2018

Added value of second biopsy target in screen-detected widespread suspicious breast calcifications

Nathalie M. Falkner

Dana A. Hince
The University of Notre Dame Australia, dana.hince1@nd.edu.au

Gareth Porter

Ben Dessauvagie

Sanjay Jeganathan

See next page for additional authors

This article was originally published as:
Falkner, N. M., Hince, D. A., Porter, G., Dessauvagie, B., Jeganathan, S., Bulsara, M., & Lo, G. (2018). Added value of second biopsy target in screen-detected widespread suspicious breast calcifications. *Journal of Medical Imaging and Radiation Oncology, 62* (3), 299-306.

Original article available here:
https://dx.doi.org/10.1111/1754-9485.12715

Title: Added Value of Second Biopsy Target in Screen-Detected Widespread Suspicious Breast Calcifications

Running Head: Added Value of Second Biopsy Target in Screen-Detected Widespread Suspicious Breast Calcifications

Authors:

Dr. Nathalie M. Falkner BSc MBBS(Hons) [1]

Dr. Dana Hince PhD(Dist) BSc(Hons) [2]

Dr. Gareth Porter MBBCh FRANZCR [1, 3]

Dr. Ben Dessauvagie MBBS(Hons) BMedSc(Hons) FRCPA [4,5,6]

Dr. Sanjay Jeganathan MBChB MMED (RAD) FRANZCR [3,5,7]

Dr. Max Bulsara PhD(Dist) MSc BSc(Hons) [2]

Dr. Glen Lo MBBS BMedSc(Hons) FRANZCR [1,3,5,8]

[1] Department of Radiology, Sir Charles Gairdner Hospital, Nedlands, Western Australia.

[2] Notre Dame University, Fremantle, Western Australia.

[3] BreastScreenWA, Perth, Western Australia.

[4] Anatomical Pathology, PathWest Laboratory Medicine WA, Murdoch, Western Australia.

[5] Fiona Stanley Hospital, Murdoch, Western Australia.

[6] School of Pathology and Laboratory Medicine, University of Western Australia, Crawley, Western Australia.

[7] Perth Radiological Clinic, Perth, Western Australia.

[8] Western Ultrasound for Women, Perth, Western Australia.

Corresponding author:
Dr. Glen Lo
Dept of Radiology
G Block
Hospital Avenue
Nedlands 6009
Western Australia
Phone: 08 6457 2233
Fax: 08 6457 2091
Mobile: 0466 724 581
Email: glen.lo@health.wa.gov.au

Abstract

Introduction: There is controversy on the optimal work up of screen-detected widespread breast calcifications: whether to biopsy a single target or multiple targets. This study evaluates agreement between multiple biopsy targets within the same screen-detected widespread (≥25 mm) breast calcification to determine if the second biopsy adds value.

Methods: Retrospective observational study of women screened in a statewide general population risk breast cancer mammographic screening program from 2009 to 2016. Screening episodes recalled for widespread calcifications where further views indicated biopsy, and two or more separate target areas were sampled within the same lesion were included. Percentage agreement and Cohen's Kappa were calculated.

Results: 293317 women were screened during 761124 separate episodes with recalls for widespread calcifications in 2355 episodes. In 171 women, a second target was biopsied within the same lesion. In 149 (86%) cases the second target biopsy result agreed with the first biopsy (κ=0.6768). Agreement increased with increasing mammography score (85%, 86% and 92% for score 3, 4 and 5 lesions). Same-day multiple biopsied lesions were three times more likely to yield concordant results compared to post-hoc second target biopsy cases.

Conclusion: While a single target biopsy is sufficient to discriminate a benign vs. malignant diagnosis in most cases, in 14% there is added value in performing a second target biopsy. Biopsies performed prospectively are more likely to yield concordant results compared to post-hoc second target biopsy cases, suggesting a single prospective biopsy may be sufficient when results are radiological-pathological concordant; discordance still requires repeat sampling.

Keywords: Breast Imaging, Screening, Widespread Calcifications, Stereotactic, Biopsy.

INTRODUCTION

For screen-detected widespread segmental breast calcifications recommended for biopsy in Western Australia there is controversy on the optimal radiological work-up: specifically, is there added value of a second biopsy target within the same lesion. In Western Australia biopsy is recommended for screen-detected breast calcifications that are interpreted on further magnification views as score 3 (possibly malignant), 4 (suspicious for malignancy) or 5 (malignant), a scoring system that overlaps BI-RADs 3/4a, 4b, and 4c/5 categories, respectively[1,2]. In the case of widespread continuous or discontinuous but isomorphic screen-detected calcifications, intuitively one expects that a single target should provide a representative sample at histopathology. In the event of radiological-pathological discordance, a repeat biopsy remains indicated, assuming sampling error[1,2]. Additionally, local surgical staging preferences sometimes request biopsy-proven malignancy, detected as calcifications, from multiple target sites, with targets > 5 cm apart to confirm a widespread transverse or craniocaudal extent of disease as this information is useful in counseling patients who may require extensive surgery (e.g. mastectomy vs. breast conserving surgery). Some radiologists prefer to anticipate this surgical staging request at the time of initial biopsy and prospectively target opposite extents of widespread calcifications,

usually anterior/posterior extent on the same day. Some surgeons, however, consider anterior and posterior extent irrelevant, as resection margins dissect to the pectoralis fascia but appreciate transverse or craniocaudal extremes targeted, with two biopsies. In the case of discontinuous calcifications, even if isomorphic, documenting malignancy at two sites is important when counseling women who are motivated to pursue breast conserving surgery. Performing two biopsies on the day of diagnostic imaging work-up utilizes a second booking slot, thereby delaying access for other scheduled patients. Each additional biopsy target is an additional invasive test that may not be justified.

