Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: Interval breast cancers at 1 year follow-up

Vittorio Corsetti a, Nehmat Houssami b, Marco Ghirardi a, Aurora Ferrari a, Michela Speziani a, Sergio Bellarosa a, Giuseppe Remida a, Cristina Gasparotti a, Enzo Galligioni a, Stefano Ciatto c,*

a Ambulatori Raphael, Calcinato (BS), Italy
b Screening and Test Evaluation Program, School of Public Health, Sydney Medical School, University of Sydney, Sydney, Australia
c Mammography Screening Health Unit 16, Padua, Italy

A R T I C L E   I N F O
Article history:
Received 10 September 2010
Received in revised form 20 November 2010
Accepted 3 December 2010
Available online 4 January 2011

Keywords:
Breast cancer
Breast density
Cancer screening
Ultrasound
Interval cancers
Sensitivity

A B S T R A C T
Surveillance of interval cancers (IC) provides a measure of breast screening efficacy. Increased breast density is a predictor of breast cancer risk and of the risk of IC in screening. Improving screening sensitivity in women with dense breasts, through adjunct ultrasound (US), may potentially reduce IC; however this has not been proven. We report on first-year IC in a retrospective cohort of 8865 women who had 19,728 screening examinations (2001–2006): women with non-dense (D1–D2) breasts received mammography (M) screening, and women with dense (D3–D4) breasts also received ultrasound. Data linkage with both hospital discharge records and cancer registry databases was used to identify IC.

Underlying cancer rates (cancers observed within 1-year from screening) were 6.3/1000 screens in the D1–D2 group and 8.3/1000 screens in the D3–D4 group. Cancer detection rate (CDR) was 5.98/1000 in all screening examinations; in D3–D4 breasts ultrasound had an additional CDR of 4.4/1000 screens. There were 21 first-year IC in the D1–D2 group and 21 in the D3–D4 group. Cancer detection rate (ICR) of 1.07/1000 screens in women <50 years and 1.16/1000 screens in women ≥50 years. ICR by breast density were 1.0/1000 negative screens in D1–D2, and 1.1/1000 negative screens in D3–D4. Interval cancers were early stage (in situ or small invasive) cancers, almost all were node-negative. Screening sensitivity was 83.5% for mammography alone in D1–D2 breasts relative to 86.7% for mammography with ultrasound in D3–D4 breasts.

Our study shows that including ultrasound as adjunct screening in women with D3–D4 breasts brings the IC rate to similar levels as IC in non-dense breasts – this suggests that additional cancer detection by ultrasound is likely to improve screening benefit in dense breasts, and supports the implementation of a randomised trial of adjunct ultrasound in women with increased breast tissue density.

© 2010 Elsevier Ltd. All rights reserved.

* Corresponding author: Address: Ambulatori Oncologici Raphael, Via Vittorio Emanuele II, 25011 Calcinato (BS), Italy. Tel.: +39 030 9969662; fax: +39 030 9982088.
E-mail addresses: v.corsetti@raphael.coop (V. Corsetti), stefano.ciatto@gmail.com (S. Ciatto).
0959-8049/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
1. Introduction

Relentless debate on the potential benefit of mammography (M) screening does not negate evidence that it is the only population health intervention shown to reduce breast cancer mortality in randomised controlled trials (RCTs).1 Interval breast cancers are monitored by screening services as a quality indicator of screening programmes and as a surrogate measure of screening efficacy.2,3 Breast tissue density is a predictor of an increased risk of breast cancer4-8 and of an increased risk of interval cancers (IC) in screening.9 High breast tissue density is one of the most consistently validated risk factors for breast cancer, and is a stronger predictor of breast cancer risk than many other risk factors (with the exception of age and specific gene mutations). The association between high breast tissue density and breast cancer risk is evident in both younger and older women.4 Thus, at a population level, improving screening sensitivity in women with dense breasts may improve screening benefit and may potentially reduce interval cancers in this group of women. A number of studies have explored and have demonstrated the detection capability of adjunct ultrasound (US) in breast screening10-12; however data on how this affects interval cancer rates are very limited in this context.

We previously reported on the incremental cancer detection and cost of adjunct breast ultrasound in screening a cohort of Italian women with dense breasts on mammography.13 In this study, we focus on evaluating outcomes in this cohort at 1 year following screening and report on interval cancers (IC): we specifically aimed to calculate first year interval cancer rates are very limited in this context.

