

RESEARCH ARTICLE

Performance Indices of Needle Biopsy Procedures for the Assessment of Screen Detected Abnormalities in Services Accredited by BreastScreen Australia

Gelareh Farshid^{1*}, Thomas Sullivan², Simeon Jones³, David Roder³

Abstract

Background: We wished to analyse patterns of use of needle biopsy procedures by BreastScreen Australia (BSA) accredited programs to identify areas for improvement. **Design:** BSA services provided anonymous data regarding percutaneous needle biopsy of screen detected lesions assessed between 2005-2009. **Results:** 12 services, from 5 of 7 Australian states and territories provided data for 18212 lesions biopsied. Preoperative diagnosis rates were 96.84% for lesion other than microcalcification (LOTM) and 93.21% for microcalcifications. At surgery 97.9% impalpable lesions were removed at the first procedure. Of 11548 Microcalcification (LOTM) biopsied, 46.9% were malignant. The final diagnosis was reached by conventional core biopsy (CCB) in 72.46%, FNAB in 21.33%, VACB in 1.69% and open biopsy in 4.52% of lesions. FNA is being limited to LOTM with benign imaging After FNAB, core biopsy was required for 38% of LOTM. In LOTM the mean false positive rate (FPR) was 0.36% for FNAB, 0.06% for NCB and 0% for VACB. Diagnostic accuracy was 72.75% for FNAB and 92.1% for core biopsies combined. Of 6441 microcalcifications biopsied 2305 (35.8%) were malignant. Microcalcifications are being assessed primarily by NCB but 6.57% underwent FNAB, 45.6% of which required NCB. False positive diagnoses were rare. FNR was 5% for NCB and 1.53% for VACB. Diagnostic accuracy was 73.52% for FNAB, 86.29% for NCB and 88.63% for VACB. Only 8 of 12 services had access to VACB facilities. **Conclusions:** BSA services are selecting lesions effectively for biopsy and are achieving high preoperative diagnosis rates. Gaps in the present accreditation standards require further consideration.

Keywords: Breast cancer - screening - mammography - needle biopsy

Asian Pac J Cancer Prev, 15 (24), 10665-10673

Introduction

Breast cancer remains one of the most common malignancies in women. Even in developing countries, its incidence is on the rise (Sepandi et al., 2014). Assessment is part of the remit of mammographic screening programs and requires care to minimize some of the harms associated with screening. In organized population based screening programs such as BreastScreen Australia, a broad range of performance standards are stipulated to monitor services' achievements in relation to assessment (BreastScreenAustralia, 2005). These include standards relating to the use of percutaneous needle biopsies, preoperative diagnosis rates and rates of benign surgical biopsy. Accreditation standards for BreastScreen Australia (BSA) are being updated at present. A recent analysis has found significant improvements in several performance indices over time, including benign open biopsy rates (Roder et al., 2014). In this study we aimed to analyse the patterns of use of needle biopsy procedures by BSA

accredited programs and identify areas for improvement.

This study was conducted with approval from the BSA National Quality Management Centre. The requirement for informed consent was waived.

Materials and Methods

Design of the breastScreen Australia program

BreastScreen Australia is a national program of free screening mammography for women in Australia, jointly funded by state and federal governments of Australia. The design of the BreastScreen Australia program has been described previously (Farshid and Rush, 2004). In brief, asymptomatic women aged 50-69 years are invited to participate at 1 or 2 yearly intervals, depending on family history. Two radiologists read two view screening mammograms independently. A third reader arbitrates discordant results. A 5-tier grading scheme is applied, grade 1, normal; grade 2, benign; grade 3, indeterminate/equivocal, grade 4, suspicious for malignancy and grade 5-

¹Discipline of Medicine, University of Adelaide and SA Pathology, Wayville, ²Discipline of Public Health, University of Adelaide, Adelaide, South Australia ³Cancer Australia, Level 14, Elizabeth, Surry Hills, New South Wales *For correspondence: gelareh.farshid@health.sa.gov.au

radiologically malignant (National Breast Cancer Centre, 2002). Women with grade 1 or 2 lesions are “cleared” to return for re-screening in 1-2 years. The remaining lesions are recalled for further assessment. At assessment further imaging is carried out, including work up mammography and if necessary ultrasound examination and clinical breast examination. Some lesions are cleared while biopsy is performed for those that remain abnormal with imaging grades 3 and above. Lesions requiring surgery are referred outside the Program. Programs are notified of interval cancers by data matching with their respective Cancer Registries.

Study design

Approval was obtained from the Breast Screen Australia National Quality Management Committee (NQMC) to undertake a collaborative benchmarking survey of the patterns of use of various percutaneous needle biopsy modalities used by BSA accredited assessment clinics across the country and to evaluate the relevant clinical outcomes. The invitation was extended to all BSA accredited jurisdictions to participate. There were no exclusion criteria for participation. Participation was on a voluntary basis at the state and territory level. For states with more than one service, the data were provided separately for each service. The survey pertains to the 5-year period 2005-2009 to optimise complete treatment data retrieval. The data are accrued prospectively by services as part of the ongoing mandatory reporting requirements for BSA accredited programs. In order to facilitate data retrieval, to the extent possible, the survey requested data items collected in line with the National Accreditation Standards (NAS) and supported by the BreastScreen Australia data dictionary. This enhances the internal consistency of the responses. In accordance with the existing NAS, the data are focused primarily on the target age group of women 50-69 years. Some data fields were not part of the NAS but were requested for this survey and could be retrieved from the service databases. Guidance was provided through documentation and support at a pilot site regarding the collection of these fields.