A recent North American study of 32 cases of only BI-RADS 4 or 5 category continuous segmental calcifications reported 100% histopathological agreement between paired anterior and posterior biopsies of morphologically similar segmental breast calcifications measuring 5 cm or more, suggesting that a second biopsy target to determine extent added no value to a single biopsy target[3]. There is a paucity of literature informing the optimal number of targets to be biopsied in cases of screen-detected, indeterminate, possibly malignant or malignant widespread breast calcification. In particular, for our Australian scoring system where any calcification interpreted as not definitely benign on magnification views biopsy is indicated (this would include BI-RADS 3 calcifications in a North American setting).

We aimed to evaluate in our Western Australian population of screen-detected widespread segmental continuous or discontinuous breast calcifications, that included score 3, 4 or 5, whether we could confirm 100% pathological agreement in biopsy pairs where 2 or more targets were sampled within the one lesion. We hypothesised that there would be 100% agreement between biopsy pairs, in all cases of screen-detected widespread calcification.

METHODS

Study Design

We conducted a retrospective observational study of histopathology reports for stereotactic core biopsies performed for widespread segmental continuous or discontinuous breast calcifications in consecutive women screened by BreastScreen WA, a government-funded general population breast screen program in Western Australia. Upon entering the BreastScreen WA screening program, women sign informed consent for information to be used for breast cancer research. Ethics approval was obtained from BreastScreen WA. In addition, institutional Quality Improvement activity approval was obtained which exempted Hospital Research Ethics Committee (HREC) review. Both the BreastScreen WA and WA Metropolitan Health Department Radiology Information System (RIS) databases were queried for women screened between 1 January 2009 and 30 April 2016 where widespread segmental continuous or

discontinuous breast calcifications were detected (coded as “WCA” or “widespread calcifications”) and, after magnification views were performed and biopsy was recommended, 2 or more biopsy targets were sampled. All widespread discontinuous calcifications were in a segmental distribution and constituted at least 3 groups of calcifications, no more than 20 mm apart.

Patient Selection

BreastScreen WA invites women via the Electoral Roll into a general population risk mammography screening program of biennial mammography between the ages of 50 and 74 years. Women at high risk of breast cancer have annual mammographic screening and women may self-present from age 40 and from age 75, without invitation. Pregnant women are excluded from screening.

Test methods

Screening consisted of bilateral craniocaudal (CC) and mediolateral oblique (MLO) 4 view digital mammography read in either hardcopy or soft copy, as the screening program transitioned to soft copy reporting during the study period, with no change in recall or cancer detection rates. Screening mammogram images were 2D. The assessment centres transitioned to include digital breast tomosynthesis (DBT) for workup of masses, distortions and asymmetries but not for the assessment of calcifications. Routine magnification and true lateral views were performed on all cases

recalled for evaluation of screen-detected breast calcifications. Screening mammograms are prospectively double-read by radiologists with subspecialty fellowship training in breast imaging, with reports structured according to the NBCC Synoptic Breast Imaging Report Guidelines[2]. Screen-detected breast calcifications are scored according to a grading system (not equivalent to BI-RADS), which classifies mammography lesions on a scale of 1 to 52. Score 1 indicates no significant abnormality (normal); score 2 is benign; score 3 is possibly malignant (which overlaps BI-RADS 3, probably benign and BI-RADS 4a, biopsy indicated but benign pathology accepted); score 4 is suspicious for malignancy and score 5, malignant[1,2]. At BreastScreen WA, calcifications are further categorised on the basis of extent. If the calcifications are widespread (25 mm or greater) they are encoded as “WCA” (widespread calcifications). If grouped calcifications are smaller in extent (smaller than 25 mm) they are reported as “LCC” (localised cluster of calcifications). A widespread area of discontinuous breast calcifications of the same morphology may be reported as “WCA” or multiple “LCC” at screening but, after magnification views when biopsy is recommended for work-up, and at subsequent surgery, may be considered a single lesion. We included cases coded as “WCA” (25 mm or greater) that included segmental continuous isomorphic calcifications (Figure 1) and segmental discontinuous

isomorphic calcifications encoded as “WCA” or multiple “LCC” but considered radiologically part of the same process/lesion (Figure 2).

The screening mammograms or further magnification views were validated for study inclusion by one of three Consultant Radiologists (all of whom were authors in the current study) for validation of WCA size and morphology.

Data was retrospectively collected from multiple institutions. Biopsies were performed at 14G, 12G or 9G, or a combination, using a Bard Magnum or Suros vacuum-assisted devices depending on institution but the screening program quality assurance mandates that at least 5 core samples are taken and specimen radiograph confirms the presence of target calcifications. Histopathology scores were reported according to a pathological 5-tier system where score indicated non-diagnostic (1), benign (2), indeterminate/atypical (3), suspicious for malignancy (4) and malignant (5) diagnoses. All pathology results were assessed for radiological-pathological concordance by the reporting radiologists.