2. Methods

This is a study of a cohort of 8865 women who had 19,728 screening examinations at the study centre from 2001 to 2006. The study centre is a charity-funded breast service for women attending for screening (including those requesting screening due to perceived risk) or for symptom evaluation13 – symptomatic women were not eligible for inclusion in the present analysis. In this self-referring population of women, annual or biennial screening was provided. Detailed methodology and description of screening processes at the study centre have been previously reported13 and will only be described briefly; however, we point out that the present study is based on a shorter timeframe than our initial report (to allow ascertainment of IC); therefore the number of subjects, screens (and hence detection data) differ from our previous publication.

The majority of screening examinations included in this study were incident (repeat) screens, and approximately 40% were performed within 11–14 months of the prior screening exam. Of 19,728 screens, 3940 (20%) were prevalent or first screens (2311 screens in women <50 years, 1629 screens in those 50 years and older), and 15,788 (80%) were incident (repeat) screens (6117 screens in women <50 years, 9671 screens in those 50 years and older).

For the timeframe of the study, film-screen mammograms were interpreted by one of six radiologists, and density was classified according to the proportion of breast volume occupied by fibroglandular density based on BI-RADS categories (D1 = 0–25%, D2 = 26–50%, D3 = 51–75% and D4 = 76–100%).14 All subjects with negative screening mammograms and with BI-RADS D3 and D4 mammography had bilateral breast ultrasound, using an Aloka ProSound SSD-5500 unit and a multifrequency linear probe operated at 7.5–10 MHz.13 Breast imaging was reported based on an established classification system used in Europe and Australia (1 – normal, 2 – benign, 3 – indeterminate, 4 – suspicious and 5 – malignant): US findings classified as 3-5 were evaluated using percutaneous needle biopsy (needle cytology or core biopsy) and surgical biopsy where indicated. Data on additional investigations caused by ultrasound screening as applied in the study protocol and by additional surgical biopsy due to ultrasound were estimated as 4.9% and 0.9% respectively, in our earlier report13, these data are updated for the screening examinations considered in this evaluation.

The mammograms of ‘mammography-negative’ and ultrasound-only detected cancers were reviewed (28 of the 32 screening examinations in this group were available for this process) by an independent expert radiologist (SC) with over 30 years experience in screen-reading, using methods that reduce bias in radiologic review of missed cancers15 including masking and case-mix. This approach was integrated into our study design for the sole purpose of providing insight into the quality (and hence the generalisability) of our findings on screening mammography in women with D3–D4 breasts.

The study centre, in Brescia, is served by several surrounding Local Health Units (LHUs: Brescia, Bergamo, Mantova (Lombardia region), and Verona (Veneto Region)). We ascertained IC through the study centre archives and data linkage with hospital discharge records databases provided by the above listed LHUs, allowing for identifying cases that may have received treatment outside the LHU. The probability that IC were not identified through this approach is estimated as <5%.  

2.1 Analysis

Cancer detection rates (CDR), interval cancers rates (ICR), and underlying cancer rates were calculated as a proportion of screening examinations (or screens) in all subjects, for the group screened with mammography alone (D1–D2), and for the dense breasts group screened with mammography and ultrasound (D3–D4). If screening is negative each woman enters observation at time 0 (mammogram) and observation is censored at day 366. If the woman has another screen then a new screening examination is considered with 0 time at mammography and follow-up to identify interval cancers occurring within 365 days. Screening sensitivity (number of cancers detected with screening/number of cancers detected with screening + number of interval cancers at 365 days) was also calculated in each group.

3. Results

There were 21 interval cancers at 1 year follow-up, an overall ICR of 1.07/1000 negative screening examinations in 8865
women who had undergone 19,728 screens: 0.95/1000 negative screens in women younger than 50 years and 1.16/1000 screens in women 50 years and older. CDR and ICR by screening strategy, density category and age-group are summarised in Table 1. ICR by breast tissue density were 1.0/1000 negative screening examinations in D1–D2, and 1.1/1000 negative screening examinations in D3–D4 breasts. Screening sensitivity was 83.5% for mammography alone in the D1–D2 group, and 86.7% for mammography with ultrasound in the D3–D4 group (Table 1). Overall, CDR for (prevalent or incident) screening in this study was 5.98/1000 screening examinations. Underlying cancer rates are also summarised in Table 1.

Descriptive data for stage, pathologic tumour size and node status are presented for screen-detected and interval cancers in Appendix A. IC had similar stage distributions in the D1–D2 and the D3–D4 groups: all were pT1 or in situ cancers, and almost all were node negative.

