazionale Screening (2010) The National Centre for Screening Monitoring Eight Report. *Epidemiol Prev* 34:1–88 (<http://win.osservatorionazionale screening.it/eng-index.php>)
5. Paci E, Ciatto S, Buiatti E, Cecchini S, Palli D, Rosselli Del Turco M (1990) Early indicators of efficacy of breast screening programmes. Results of the Florence District Programme. *Int J Cancer* 46:198–202
6. Morrison AS (1982) The effects of early treatment. lead time and length bias on the mortality experienced by cases detected by screening. *Int J Epidemiol* 11:261–267
7. Houssami N, Irwig L, Ciatto S (2006) Radiological surveillance of interval breast cancers in screening programmes. *Lancet Oncology* 7:259–265
8. Comen EA, Norton L, Massagué J (2011) Breast cancer tumor size, nodal status, and prognosis: Biology trumps anatomy. *J Clin Oncol* 9:2610–2612
9. Carter CL, Allen C, Henson DE (1989) Relation of Tumor Size, Lymph Node Status and Survival in 24,740 Breast Cancer Cases. *Cancer* 63:181–187
10. Pellegrini M, Bernardi D, Di Michele S, Tuttobene P, Fantò C, Valentini M, Peterlongo P, Caumo F, Frigerio A, Ciatto S (2011) Analysis of proportional incidence and review of interval cancer cases observed within the mammography screening programme in Trento province, Italy. *Radiol Med*. [Epub ahead of print] PMID 21744249 2011.07.09
11. Ciatto S, Catarzi S, Lamberini MP, Risso G, Saguatti G, Abbattista T, Martinelli F, Houssami N (2007) Interval breast cancers in screening: The effect of mammography review method on classification. *Breast* 16:646–652
12. Vettorazzi M, Stocco C, Chirico A, Recanatini S, Saccon S, Mariotto R, Cinquetti S, Moretto T, Sartori P, Stomeo A, Ciatto S (2006) Quality control of mammography screening in the Veneto Region. Evaluation of four programmes at a Local Health Unit level—Analysis of the frequency and diagnostic pattern of interval cancers. *Tumori* 92:1–5
13. Caumo F, Vecchiato F, Pellegrini M, Vettorazzi M, Ciatto S, Montemezzi S (2009) Analysis of interval cancers observed in an Italian mammography screening programme (2000–2006). *Radiol Med* 114:907–914