Histopathology report scores were validated by a single consultant pathologist. Lesions were considered in histopathological agreement if the reported numerical pathology scores categories matched. Diagnostic pathology scores were also further categorised into binary clinical management categories of “benign” (return to routine screening) vs. “not benign” (atypical/suspicious/malignant –

requires repeat or excision biopsy or definitive surgery). Lesions were considered in clinical agreement if the binary categorisation of each biopsy matched. Percentage agreement was calculated by dividing the number of paired biopsies in agreement by the total number of paired biopsies for each dataset.

Statistical Analysis

Agreement between biopsies was assessed using Cohen's kappa. Firstly, a weighted kappa was calculated for the histopathology scores, which penalised disparate scores progressively as the difference between them increased. Secondly, kappa was calculated for binary clinical management categories (benign vs. not benign) for each biopsy pair, both for the whole cohort and for specific subgroups. In addition, to investigate the association between clinical agreement and demographic and clinical factors, univariate logistic regression models were fitted to the data. Robust standard error estimates were used to account for the three participants who had two separate encounters with the service for biopsy of different widespread calcifications. $P<0.05$ was considered statistically significant and all analyses were conducted using Stata v14.1.

RESULTS

BreastScreen WA conducted 761124 screening episodes for 293317 women between 01 Jan 2009 and 30 April 2016. During this time,

there were 2355 recalls for further imaging (magnification views of widespread calcifications), of which 443 were benign and patients were returned to routine screening. In 1912 recalls for WCA, further views were read as indeterminate, suspicious or mammographically malignant and biopsy was recommended. In 1295 of those cases multiple biopsies were performed. Of those 1295 cases where two or more biopsies were performed, in 174 cases two or more stereotactic core biopsy targets were performed within a single widespread calcification (Figure 3). Other women had biopsies of multiple different lesions.

There were 174 lesions in 171 women that underwent paired biopsies of 2 or more target areas within the same lesion. The distribution of the two most different (if more than 2 targets) biopsy results of reported histopathology scores (1 = non-diagnostic, 2 = benign, 3 = indeterminate/atypical, 4 = suspicious for malignancy, 5 = malignant) is displayed in Table 1. Overall percentage agreement for the 174 lesions biopsied was 79% and 86% for histopathological and clinical agreement, respectively (Table 2). Of the 174 lesions, 143 paired biopsies were performed prospectively on the same day, while the remaining 31 were performed following a call back following initial biopsy for wider sampling (surgical staging) or radiological-pathological discordance. The latter cases are referred to as post-hoc second target biopsy cases. For the 143 prospective paired biopsies, there was 84% and 89%

histopathological and clinical agreement, respectively. For post-hoc second target biopsy cases, histopathological agreement was 55% and clinical agreement was 71%.

Of 174 paired biopsies, 94 were performed both with a 14G Bard Magnum biopsy device, 37 were performed with vacuum assisted devices and 4 were performed with a combination of the two (Table 3). In 39 biopsy pairs the device used was not recorded. We observed no statistically significant difference in agreement (either histopathological score or clinical assessment) if the biopsy was performed at 14G, or with vacuum-assistance (Table 3). For those biopsies performed with a 14G Bard Magnum biopsy device, there was 77% and 85% histopathological and clinical agreement, respectively. For biopsies performed with vacuum-assistance, histopathological agreement was 81% and clinical agreement was 84%. Biopsies performed with a combination of devices yielded 50% histopathological and clinical agreement while those where the biopsy device was not stated demonstrated 85% histopathological agreement and 92% clinical agreement. These latter large differences are observed in only 4 cases and are likely due to chance.

We observed a trend towards increasing percentage agreement with increasing degree of mammography suspicion, with histopathological percentage agreement of 77%, 81% and 85% and

clinical agreement of 85%, 86% and 92% for subsets of mammographic scores of 3, 4 and 5, respectively (Table 2). Lesions measuring 50 mm or larger (n=104) were in histopathological agreement in 82% of cases and in clinical agreement in 87% of cases (Table 2).

Cohen's kappa was calculated to determine if there was agreement between all biopsy pairs (Table 4). Overall kappa for concordance was 0.68 whether by 5-tier histopathology score (CI 0.55-0.81) or binary benign-malignant score (CI 0.53-0.82). However, agreement was never perfect.

Subset analysis (Table 4) demonstrated statistically significant agreement between biopsy pairs in WCA measuring 50 mm or more (kappa=0.6846, CI 0.49-0.88) and in those WCA with a mammography score of 3 (kappa =0.6902, CI 0.52-0.86). There was only slight agreement between biopsy pairs in those WCA with a mammography score of 4 or 5, which was not statistically significant (kappa=0.1848, CI -0.09-0.46). For prospective biopsy pairs, there was statistically significant agreement between biopsy pairs (kappa = 0.7505, CI 0.59-0.91) while in post-hoc second target biopsy cases, agreement was not statistically significant (kappa = 0.3178, CI -0.03-0.67).

Table 5 displays the odds ratios from the univariate logistic regression models for clinical agreement. Only biopsy timing was

significantly associated with agreement, with prospectively (same day) biopsied patients being over 3 times more likely to have results in agreement compared to those in the call back group (OR 3.25, p = 0.014). The aforementioned trend towards increasing percentage concordance with increasing degree of mammography suspicion was not statistically significant when analysed for clinical agreement (OR 1.16, p = 0.787 and OR 2.17, p = 0.47).