All data were de-identified with respect to state and service. No patient level data were collected.

The NQMC provided input in the design of the survey. Following a pilot at one site, the survey was administered to all participating sites. Support was available at the pilot

site to address any ambiguities in the survey questions for participants. The NQMC monitored the progress of the study through regular updates and consultation regarding the presentation of the results.

The raw data were aggregated in a combined database for the purposes of statistical analysis. Since microcalcifications are often not detectable by sonography, precluding some needle biopsy approaches, we have separated microcalcifications from other lesions in the analysis of outcomes. We present the results separately for microcalcifications and group all other lesions as lesions other than microcalcifications (LOTM).

In BSA services the false positive rate (FPR) is specified as the number of false positive diagnoses established by each needle biopsy modality, as the proportion of cases with final malignant outcome. In BSA services the false negative rate (FNR) is specified as the number of false negative diagnoses established by each needle biopsy modality, as the proportion of cases with final malignant outcome. Two methods were used to present the FPR and FNRs. One was based on the average of the rates reported from the participating services. The other was calculated by averaging the combined raw data from the services. Although these figures were generally close to each other, the first method is closer to real world practice, as it reflects the differing lesion selection criteria used by services.

In the following account, definitive results refer to the sum of clear cut benign or malignant diagnoses as a proportion of all biopsies performed for the lesion subtype (microcalcifications or LOTM). Other diagnostic categories, i.e. inadequate, atypical/equivocal and suspicious are thus excluded from the numerator of definitive results.

Results

For the five year period 2005-2009, five of seven Australian states and territories, collectively representing 12 services, participated in the survey. The remaining two states expressed support but were not in a position to participate due to local factors.

Service snapshot

During the 5-year period 2005-2009, screening mammograms were performed for a total of 1494809 women, of whom 67956 were recalled, representing

Table 1. Summary Overview

Lesion	Mean	Standard Deviation
Number of screening mammograms in the target age group (TAG)	24913.48	19945.89
Number of women in TAG assessed by the service	1132.6	659.26
Number of radiologists per year	11.17	9.89
Recall rate (%) - round 1 for TAG (NAS < 10%)	12.09	2.42
Recall rate (%) - round 2+ for TAG (NAS < 5%)	5.21	3.17
Pre-operative diagnosis rate for invasive cancer & DCIS (%) (NAS > 75%)	91.8	4.53
Benign open biopsy rate for women assessed - round 1 (%) (NAS < 4%)	2.35	0.83
Benign open biopsy rate for women assessed - round 2+ (%) (NAS < 3.2%)	1.89	0.64
Ratio of benign to malignant open biopsies (based on local excision result)	2.73	1.59
% Impalpable lesions correctly identified at first open biopsy	97.9	2.28

an overall recall rate of 4.55%. Of these, 6379 microcalcifications and 11833 LOTM were biopsied, amounting to 18212 lesions. This figure represents 1.22% of women screened and 26.80% of women recalled for assessment. A total of 20,724 biopsies were performed, as some lesions required more than 1 type of biopsy.

The mean number of women 50-69 screened by the services was 24913 (SD 199945.9, range 3685-66844). The mean number of women assessed per service was 1132.6 (range 216-1971). The mean number of radiologists employed in any year was 11 (range 1-38). Three of 12 services reported access to on site pathologists.

Five of 12 (41.7%) services had a prone table and eight services (66.7%) had an add-on device available. Overall, facilities for the performance of vacuum assisted core biopsies (VACB) were available at eight of 12 sites (66.7%).

Core imprint cytology was utilised by four services (33.3%); three reported discussing the core imprint findings with the clients.

Lesion localisation was undertaken using hook wires in six services (50%), carbon ink in one service and a combination of hook wire, carbon and other methods at one centre. One centre reported not using localisation techniques. No centre reported using intra-operative ultrasound for lesion localisation.

Of the 11 services responding to this question, 9 reported using site markers (81.8%) to assist with the localisation of stereotactic guided biopsies and 8 (72.7%) reported using site markers for ultrasound guided biopsies.

Overall outcomes

Table 1 presents some of the important data relating to recall and assessment of women undergoing screening mammography. The mean recall rate was 12.09% for women during the first round of screening and 5.21% for subsequent rounds. The pre-operative cancer diagnosis rate, inclusive of invasive cancer and DCIS, was 91.8%. In total, 2.35% of women recalled to assessment at the first screening round and 1.89% of those recalled at subsequent rounds, were referred to surgical biopsy to establish the nature of their screen detected lesion. The ratio of benign to malignant diagnoses after surgical biopsy was 2.73: 1.

Overall, in 97.9% of all surgery undertaken for impalpable lesions, the intended screen detected abnormality was identified during the first procedure. As shown in Table 2 the preoperative biopsy rate varied by lesion type, being higher for LOTM. Two rates are provided for preoperative biopsy rates, one based on the average of all raw data combined, the other is the mean of rates reported by services.

Lesions other than microcalcifications (LOTM)

Biopsy results for a total of 11,548 non-calcified lesions are reported, of which 5420 (46.9%) were found ultimately to be malignant. Table 3 summarises the performance indices for LOTM. Overall, 96.44% of malignant LOTM were diagnosed without the need for surgical biopsy. The mean preoperative biopsy rate reported by services was 96.84%.