Radiologic review of mammography films for the ‘mammography-negative’ and ultrasound-only detected cancers (see Section 2 on independent review) indicated that 5 of 28 were false negatives and had been missed by the interpreting radiologist: two were evident as suspicious lesions, one was a well-defined density that had been misdiagnosed as a lymph node and two were ‘minimal signs’ having only slightly atypical features – so overall about 18% were false negatives on screening mammography.

Additional testing (mostly fine needle biopsy) or surgery due to false positive ultrasound in women with dense breasts occurred in 5.5% (395/7224) of screening examinations – this included 61 surgical biopsies (0.84% of screens) with benign outcomes. Surgical biopsy due to false positive ultrasound was 1% of screening examinations in women <50 years and 0.6% of screening examinations in women 50 years and older.

### 4. Discussion

One of the challenging aspects of contemporary breast screening is to determine whether the use of adjunct imaging increases screening sensitivity relative to mammography alone, and whether that increased sensitivity in fact leads to a reduction in interval cancers (IC), and hence likely to lead to improved outcomes. Assumptions are usually made that additional cancers detected by adjunct screening (such as those detected on ultrasound-only in our study) would have emerged as IC had the screening been based on mammography alone. Such an assumption may be incorrect – it is just as reasonable to consider that some of the additional cancers detected with adjunct screening might not emerge as IC, as it is to consider that some will progress and be identified as IC: ideally this should be tested through a randomised controlled trial (RCT). To date, there have been no reports of RCTs of adjunct imaging in breast screening, although two studies of adjunct ultrasound have been initiated in Japan16 and in Italy.17

Our observational study design allows us to interpret the ICR in the higher-density group screened with mammography and ultrasound relative to that of the lower-density group screened with mammography only. In the former D3–D4 group an approximate fourfold rate of IC would be expected at 1 year follow-up relative to D1–D2 subjects, based on data for screening with mammography only, estimated from Kerlikowske.18 While we cannot claim that all the cancers detected with ultrasound-only screening of D3–D4 subjects

---

**Table 1 – Outcomes of screening mammography (M) in all subjects, and adjunct ultrasound (US) in women with dense (D3–D4) breasts, according to age-group and density categories (19,728 screening examinations).**

<table>
<thead>
<tr>
<th>Breast density</th>
<th>Age &lt; 50 years</th>
<th>Age ≥ 50 years</th>
<th>All ages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BIRADS D1–D2</strong>&lt;br&gt;(M only)&lt;br&gt;[5509 women]</td>
<td>Total screening examinations [screens&lt;sup&gt;a&lt;/sup&gt;]</td>
<td>4434</td>
<td>8070</td>
</tr>
<tr>
<td></td>
<td>Cancers detected by M at screening (CDR per 1000 screens&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>12 (2.7)</td>
<td>54 (6.7)</td>
</tr>
<tr>
<td></td>
<td>First year interval cancers (IC)&lt;br&gt;[number/number negative screens]</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Screening sensitivity&lt;sup&gt;b&lt;/sup&gt; (M only)</td>
<td>85.7%</td>
<td>83.1%</td>
</tr>
<tr>
<td></td>
<td>Cancer rates&lt;sup&gt;c&lt;/sup&gt; per 1000 screens (number of cancers/number of screens)</td>
<td>3.2</td>
<td>8.1</td>
</tr>
<tr>
<td><strong>BIRADS D3–D4</strong>&lt;br&gt;(M and US)&lt;br&gt;[3356 women]</td>
<td>Total screening examinations [screens&lt;sup&gt;a&lt;/sup&gt;]</td>
<td>3994</td>
<td>3230</td>
</tr>
<tr>
<td></td>
<td>Cancers detected by M at screening (rate per 1000 screens&lt;sup&gt;d&lt;/sup&gt;)</td>
<td>11 (2.7)</td>
<td>9 (2.8)</td>
</tr>
<tr>
<td></td>
<td>Additional cancers detected only by adjunct US&lt;br&gt;(rate per 1000 screens&lt;sup&gt;d&lt;/sup&gt;)</td>
<td>14 (3.5)</td>
<td>18 (5.6)</td>
</tr>
<tr>
<td></td>
<td>First year interval cancers (IC)&lt;br&gt;[number/number negative screens]</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Screening sensitivity&lt;sup&gt;b&lt;/sup&gt; (M + US)</td>
<td>80.6%</td>
<td>93.1%</td>
</tr>
<tr>
<td></td>
<td>Cancer rates&lt;sup&gt;c&lt;/sup&gt; per 1000 screens (number of cancers/number of screens)</td>
<td>7.8</td>
<td>9.0</td>
</tr>
</tbody>
</table>

---

<sup>a</sup> Data are based on 19,728 screening examinations or ‘screens’ in 8865 women: a screening examination is a mammogram (M) only in the D1–D2 group, and a mammogram (M) followed by adjunct ultrasound (US) in the D3–D4 group; there are more screening examinations than women since this cohort was screened for a maximum of 6 years, and includes prevalent and incident screening examinations (described in Section 2). If screening is negative each woman enters observation at time 0 (mammogram) and observation is censored at day 366 (see ‘Section 2.1’ for further details).