DISCUSSION

For screen-detected widespread segmental breast calcifications recommended for biopsy in Western Australia there is controversy on the optimal radiological work-up: specifically, whether there is added value of a second biopsy target within the same lesion. The surgical decision between breast conservation and mastectomy is influenced by several factors, including the extent of disease. Larger lesions of 50 mm or greater require more extensive surgery to achieve clear margins with recurrence largely being influenced by margin status[4]. Anecdotally, it is useful to have histopathological results consistent with mammographic appearances of widespread cancer when counselling women for more aggressive therapy.

Our data demonstrate substantial and statistically significant but imperfect agreement between reported histopathology scores obtained from two or more sites within screen-detected widespread continuous or discontinuous calcifications. These results differ from

those of Raj et al (2016) who demonstrated 100% agreement between anterior and posterior biopsies in segmental breast calcifications 50 mm or greater[3]. Results of these two studies may differ for a number of reasons. For example, our study included cases between 25 and 50 mm and was not limited to anterior-posterior lesion extent, whilst Raj et al (2016) excluded calcifications < 50 mm in size. However, in the current study's subgroup of patients with widespread calcifications measuring 50 mm or greater, where the majority had two targets biopsied prospectively at anterior and posterior margins anticipating a surgical staging request, 100% agreement in histopathological result (benign vs. not benign) was not observed (we observed 87%). This suggests that in up to 13% of cases with clinically divergent results, sampling of multiple sites within widespread calcifications is arguably justifiable.

The unexpected finding of only slight agreement between biopsy pairs for mammography scores 4 or 5 can be explained by the smaller sample size of this subset (n = 50) and the inherent greater probability of concordance being due to chance alone. The observed probability of agreement (0.88) is not that much greater than that expected due to chance (0.85), hence kappa is small. In comparison, the probability of agreement due to chance for mammography score 3 lesions (n = 124) was 0.50 and the observed agreement was 0.84. Therefore, although the percentage

agreement is equivalent in both groups, the expected agreement is very different and hence, so are the kappa values.

The finding that prospectively biopsied (same day) cases were 3 times more likely to have results in agreement suggests that performing paired biopsy targets rather than a single biopsy target may not be necessary for screen-detected widespread breast calcifications. The assessment of radiological-pathological concordance (e.g. accept a benign histopathology result) is made at the time of initial biopsy. In 3 of the 11 post-hoc second biopsy cases recalled for radiologic-pathologic discordance, where a second biopsy target was sampled at a later date, the decision to repeat biopsy was made following second opinion or multidisciplinary meeting. It should be noted that this analysis was exploratory in nature and the sample size for some models was quite small. Therefore, given their potential clinical utility, it is important to demonstrate that these results can be replicated in a larger, prospectively collected, cohort.

The main limitation of this study is its retrospective design, and consequent limited availability of desired data. However, the cases were prospectively enrolled and screened in a statewide program, a population applicable to routine general risk women. Study population heterogeneity limits ability to generalise findings to change program policy. For example, there was heterogeneity in the

gauge of biopsy and use of vacuum assisted techniques and as such the biopsy sensitivity, accuracy and risk of underestimation[5,6] varied between study population subsets. However, the aim of the study was to identify the presence of cases where paired biopsies within one lesion yielded discordant results, and biopsies performed with or without vacuum assistance showed this.

A further study limitation is the potential for selection bias from retrospective study design: not all cases of screen-detected widespread breast calcifications where further views recommended biopsy and biopsy was performed, had paired biopsies. Of 1912 screens with widespread calcifications recalled where biopsy was recommended, 617 were excluded because either only single target biopsy or no biopsy was performed. Of the remaining 1295 cases where multiple biopsies were performed, the majority were biopsies of different lesions, for example a mass or contralateral breast lesion.

CONCLUSION

Our data demonstrate statistically significant but imperfect agreement between reported histopathology scores obtained from two or more sites within single screen-detected widespread continuous or discontinuous calcifications (considered single lesions). In 174 lesions in 171 women that underwent paired biopsies the majority (86%) of biopsy pairs were in pathological

agreement, with the second biopsy target adding value in 14% of cases where there was disagreement between biopsy pairs. Our data suggest that the second biopsy target is particularly valuable in cases of radiological-pathological discordance or if the calcifications are interpreted as indeterminate, rather than definitively malignant in appearances. Further research is needed to identify factors that predict cases of pathological disagreement.

Acknowledgements

We thank the participating women; Dr Elizabeth Wylie for advice on study design and data interpretation; Ruth Bostock and Kim Ooi for BSWA data extraction; Dr Anita Bourke for advice on study design and interpretation; Mr Lee Jackson for a surgical perspective on data interpretation.

References

1. Sickles EA, DOrsi CJ, Bassett LWEA. ACR BI-RADS® Mammography. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. American College of Radiology. Reston, VA; 2013.