A total of 13,953 biopsy procedures were carried out for these 11,548 lesions, 2405 lesions (17.2%) requiring assessment by more than one modality. Collectively of all biopsy procedures for LOTM, 5259 (37.7%) were FNAB, 8500 (60.9%) conventional NCB and only 194 (2.2%) VACB.

Services specified the modality that ultimately established the diagnosis for each lesion they assessed. Among LOTM, the final diagnosis was established by means of FNAB in 21.33% of cases, by NCB in 72.46%, by VACB in 1.69% and after diagnostic open biopsy in 4.52% of lesions.

The number of LOTM undergoing assessment by FNA varied from 3-426 per year. The proportion of histologically proven malignant lesions that were being assessed by FNAB ranged from 0 to 52.72% among the services, with a mean of 25.78%. This figure was lower than 50% for all but one service. This variation indicates that most services are preselecting lesions likely to be benign for assessment by FNAB.

Of the 5259 lesions undergoing FNAB, the mean value for the proportion of lesions requiring further assessment by core biopsy was of 38% (range 12.5 to 80.5%). Focusing on the 6 of these 12 programs performing more than 50 FNA procedures per year for LOTM, this range was (20.9% to 67.4%).

Table 2. Biopsy Outcomes by Lesion Type

Lesion	Non-calcified lesions	Microcalcifications
Mean number of lesions biopsied	11548	6441
Mean number found to be malignant	5420 (46.9%)	2305 (35.8%)
Malignancy diagnosed lesions without open biopsy	5227 (96.4%)	2100 (91.1%)
Mean reported preoperative diagnosis rate for malignant lesions	96.84%	93.21%

Table 3. Performance Indices for Lesions Other Than Microcalcifications

	FNA	NCB	VACB	Calculated for all cores
Number assessed	5259	8500	194	8694
Modality for final diagnosis	21.33%	72.46%	1.69%	74.15%
Rate definitive results	74.20%	93.10%	88.70%	93.01%
Mean reported false positive rate	0.36%	0.06%	0%	0.06%
Mean reported false negative rate	4.91%	1.98%	2.94%	1.71%
Mean reported inadequate rate	14.47%	0.81%	2.72%	0.83%
Accuracy	72.75%	92.24%	86.08%	92.10%

The mean value calculated for the proportion of lesions with definitive benign or malignant results was 74.2% for FNAB, 93.1% for NCB and 88.7% for VACB. The figure for VACB is less robust than the other two rates since it is based on only 194 biopsies compared to 5259 FNAs and 8500 conventional core biopsies.

Accuracy was calculated as a sum of cases with true benign and true malignant results as a proportion of all cases assessed by each modality. Considering the findings for LOTM, the accuracy was 72.75% for FNAB, 92.24% for conventional core biopsy and 86.08% for VA core biopsy. Combining all core biopsies, the accuracy was 92.1%.

False positive diagnoses of cancer for LOTM

For women 50-69 assessed for screen detected LOTM, there were 8 false positive cytologic diagnoses among the 2036 malignant lesions assessed by FNAB, 2 among the 4431 assessed by conventional core biopsy and none among the 83 that underwent VA core biopsy. Rather than calculating averages from amalgamated data, the service performance data will be presented. The mean of the false positive rates (FPR) reported by the services was 0.36% for FNAB, 0.06% for NCB and 0% for VACB. Given the small numbers of VACBs, the data for all cores were combined, giving a mean figure of 0.06%.

Among the 5 sites that consistently reported more than 10 malignant FNA results per year for LOTM, 4 had experienced at least one false positive FNA diagnosis during the survey period for women in the target age group. Considering LOTM alone, at the individual service level, during the survey period the FPR of FNA exceeded the NAS threshold of 1%, at least once in 3 of these 5 services.

Among the 4514 non-calcified malignant lesions assessed by core biopsies (conventional and VA combined) during this five year period, there were only 2 false positive core biopsy diagnoses reported among all sites

combined. Apart from these two isolated events, the FPR for core biopsies was 0% across all services at all other time periods.

In addition to the false positive rate of FNA, a proportion of non-calcified lesions with positive FNA results were found to be DCIS rather than invasive cancers on final histology. During the 5-year survey period, 7 of the 9 services that consistently performed more than 10 FNAs per year reported diagnosing malignancy in LOTM by FNAB that on final histology proved to be DCIS. The number of such lesions was between 1- 4 per service for women in the TAG.

False negative diagnoses for LOTM

In BSA services the false negative rate is specified as the number of false negative diagnoses established by each needle biopsy modality, as the proportion of cases with final malignant outcome. Using the figures reported by each service, the mean FNR was 4.91% for FNA, 1.98% for Non-VA core biopsy and 2.94% for VA core biopsy. Given the small numbers of VACBs, the data for all cores were combined, giving a mean FNR rate of 1.71% for all core biopsies.

Four the 10 services that perform more than 10 FNAs per year exceeded the NAS threshold of 6% for the FNR of FNA at least once in the 5 year period of the survey. The number of malignant diagnoses established by FNA was consistently low, suggesting lesions likely to be benign were being pre-selected for this test.