<sup>b</sup> Screening sensitivity = cancers detected at screening/cancers detected at screening plus interval cancers.

<sup>c</sup> Cancer rates are the underlying cancer rates per 1000 screens based on all cancers observed at 1 year from screening (cancers detected at screening plus interval cancers).

---
would have evolved as IC, it is clear from our data that the use of adjunct ultrasound in D3–D4 screens has essentially resulted in similar IC rates in this group of subjects as their D1–D2 counterparts (counting all ages). Careful examination of the data pattern shown in Table 1 indicates that this was largely achieved through an effect of screening women 50 years and older with dense breasts in whom the inclusion of ultrasound in the screening strategy was associated with additional detection of 5.6 cancers per 1000 screens, and this appears to be followed by a markedly diminished first-year ICR. However, the apparent effect of adjunct ultrasound was different in women aged 49 years and younger: although ultrasound was associated with additional detection of 3.5 cancers per 1000 screens, there was a threefold rate of IC relative to their less dense (D1–D2) counterparts (approximately 4.5 times the ICR would have been expected in the D3–D4 group in this age-group relative to the D1–D2 group). So, although adjunct ultrasound seems to be associated with some decrease in the ICR this has not been as large an effect in the 49 years and younger age-group as that observed in those 50 years and older – this observation should be interpreted with caution since we had relatively modest numbers of cancers and screens in the 49 years and younger age-group.

We were unable to estimate the expected number of cancers in our cohort, to do so would require making several assumptions about underlying risk in this self-selected cohort of women – therefore we did not report proportionate ICR. Instead, we chose to simply report relative screening sensitivity and ICR (for each screened group) to avoid over-interpretation of these observational data. We have also highlighted the underlying cancer rates in our subjects (Table 1) based on all cancers observed at 1 year follow-up from screening. It is possible that a small number of IC may not have been identified using the data linkage process we described in Section 2 – however even if this has occurred we would expect the distribution to be random across density and age categories. So this is unlikely to have a substantial effect on the pattern of IC rates discussed above. It could be argued that mammography-only screening in our D3–D4 group had modest cancer detection rates, raising the possibility that mammography quality may not have been optimal (and hence this may have led to high additional detection rates with ultrasound). Review of mammography films for the ‘mammography-negative’ and ultrasound-only detected cancers (see Section 2 on independent review) indicated that even an expert radiologist considered only five cases (about 18%) to be false negatives (screening error) for mammography, and the distribution of these cases was similar to what might be expected in a series of missed cancers in mammography screening.\(^\text{15}\) The expert review of ‘mammography-negative’ films was performed by a radiologist who was not involved in screen-reading at the study centre and was based on a mixed case-set to minimise bias in classification\(^\text{15}\); however the reviewer was not blinded to the study question – nonetheless the findings of the mammography review process suggest that the quality of screening mammography in the study centre was adequate. This is also supported by the detection data shown in Table 1 for the non-dense (D1–D2) cohort.

In this study, underlying cancer rates (counting all cancers at 1 year follow-up from screening) were higher in association with the screening examinations in the D3–D4 group than those in the D1–D2 group (8.3/1000 screens relative to 6.3/1000 screens) as would be expected for breast density as a risk factor. Of note, differences in underlying cancer rates associated with tissue density were evident to a lesser extent in women aged 50 years and older relative to the younger group (Table 1) possibly due to the duration of follow-up (1 year) in which to ascertain interval cancers; this could also be due to the fewer numbers of D3–D4 screens in older women (Table 1).