2. Zorbas H, Ahern V, Bilous M, Kossof M, Mann B, Masters J, et al. Synoptic Breast Imaging Report. 2007 Jan 1;(6). Available from: http://www.worldcat.org/title/recommendations-for-a-synoptic-breast-imaging-report-including-imaging-classification-system/oclc/225576222

3. Raj SD, Sedgwick EL, Severs FJ, Hilsenbeck SG, Wang T, Sepulveda KA. Stereotactic Biopsy of Segmental Breast Calcifications. Academic Radiology [Internet]. 2016 Jun 1;23(6):682–6. Available from: http://dx.doi.org/10.1016/j.acra.2015.12.025

4. Reeve T, Bandaranayake R, Barraclough B, Bilous M, Coates A, Draper M, et al. National Health and Medical Research Council. National Breast Cancer Centre: Clinical Practice Guidelines for the Management of Early Breast Cancer [Internet]. 2001. Available from: https://canceraustralia.gov.au/system/tdf/publications/cpg-clinical-practice-guidelines-management-early-breast-cancer_504af03111a52.pdf?file=1&type=node&id=2941

5. Park H-L, Kim LS. The Current Role of Vacuum Assisted Breast Biopsy System in Breast Disease. Journal of Breast Cancer. 2011;14(1):1-7. doi:10.4048/jbc.2011.14.1.1.

6. O'Flynn EA, Wilson AR, Michell MJ. Image-guided breast biopsy: state-of-the-art. Clinical Radiology. 2010: 65(4):259-70. doi: 10.1016/j.crad.2010.01.008.

Figure Legends

Figure 1 - Widespread segmental breast calcifications.
Female age 50 years, 3rd round screening mammogram, recalled for further views for right upper inner quadrant widespread segmental calcifications > 40 mm, Score 5 (malignant). Two targets, anterior and posterolateral, were biopsied with histopathologic and clinical agreement showing high grade DCIS, no invasive malignancy.

Figure 2 - Widespread discontinuous calcifications.
Female age 61 years, 6th round screening mammogram. Recalled

for left lower inner quadrant widespread discontinuous but isomorphic segmental calcifications, 70 mm diameter, score 5. Anterior and posterior biopsy targets, with marker clip placement, with histopathologic and clinical biopsy result agreement: malignant, high nuclear grade, predominantly solid pattern ductal carcinoma in-situ (DCIS) with central comedo necrosis and calcification.

Figure 3 – Patient Flow

Tables

Table 1 - Comparison of biopsy results (histopathology score) between paired biopsy targets for all cases (n=174). Path score 1 = non-diagnostic, 2 = benign, 3 = indeterminate/atypical, 4 = suspicious for malignancy, 5 = malignant.

Table 2 - Baseline demographic and clinical characteristics of participants

Table 3 – Odds ratios from univariate logistic regression models for the association between agreement and biopsy method

Table 4 – Statistical Summary whole cohort and prospective subsets

Table 5 – Odds ratios from univariate logistic regression models for the association between agreement and each independent variable

		BIOPSY 2nd TARGET					
BIOPSY 1st TARGET		**PATH SCORE 1**	**PATH SCORE 2**	**PATH SCORE 3**	**PATH SCORE 4**	**PATH SCORE 5**	**TOTAL**
		n	n	n	n	n	
PATH SCORE 1	n	0	0	0	0	0	0
PATH SCORE 2	n	0	45	3	1	3	52
PATH SCORE 3	n	0	7	14	0	2	23
PATH SCORE 4	n	0	0	0	2	1	3
PATH SCORE 5	n	1	10	10	0	75	96
TOTAL		1	62	27	3	81	174

Table 1. Comparison of biopsy results (histopathology score) between paired biopsy targets for all cases (n=174). Path score 1 = non-diagnostic, 2 = benign, 3 = indeterminate/atypical, 4 = suspicious for malignancy, 5 = malignant.

Table 2 - Baseline demographic and clinical characteristics of participants

	n	Median age of women at biopsy (range) years	Median WCA size (range) mm	Histopathological Agreement %	Clinical Agreement %
Number of cases	174	55 (31-78)	50 (25-160)	79% (137/174)	86% (149/174)
Mammographic Score 3	124	55 (31-78)	50 (25-125)	77% (96/124)	85% (105/124)
Mammographic Score 4	37	57 (43-78)	55 (26-160)	81% (30/37)	86% (32/37)
Mammographic Score 5	13	60 (49-78)	66 (35-100)	85% (11/13)	92% (12/13)
Size ≥ 50mm	104	56 (41-78)	61 (50-160)	82% (85/104)	87% (90/104)
Prospective cases	143	55 (31-78)	50 (25-150)	84% (120/143)	89% (127/143)
Post-hoc second target biopsy cases	31	55 (41-78)	60 (30-160)	55% (17/31)	71% (22/31)

Table 3 – Odds ratios from univariate logistic regression models for the association between agreement and biopsy method

Variable		Agreement	Disagreement	OR (95% CI)	p-value
Biopsy gauge	14 gauge	80	14	0.911 (0.387 to 2.14)	0.831
	Vacuum	31	6	0.832 (0.305 to 2.27)	0.719