The mean FNR for core biopsies was 1.71%. There are currently no NAS for this performance measure. If a figure of 2.0% were used as a potential maximum FNR, 4 sites would have exceeded this figure in at least 3 of the 5 years of the survey. A common feature in 3 of these 4 sites was a high rate of use of FNAB, and specifically the higher proportions of malignant FNA results, which was in the region of 35%. The practice of assessing a high proportion of malignant cases by FNA reduces the

Table 4. Performance Indices for Microcalcifications

	FNA	NCB	VACB	Calculated for all cores
Number assessed	438	3757	2576	6333
Modality for final diagnosis	4.70%	48.90%	40.00%	88.90%
Rate definitive results	74.70%	87.40%	92.89%	89.57%
Mean reported false positive rate	0%	0.31%	0%	0.18%
Mean reported false negative rate	1.67%	5.00%	1.53%	2.46%
Mean reported inadequate rate	15.80%	1.43%	1.28%	1.99%
Accuracy	73.52%	86.29%	88.63%	87.24%

Table 5. Performance Indices All Lesions

	FNA		NCB		VACB		Calculated for all cores	
	Calcs	Other	Calcs	Other	Calcs	Other	Calcs	Other
Number assessed	438	5259	3757	8500	2576	194	6333	8694
Modality for final diagnosis	4.70%	21.33%	48.90%	72.46%	40.00%	1.69%	88.90%	74.15%
Rate definitive results	74.70%	74.20%	87.40%	93.10%	92.89%	88.70%	89.57%	93.01%
Mean reported false positive rate	0%	0.36%	0.31%	0.06%	0%	0%	0.18%	0.06%
Mean reported false negative rate	1.67%	4.91%	5.00%	1.98%	1.53%	2.94%	2.46%	1.71%
Mean reported inadequate rate	15.80%	14.47%	1.43%	0.81%	1.28%	2.72%	1.99%	0.83%
Accuracy	73.52%	72.75%	86.29%	92.24%	88.63%	86.08%	87.24%	92.10%

proportion of malignant lesions undergoing core biopsy, such that a small number of false negative results lead to relatively high FNRs. Frequent use of FNAB did not invariably lead to high FNRs of core biopsy, as two other sites with frequent FNA use had core biopsy FNRs of under 2%. The remaining site with higher than average core biopsy FNRs did not have a high rate of FNA use. We do not have data on numbers of cores retrieved per lesion and the level of teams' expertise to assess the significance of these variables on FNRs.

The mean reported rate of inadequate specimens for LOTM was 14.47% for FNAB, 0.81% for conventional core biopsy and 2.27% for VA core biopsy. In consideration of the small number of VA core biopsies performed for LOTM all core biopsy results were combined, producing overall inadequate rate of 0.83% for all core biopsies.

Time trends

For the 9 services that reported more than 10 FNA biopsies for LOTM per year, there was a decline in the number of FNA procedures over the 5 year time frame of the study. In these services compared to the numbers of FNA procedures performed in 2005, there was a contraction in the use of FNA in 2009 by 9.2 to 51.2%. No service had expanded its FNA use during this time frame.

Outcomes for microcalcifications

A total of 6441 microcalcifications were biopsied during this 5 year time frame, of which 2305 (35.8%) ultimately proved to be malignant. Of the 2305 malignancies, 2100 (91.1%) were diagnosed on the basis of needle biopsies. The mean preoperative biopsy rate for microcalcifications as reported by services was 93.21%. Table 4 summarises the performance indices of biopsies for microcalcifications. A total of 6771 biopsy procedures were performed for these 6441 microcalcifications, as some lesions required assessment by more than one modality. Collectively, of all biopsy procedures for microcalcifications 438 (6.57%) were FNAB, 3757 (56.32%) conventional NCB and 2476 (37.12%) VACB. Together core biopsies accounted for 93.43% of biopsies performed for microcalcifications. Services specified the biopsy modality that ultimately established the diagnosis of each lesion they assessed. Overall, among microcalcifications, the final diagnosis was established by means of FNAB in 4.7%, NCB in 48.9%, VACB in 40.0% and diagnostic open biopsy in 6.4%.

Only one services reported consistently performing more than 10 FNAs for microcalcifications per year during this 5-year time frame. On average 45.6% of calcified lesions assessed by FNA required further assessment by

core biopsy.

Time trends

Over the last 5 years, one service has increased its use of FNA for microcalcifications slightly from 15 to 23 lesions per year. The other two services have reduced their reliance on FNA from 82 to 6 lesions and from 16 to 7 lesions. In 2009, only one service had performed more than 10 FNAs for microcalcifications.

Definitive results

The proportion of microcalcifications with definitive benign or malignant results was 74.7% for FNAB, 87.4% for CNCB and 92.89% for VACB. This figure was 89.57% combining all core biopsies.

Accuracy of assessment of microcalcifications

Accuracy was calculated as a sum of cases with true benign and true malignant results as a proportion of all cases assessed by each modality. Overall, 30.75% of biopsies for microcalcifications constituted true positive malignancies, 55.60% true negative benign lesions and the remainder false positives, false negatives and non-representative/inadequate, atypical and suspicious results. Table 4 shows an accuracy of 73.52% for FNAB, 86.29% for NCB and 88.63% for VACB when assessing microcalcifications. Considering all core biopsies together, the composite figure for accuracy of core biopsies in assessing microcalcifications is 87.24%.

False positive diagnoses of cancer for microcalcifications

For women 50-69 assessed for screen detected microcalcifications, there were no false positive cytologic diagnoses among the 167 malignant lesions assessed by FNAB, 4 false positive diagnoses among the 1320 malignant lesions assessed by conventional core biopsy and none among the 875 malignant VA core biopsy results. Rather than calculate averages from amalgamated data, the service performance data will be presented as it reflects real world service practice more closely. The mean of the false positive rates (FPR) reported by the services was 0% for FNAB, 0.31% for NCB and 0% for VACB. When all results for malignant microcalcifications assessed by core biopsies are combined the calculated mean figure for FPR is 0.18%.