Several studies, reviewed in two recent publications\(^\text{11,12}\), have reported on the accuracy of ultrasound as adjunct screening in women with dense breasts. However, to date, only two studies have provided information on subsequent interval cancers in women with dense breasts screened with ultrasound – the work reported in this paper and the earlier work of Berg et al.\(^\text{19}\) The study from Berg et al. was based on both increased risk and tissue density (comprising D2, D3 and D4) as entry criteria into the trial, and all subjects were screened with both M and US. Thus the two different designs and study populations cannot be directly compared (and comparison of interval cancer data across countries is not necessarily valid\(^\text{3}\)); however the two studies are discussed here as the only sources of evidence on this issue: overall, both studies indicate that the inclusion of ultrasound in a breast screening strategy for women with dense breasts appears to be associated with (a) substantial additional ultrasound-only detection in dense breasts (Berg’s trial had an additional detection yield of 4.2 per 1000 screened women) and (b) a modest number of IC at first-year follow-up (at the least, fewer IC than expected with conventional mammography-only screening in dense breasts). Berg’s trial reported 8 first-year interval cancers from 2637 screened D2–D4 women.\(^\text{19}\) Our work has indicated that 8 first-year interval cancers occurred following 7224 screens in 3356 D3–D4 women. While we do not advocate that ultrasound should be used to screen all women with dense breast tissue, we take the view that both these studies (ours and that of Berg et al.)\(^\text{15}\) provide evidence to support consideration of an RCT of adjunct breast ultrasound in women with increased breast tissue density.

Since double-reading is associated with improved cancer detection in mammography screening\(^\text{16}\), it is possible that the use of single-reading of mammograms in our study has slightly over-estimated the incremental contribution or effect of adjunct ultrasound. While this potential limitation is acknowledged, it is unclear to what extent double-reading improves perception errors specifically in dense breasts, and much of the evidence on the risk of interval cancers in dense breasts comes from screening programmes using double-reading. Another potential limitation of this work is that the ultrasound technology used (7.5–10 MHz) was of lower frequency than the current state-of-the-art ultrasound units (frequently operating at higher frequency), so this may have slightly reduced ultrasound sensitivity, and possibly underestimated the contribution of adjunct ultrasound.

We conclude that our data are consistent with an effect of a reduction of IC in women with dense (D3–D4) breasts screened with ultrasound as adjunct to mammography screening, and the overall pattern of our data suggests that adjunct ultrasound in women with dense breasts may potentially have the most benefit in those aged 50 years and...
older. These findings may justify evaluation in a randomised study and with a larger number of subjects. We note one RCT of adjunct breast ultrasound, under progress in Japan,16,17 restricted to women aged 40–49 years: there may be more value in implementing a screening trial in women in the highest breast tissue density (D3–D4) who are 50 years and older, as in the ongoing Italian RCT,17 which is implemented within the framework of the national screening programme, instead of recruiting younger women only or a broader range of breast density categories.

Conflict of interest statement

None declared.

Acknowledgements

Nehmat Houssami is supported by National Health and Medical Research Council (NHMRC) programme Grant 402764 to the Screening and Test Evaluation Program.

The authors are grateful to Local Health authorities for providing Hospital Discharge database for linkage: Dr. G. Imbalzano (Bergamo); Dr. F. Vassallo (Brescia); Dr. M. Galavotti, Dr. E. Anghinoni (Mantova); Dr. M. Vettorazzi (Veneto Region) and Dr. Pierluigi Chiodera, Istituto Clinico S.Anna (Brescia) for pathological data retrieval.

Appendix A

See Table A.1

<table>
<thead>
<tr>
<th>Stage variable</th>
<th>M-detected in non-dense breasts (D1–D2)</th>
<th>Interval cancers (1-year) in non-dense breasts (D1–D2)</th>
<th>M-detected in dense breasts (D3–D4)</th>
<th>US-only detected in dense breasts (D3–D4)</th>
<th>Interval cancers (1-year) in dense breasts (D3–D4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumour size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTis</td>
<td>9</td>
<td>–</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>pT1mic</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pT1a</td>
<td>2</td>
<td>1</td>
<td>–</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>pT1b</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>pT1c</td>
<td>33</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>pT1</td>
<td>2</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pT2</td>
<td>7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pT3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pT4</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pTx (unknown)</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>Node status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pN0</td>
<td>45</td>
<td>11</td>
<td>15</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>pN1</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>–</td>
</tr>
<tr>
<td>pNx (unknown)</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>All cancers</td>
<td>66</td>
<td>13</td>
<td>20</td>
<td>32</td>
<td>8</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 0</td>
<td>9</td>
<td>–</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Stage I</td>
<td>31</td>
<td>12</td>
<td>11</td>
<td>24</td>
<td>6</td>
</tr>
<tr>
<td>Stage II+</td>
<td>23</td>
<td>–</td>
<td>1</td>
<td>3</td>
<td>–</td>
</tr>
<tr>
<td>Stage unknown</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>–</td>
</tr>
</tbody>
</table>

REFERENCES