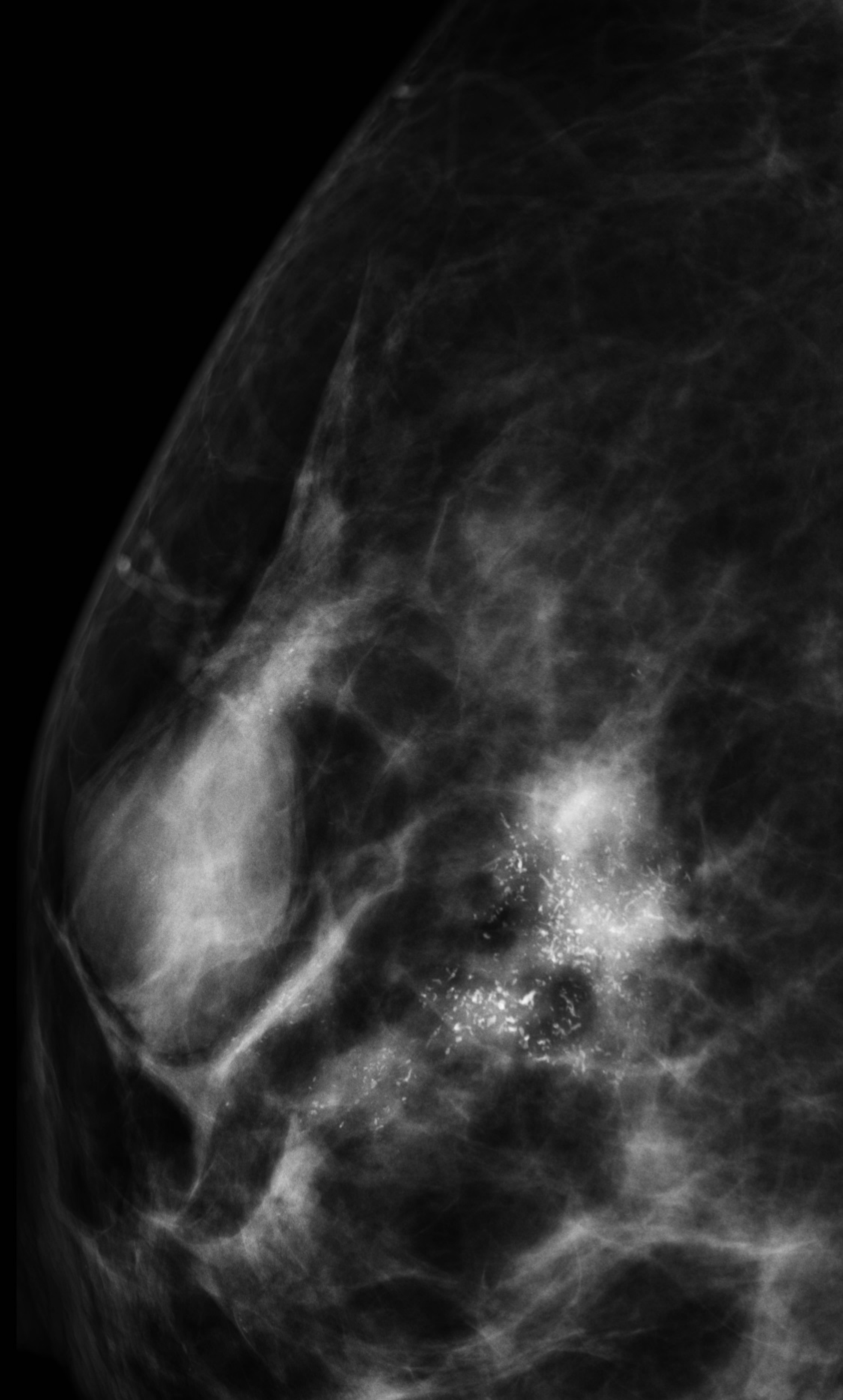
Table 4 – Statistical Summary whole cohort and prospective subsets