At the individual service level, there were no false positive diagnoses with VA core biopsy or FNAB, although only 167 malignant lesions were assessed by FNA versus 875 by VA core biopsy, rendering this comparison unreliable. With a total of 1320 malignant microcalcifications assessed by conventional core biopsy,

Table 6. BSA National Accreditation Standards for Needle Biopsies

Standard	FNAB	Core Biopsy
Inadequate/insufficient results	<6%	No value set (2.18.2)
False negative result	<25% (2.18.1- L2)	No value set (2.18.2)
False positive result	<1% (2.19.6 (a) &(b))	<0.5% (2.19.7)
Absolute sensitivity	>60% (2.19.8)	>70% (2.19.10)
Complete sensitivity	>80% (2.19.9)	>80% (2.19.11)
PPV cancer diagnosis	>98% (2.19.12)	>99% (2.19.13)

in 2006 there were 3 false positive diagnoses on the basis of NCB at the one service. In 2007 another service reported a false positive diagnosis in microcalcifications assessed by conventional core biopsy. Since then, in the ensuing 3 years, no service reported any false positive core biopsy diagnosis for microcalcifications, either conventional or VA assisted. Further details of the circumstances of these false positive diagnoses are not available. The NAS has set the false positive rate of core biopsies at 0.5%.

False negative diagnoses for microcalcifications

The false negative rate is the number of false negative diagnoses established by each needle biopsy modality as the proportion of cases with final malignant outcome. Using the figures reported by each service, Table 4 highlights differences in the false negative rates of the three modalities. For assessing microcalcifications the FNR was 1.67% for FNAB, 5.00% for NCB and 1.53% for VACB. The far small number of FNABs performed compared to the other two modalities limits the opportunity for false positive and negative results to be experienced. There were 5 false negative cytologic diagnoses among 167 histologically proven malignant microcalcifications assessed by FNA (1.67%). For all core biopsies combined, there were 54 false negative diagnoses among a total of 2195 histologically malignant microcalcifications assessed by core biopsies, giving a calculated mean of 2.46%.

Only four services consistently assessed more than 15 microcalcifications annually with nonVA core biopsy. The average FNR of these services was 1.38%. For sites using few nonVA cores even isolated false negative diagnoses lead to relatively high FNRs. More specific information regarding number of cores, specimen x-ray facilities and team skill levels are not available.

For VA core biopsies the average FNR as reported by services was 1.53%. All but 4 services assessed fewer than 15 malignant microcalcifications by VA cores per year. When the performance of these 4 sites was considered, the average FNR was 2.14%, suggesting that 2-2.5% reflects the range of FNR that can be expected when services use VA core biopsy regularly for microcalcifications.

The mean reported rate of inadequate specimens for microcalcifications was 15.8% for FNAB, 1.43% for conventional core biopsy and 1.28% for VA core biopsy. Considering all 6233 core biopsy results reported for microcalcifications, there were a total of 124 inadequate results, for a calculated mean of 1.99%.

Discussion

The assessment of screen detected breast lesions is an important aspect of population screening for breast cancer and is a central responsibility of screening services. Recall to assessment, regardless of the final outcome, engenders anxiety and entails additional medical procedures, adding to the morbidity and cost of screening. Recall may also discourage participation (Sim et al., 2012). Organised population screening programs, including BreastScreen Australia, have determined a range of standards to monitor performance in relation to assessment. BreastScreen Australia's National Accreditation Standards (NAS) are

being updated to reflect current practice.

We report the performance of 12 BSA services over a 5 year period in assessing 20724 biopsies by various percutaneous needle biopsy techniques. The size range of services and their staffing and infrastructure facilities are representative of the overall BSA program.

We have presented outcomes for microcalcifications separately from other screen detected lesion types since the nature of the lesion impacts test performance. Also, rather than amalgamate all programs' results and calculate average figures, we base most of the analysis on the data provided for each service to allow for the impact of variations across sites. The important performance indices for the various biopsy techniques are presented in Table 5.

Overall the mean recall rates for the first and subsequent rounds of screening both exceed the NAS thresholds of 10% and 5% respectively. However, of the lesions biopsied 46.9% of LOTM and 35.8% of microcalcifications ultimately proved to be malignant, suggesting that although recall rates are higher than desirable, leading to additional imaging workup, lesions are being selected appropriately for biopsy, in that a high proportion of those biopsied are malignant.

The preoperative diagnosis rate is the proportion of all screen detected malignancies, inclusive of invasive cancers and DCIS, diagnosed on the basis of needle biopsies. Among the 10970 screen detected malignancies biopsied, the overall reported preoperative diagnosis rate was 91.8%. Based on the services reported data, these figures were 93.21% for the 2305 malignancies classified as microcalcifications and 96.84% for the 5420 malignant LOTM biopsied.

The preoperative diagnosis rates of both microcalcifications and LOTMC well exceed the current 75% minimum figure for this standard. The advances in imaging and biopsy techniques, particularly the improvements in conventional and vacuum assisted core biopsy facilities in the last 15 years permit accurate evaluation of screen detected lesions (Lieberman, 2002). The larger volume of the vacuum assisted core biopsies and the ability to retrieve multiple samples during the same procedure represent advantages, particularly for microcalcifications.