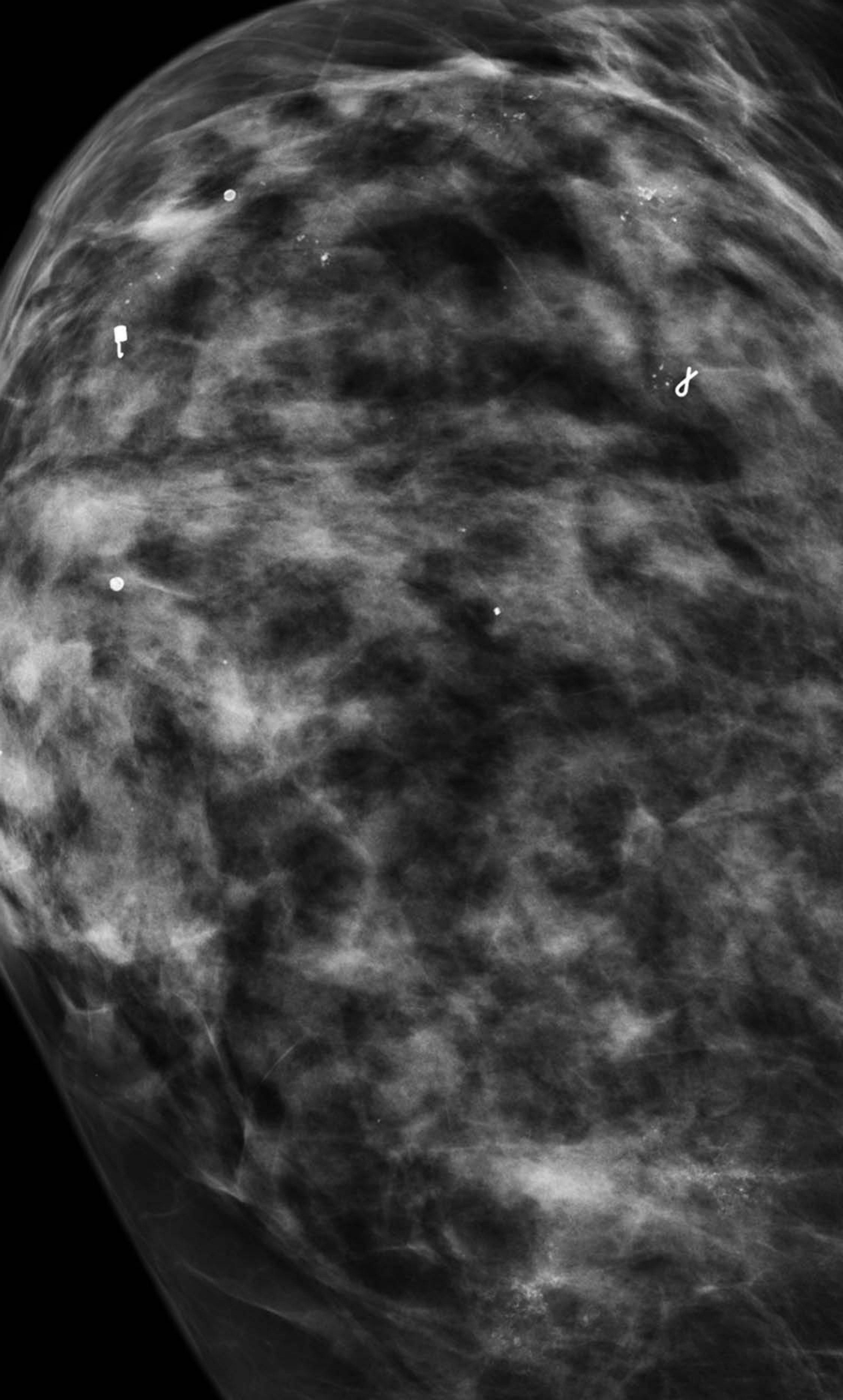
		n	Agree Benign	Agree Not Benign	Disagree	kappa	Lower CI	Upper CI	Standard error
WHOLE COHORT	Whole cohort	174	45	104	25	0.6768	0.53	0.82	0.0750
	Age <50 years	36	9	19	8	0.5200	0.20	0.84	0.1655
	Age ≥50 years	138	36	85	17	0.7194	0.55	0.88	0.0842
	Continuous WCA	149	40	85	24	0.6473	0.81	0.49	0.0810
	Discontinuous WCA	25	5	19	1	0.8837	0.49	1.27	0.1986
	Size ≥50mm	104	25	65	14	0.6846	0.49	0.88	0.0977
	Mammogram Score 3	124	44	61	19	0.6902	0.52	0.86	0.0883
	Mammogram Score 4 or 5	50	1	43	6	0.1848	-0.09	0.46	0.1414
	Dense Breasts	79	24	42	13	0.6534	0.43	0.87	0.1121
	Non-dense Breasts	94	20	62	12	0.6839	0.52	0.85	0.1008
	PHx Breast or Ovarian CA	8	1	6	1	0.6000	-0.04	1.24	0.3240
	No PHx Breast or Ovarian CA	165	43	98	24	0.6741	0.52	0.83	0.0771
	Family History of Breast CA	39	11	22	6	0.6667	0.35	0.98	0.1591
	No Family History of Breast CA	133	32	82	19	0.6690	0.50	0.84	0.0856
	Post-hoc second target biopsy cases	31	5	17	9	0.3178	-0.03	0.67	0.1791
PROSPECTIVE CASES ONLY	Prospective	143	40	87	16	0.7505	0.59	0.91	0.0826
	Age <50 years	31	8	18	5	0.6437	0.30	0.99	0.1755
	Age ≥50 years	112	32	69	11	0.7805	0.60	0.96	0.0936
	Continuous WCA	124	36	72	16	0.7202	0.55	0.89	0.0884
	Discontinuous WCA	19	4	15	0	1.0000	0.55	1.00	0.2294
	Size ≥50mm	84	22	55	7	0.8037	0.59	1.01	0.1080
	Mammogram Score 3	97	39	46	12	0.7530	0.56	0.95	0.1001

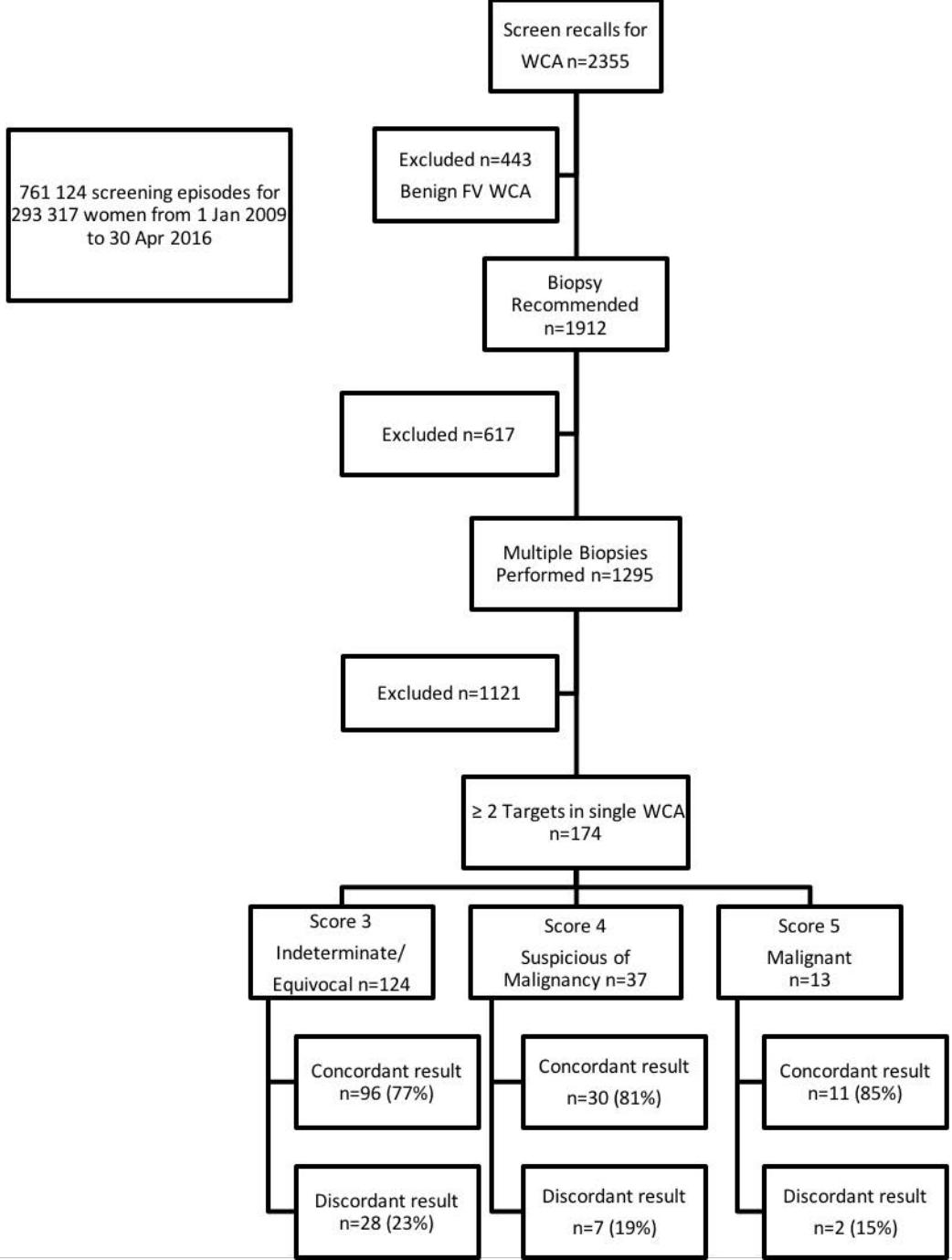
Mammogram Score 4 or 5	46	1	41	4	0.2923	0.02	0.56	0.1379
Dense Breasts	64	22	33	9	0.7120	0.47	0.95	0.1234
Non-dense Breasts	78	17	54	7	0.7696	0.55	0.98	0.1117
PHx Breast or Ovarian CA	5	1	4	0	1.0000	0.12	1.00	0.4472
No PHx Breast or Ovarian CA	137	38	83	16	0.7397	0.57	0.88	0.0843
Family History of Breast CA	34	10	20	4	0.7434	0.41	1.08	0.1701
No Family History of Breast CA	107	28	67	12	0.7430	0.56	0.93	0.0952

Table 5 – Odds ratios from univariate logistic regression models for the association between agreement and each independent variable

Variable		Agreement	Disagreement	OR (95% CI)	p-value
Biopsy timing	Post-hoc second target biopsy	22	9	1	0.014
	Prospective	127	16	3.25 (1.27 to 8.29)	
Age group	< 50	28	8	1	0.141
	≥ 50	121	17	2.03 (0.79 to 5.23)	
Breast density	Not dense	82	12	1	0.496
	Dense	66	13	0.74 (0.32 to 1.74)	
Family history breast cancer	No	114	19	1	0.865
	Yes	33	6	0.92 (0.34 to 2.50)	
WCA Continuous	No	24	1	1	0.145
	Yes	125	24	0.22 (0.03 to 1.69)	
Personal history breast or ovarian cancer	No	141	24	1	0.873
	Yes	7	1	1.19 (0.14 to 10.19)	
Mammogram Score	3	105	19	1	0.787
	4	32	5	1.16 (0.40 to 3.36)	0.47
	5	12	1	2.17 (0.46 to 17.81)	






761 124 screening episodes for 293 317 women from 1 Jan 2009 to 30 Apr 2016
Screen recalls for WCA n=2355
Excluded n=443 Benign FV WCA
Biopsy Recommended n=1912
Excluded n=617
Multiple Biopsies Performed n=1295
Excluded n=1121
≥ 2 Targets in single WCA n=174
Score 3 Indeterminate/ Equivocal n=124
Score 4 Suspicious of Malignancy n=37
Score 5 Malignant n=13
Concordant result n=96 (77%)
Discordant result n=28 (23%)
Concordant result n=30 (81%)
Discordant result n=7 (19%)
Concordant result n=11 (85%)
Discordant result n=2 (15%)