The high preoperative cancer diagnosis rates are a significant advance in the screening programs' performance. Almost uniformly, over 90% of all malignant lesions are being correctly diagnosed by percutaneous needle biopsies. A high preoperative diagnosis rate enables women and their doctors to discuss treatment options and plan for one stage, definitive surgical management, including axillary staging for invasive cancer (Lieberman et al., 1997; Smith et al., 1997; Whitten et al., 1997; Kaufman et al., 1998; Verkooijen et al., 2001; White et al., 2001). For DCIS there is evidence that a definitive preoperative diagnosis of malignancy reduces the need for reoperation to achieve clear margins. Brunning's systematic review of the use of core biopsies reported the odds ratio of completing treatment with a single surgical procedure was improved by 13.9 times when core biopsies were used (Bruening et al., 2010).

These data indicate that to reflect modern practice,

the bar for this important performance standard should be raised to 90%.

The preoperative diagnosis rate for microcalcifications is somewhat lower than for other lesions partly because their more dispersed distribution hampers targeting by needle biopsies. There is also a range of borderline lesions that may present as mammographic microcalcifications for which definitive exclusion of malignancy often requires examination of a larger volume of tissue than is available on core biopsies. These lesions include atypical ductal hyperplasia, lobular neoplasia and flat epithelial atypia (Bianchi et al., 2011). Less commonly mucocoele like lesions and papillary lesions may present as microcalcifications.

Surgery for non-malignant processes is an unintended consequence of assessment of screen detected lesion. While preoperative diagnosis rates were high in this survey, overall diagnostic surgical biopsy was required in 8.65% of microcalcifications and 4.96% of LOTM assessed by needle biopsies. The mean ratio of benign to malignant open biopsies was 2.73, indicating that most surgical procedures are for benign processes and highlighting this as an area for further improvement. Currently four standards monitor benign open biopsy rates in BSA services. The NAS specify that no more than 4% of women assessed in the first round and 3.2% of those undergoing assessment at subsequent rounds of screening should be found to have had non-malignant diagnoses at open biopsy. The mean of service data show that at 2.3% for round 1 and 1.89% for subsequent rounds of screening, both sets of figures are well within the stipulated bounds. With the greater awareness of borderline lesions in recent times, the proportion of benign open biopsies may rise in the coming years. Borderline lesions are a range of non-malignant lesions whose diagnosis on core biopsy serves as indications for surgical biopsy so as to avoid under diagnosis of concurrent malignancies. They include radial scars, atypical ductal hyperplasia, papillary lesions, flat epithelial atypia, lobular neoplasia, mucinous lesions and fibroepithelial lesions that raise the possibility of phyllodes tumour. These categories have varying upgrade rates of 15-50% (Dillon et al., 2007; El-Sayed et al., 2008; Bianchi et al., 2011; Rakha et al., 2011; Ohsumi et al., 2012). Repeat core biopsy using VA cores and other non-surgical approaches have been proposed to limit surgery. Further study of these lesions is needed to refine and if possible limit the indications for surgery.

Performance indices by biopsy modality: at present, the BSA accreditation standards do not provide guidance as to the preferred method of percutaneous needle biopsy of screen detected breast lesions. They specify a range of separate performance indicators for FNAB and core biopsy. The thresholds for satisfactory performance are different between the two modalities as shown in Table 6, the bar being set higher for core biopsies.

In analysing the outcomes for 20,724 lesions biopsied we have found important differences in the pattern of use and accuracy of the various needle biopsy modalities and have shown that test performance indicators vary by lesion type.

For microcalcifications, services are overwhelmingly

utilising core biopsies, the final diagnosis being made by VA and conventional core biopsies in 88.9% of cases. Overall, among microcalcifications FNA biopsy established the diagnosis in only 4.7% of cases. Time trends indicate that all but one service have abandoned FNAB for microcalcifications.

FNA is being used more often for assessing LOTM where a mean of 21.33% of lesions were diagnosed on the basis of FNA biopsy. However in this setting services are selecting lesions with benign imaging features for this test. Thus while overall 46.9% of all LOTM were histologically malignant, only 25.78% of histologically malignant non-calcified lesions had been assessed by FNAB. Only at one site did the mean proportion of malignant lesions assessed by FNAB reach 50%. This pre-selection affects test performance statistics. Time trends show a contraction in the use of FNA over the survey period.

The rate of inadequacy is a point of difference between FNA and core biopsies (Australian Government Department of Health and Ageing, 2009). For microcalcifications, the mean rate of inadequacy for FNAB was 15.8%, compared to 1.43% for conventional core biopsy and 1.28% for VA core biopsy. For LOTM the inadequacy rate for FNA was also high at 14.47%, contrasting with 0.81% for conventional core biopsy and 2.27% for VA core biopsy. These high rates of inadequate FNAs are reflected in the 15-20% lower rates of definitive diagnoses compared to core biopsies. To combat this, some services have on site pathologists to provide rapid assessment so women may have immediate repeat biopsies if needed. Disregarding repeat FNAs, on average 38% of LOTM undergoing FNA required further assessment by core biopsy (range 12.5 to 80.5%). There are presently no standards in place to monitor the number of biopsy episodes before a final diagnosis is reached.

False positive diagnoses of cancer were uncommon among LOTM. The false positive rate for FNAB was 0.36%, 6 times higher than that of conventional needle core biopsy at 0.06%, while there were no false positive diagnoses among VA core biopsies. False positive diagnoses were rare among microcalcifications regardless of biopsy technique used. No false positive diagnoses were reported by VA core biopsy or for the small number of FNAs performed. For conventional cores, the FPR was 0.31%. Among the 3757 biopsies there were a total of 4 false positive diagnoses, 3 from the same service early in the survey period and another the following year at another site. Since then, there were no other false positive core biopsy diagnoses reported for microcalcifications.

The current standards set the maximum false positive rate for core biopsies at 0.5% and allow twice that rate for cytology at 1% (BreastScreenAustralia, 2005). Clearly these thresholds far exceed what services are achieving with current techniques and require updating.

False positive diagnoses of breast cancer result in women being incorrectly informed that they have breast cancer and can lead to unnecessary staging investigations and unnecessary breast and axillary surgery. In terms of risk management, any false positive diagnoses, regardless of modality, may be regarded as serious clinical incidents. In mitigating this risk to client safety services should be

supported to utilize biopsy procedures with the highest accuracy rates and accreditation standards can contribute to this aim.

The false negative rate varied across modalities and lesion types. For LOTM the FNR was 4.91% for FNA compared to 1.98% for conventional core biopsy. For the very few LOTM assessed by VACB, the reported FNR is 2.94%. It is possible that some mass lesions are being selected for assessment by VACB if other techniques have failed or if there are particular circumstances guiding this decision. As such, these lesions may not be representative of all LOTMs. For microcalcifications, the few FNAs performed recorded a mean FNR of 1.67%. The FNR of conventional core biopsy was 5% compared to 1.53% for VACB. With large numbers of lesions being assessed by the 2 core biopsy modalities, this difference is reliable.

Both the higher sensitivity and the higher specificity of VACB found in this survey confirm prior studies (Lieberman, 2002). The larger tissue volume available with VACB improves calcium retrieval. In view of the superior performance indices of VA core biopsies for assessing microcalcifications, it is the recommended biopsy modality for microcalcifications. VACB may be performed under ultrasound or mammographic guidance.

The use of VACB has been expanded in parts of the UK screening program where rather than reflex surgical biopsy, lesions with borderline diagnoses on conventional core biopsy are being reassessed by VACB and lesion excision through VACB is being undertaken for radial scars and papillary lesions lacking atypia (Humber and Yorkshire Coast Cancer Council, 2011). However, at the time of the survey only 8 of the 12 services had access to VA core biopsy facilities in Australia.

Test accuracy is a useful summary statistic in that it is calculated as the sum of cases with true benign and true malignant results as a proportion of all cases assessed by each modality. The rate of accuracy of FNA was 72.75% for LOTM and 73.52% for microcalcifications. For core biopsies these figures were 92.1% for LOTM and 87.24% for microcalcifications.

The current accreditation standards do not address the additional problem of the inability of cytology to discern between DCIS and invasive cancer. Since both diagnoses are malignant, the failure to make this distinction is not recorded as an adverse event, but arguably it should be. Among services that used FNA, during the survey period, in addition to the 8 false positive cytologic diagnoses of cancer there were 17 cases of DCIS with positive cytology for lesions other than microcalcifications. Individual level data were not reviewed to establish if further biopsies were taken and determine patient consequences. Women with DCIS that has imaging features other than microcalcifications and positive cytology are potentially at risk of being counseled as if they have invasive cancer, may be exposed to unnecessary staging investigations and may undergo axillary surgery, which may not have been indicated if the in situ nature of the lesion had been established pre-operatively. In terms of risk management, the consequences of such misclassification represent a serious clinical incident. Current standards do not address this issue.

In the light of the above analysis, the broader issue for accreditation is that the maintenance of parallel targets for FNA and core biopsies, the disparity which accepts twice the rate of false positives with cytology as well as lower rates of sensitivity, when combined with the absence of guidelines for lesion selection may convey the message that both FNA and core biopsies are endorsed by accreditation bodies as appropriate for the assessment of all lesions, either choice being equally valid. The attractions of cytology, namely its lower immediate cost in consumables, limited infrastructure requirements, convenience and rapid turnaround may, in some services justify the inferior accuracy of the test and it continues to be used. This argument does not take into account the consequences of inadequate samples and inaccurate diagnoses and the need for on site pathologists.

Core biopsies, while by no means perfect are the current standard of care as they have demonstrably higher rates of definitive results and higher accuracy rates than FNAB. They permit distinction between in situ and invasive cancer and determination of grade, subtype and biomarkers evaluation status. In some situations, such as the diagnosis of invasive lobular carcinoma, this information may be an indication of preoperative MRI evaluation for surgical planning. There is evidence that biomarker studies carried out on core samples are highly reliable and perhaps more so than those undertaken in resection specimens.

Core biopsies have certain limitations. As noted previously, borderline lesions diagnosed on core biopsy may be harbingers of co-existing more established neoplasia and require further tissue diagnosis. In addition, due to sampling limitations, there is an upgrade rate for DCIS diagnosed on core biopsy in that detailed pathology examination reveals an invasive component in the resection specimen, not represented in the original core biopsy sample. Finally, as shown in this study, false negative diagnoses and inadequate samples are experienced with core biopsies, particularly when conventional core biopsy is used for microcalcifications. These cases are almost invariably due to the intended lesion not being sampled by the core biopsy, rather than being present and being misinterpreted during the histopathologic examination.

Recently the UK screening program has mandated that cytology should no longer be used alone to obtain a non-operative diagnosis of breast cancer (NHSBSP, 2010). They indicate if a rapid diagnosis is needed, FNAB may be used in addition to core biopsy but not instead of it. Imprint cytology of core samples, offered by 3 services, is another mechanism for reaching a rapid cytologic diagnosis. Similar guidelines on test selection are yet to be published by other screening programs.

Finally, the limited access to VA core biopsy facilities by some services should be addressed as this hampers optimal assessment for microcalcifications, leading to higher open biopsy rates.

In conclusion, this multicenter study benchmarks for the first time the performance of a range of BSA accredited services in over 20000 needle biopsies for screen detected lesions. Program performance across a range of indicators surpasses current minimum accreditation requirements and

in the light of this experience the bar may be reasonably raised in future updates of national accreditation standards. Gaps are identified in the current monitoring system for assessment and more specific guidance regarding test selection are recommended, along the lines mandated in the UKBSP.

Acknowledgements

We acknowledge the support of the BreastScreen Australia National Quality Management Committee and its individual members for their support of this audit. We acknowledge all services that have contributed their de-identified data for this collaborative effort. Their assistance has enhanced the scope and rigor of the undertaking. Our clients and the treating teams outside of the screening programs who have provided clinical care for BSA clients are gratefully acknowledged. Ms Ada Childs from the BreastScreen South Australia Monitoring Evaluation Unit provided valuable assistance in the design and pilot phase of this study.

References

- Bianchi S, Caini S, Renne G, et al (2011). Positive predictive value for malignancy on surgical excision of breast lesions of uncertain malignant potential (B3) diagnosed by stereotactic vacuum-assisted needle core biopsy (VANCB): a large multi-institutional study in Italy. *Breast*, **20**, 264-70.
- Bruening W, Fontanarosa J, Tipton K, et al (2010). Systematic review: comparative effectiveness of core-needle and open surgical biopsy to diagnose breast lesions. *Ann Intern Med*, **152**, 238-46.
- Dillon MF, McDermott EW, Hill AD, et al (2007). Predictive value of breast lesions of "uncertain malignant potential" and "suspicious for malignancy" determined by needle core biopsy. *Ann Surg Oncol*, **14**, 704-11.
- El-Sayed ME, Rakha EA, Reed J, et al (2008). Predictive value of needle core biopsy diagnoses of lesions of uncertain malignant potential (B3) in abnormalities detected by mammographic screening. *Histopathology*, **53**, 650-7.
- Farshid G, Rush G (2004). Assessment of 142 stellate lesions with imaging features suggestive of radial scar discovered during population-based screening for breast cancer. *Am J Surg Pathol*, **28**, 1626-31.
- Humber and Yorkshire Coast Cancer Council. 2011. Guidelines for B3 vacuum assisted biopsy [Online]. Available: http://www.hyccn.nhs.uk/Downloads/HYCCN_Final_Guidelines_for_B3_Vacuum_Assisted_Biopsy_April_2011_v1.3.pdf [Accessed November 2013 2013].
- Kaufman CS, Delbecq R, Jacobson L (1998). Excising the reexcision: stereotactic core-needle biopsy decreases need for reexcision of breast cancer. *World J Surg*, **22**, 1023-7.
- Liberian L (2002). Percutaneous image-guided core breast biopsy. *Radiol Clin North Am*, **40**, 483-500.
- Liberian L, LaTrenta LR, Dershaw DD (1997). Impact of core biopsy on the surgical management of impalpable breast cancer: another look at margins. *AJR Am J Roentgenol*, **169**, 1464-5.
- Ohsumi S, Taira N, Takabatake D, et al (2012). Breast biopsy for mammographically detected nonpalpable lesions using a vacuum-assisted biopsy device (Mammotome) and upright-type stereotactic mammography unit without a digital imaging system: experience of 500 biopsies. *Breast Cancer*.
- Rakha EA, Ho BC, Naik V, et al (2011). Outcome of breast lesions diagnosed as lesion of uncertain malignant potential (B3) or suspicious of malignancy (B4) on needle core biopsy, including detailed review of epithelial atypia. *Histopathology*, **58**, 626-32.
- Roder DM, Ward GH, Farshid G, et al (2014). Influence of service characteristics on high priority performance indicators and standards in the BreastScreen Australia program. *Asian Pac J Cancer Prev*, **15**, 5901-8.
- Sebandi M, Akrami M, Tabatabaee H, et al (2014). Breast cancer risk factors in women participating in a breast screening program: a study on 11,850 Iranian females. *Asian Pac J Cancer Prev*, **15**, 8499-502.
- Sim MJ, Siva SP, Ramli IS, et al (2012). Effect of false-positive screening mammograms on rescreening in Western Australia. *Med J Aust*, **196**, 693-5.
- Smith DN, Christian R, Meyer JE (1997). Large-core needle biopsy of nonpalpable breast cancers. The impact on subsequent surgical excisions. *Arch Surg*, **132**, 256-9.
- Verkooijen HM, Borel Rinke IH, Peeters PH, et al (2001). Impact of stereotactic large-core needle biopsy on diagnosis and surgical treatment of nonpalpable breast cancer. *Eur J Surg Oncol*, **27**, 244-9.
- White RR, Halperin TJ, Olson JA, et al (2001). Impact of core-needle breast biopsy on the surgical management of mammographic abnormalities. *Ann Surg*, **233**, 769-77.
- Whitten TM, Wallace TW, Bird RE, et al (1997). Image-guided core biopsy has advantages over needle localization biopsy for the diagnosis of nonpalpable breast cancer. *Am Surg*, **63**, 1072-7.