Original Article

Stratification of risk of malignancy in atypical breast fine needle aspiration: A cytomorphological approach

J. Weigner [1&3], I. Zardawi [2&3], S. Braye[1&3], P. McElduff[3]

[1]*Pathology North, John Hunter Hospital, Newcastle,* [2]*Pathology Queensland, Cairns, Queensland and* [3]*University of Newcastle, NSW, Australia*.

Summary

Although atypical (C3) breast cytology is a legitimate reporting category, it is not very useful clinically because it yields benign as well as malignant histological outcomes. Therefore, a cytomorphological approach to minimise inappropriate use of the C3 category may improve the clinical usefulness of this category. In a previous training set of 180 atypical breast cytology (C3) cases we had identified the best discriminating cytological criteria to predict the histological outcomes (malignant, benign proliferative and benign non-proliferative). Using a selection of these previously identified statistically significant criteria (cystic background, cohesiveness, myoepithelial cells or bare bipolar nuclei, papillary fragments and tubules) we tested their ability to predict the same outcomes on 182 subsequent C3 cases (validation set). The diagnostic accuracy for each outcome was compared with the training set. A probability calculator with nominated cut-points was developed to stratify and re-grade C3 cases. Cases outside the cut-points were either up-graded to suspicious (C4) or down-graded to benign (C2). Re-graded cases were reviewed to identify reasons for over or under-interpretation. Statistical analysis showed comparable diagnostic accuracy between the training set and the validation set for malignant and benign proliferative outcomes but not for the benign non-proliferative outcome. The probability calculator resulted in an up-grade of 18% (33/182) ofthe cases. Malignant histology was seen in 25/33 (76%) of the up-graded cases but 8/33 (24%) produced a proliferative non-malignant outcome. A meaningful lower cut-point could not be established without including malignant cases. The C3 category remains a legitimate with a heterogeneous mix of pathological entities. Attempts to minimise inappropriate allocation of C3 cases into this category have met with limited success.

Key words: C3, atypical, breast FNA, probability calculator, risk of malignancy

Author for correspondence: *Ibrahim ZardawiMBChB, MSc, FRCPA, FRCPath, FIAC, Dip Cytopath (RCPA), FFSc (RCPA), Dip Euro Board of Pathology. Conjoint Professor, University of Newcastle and Senior Pathologist at* Pathology Queensland, Cairns Hospital, *Australia. E mail:* ibrahim.zardawi@newcastle.edu.au

http://orcid.org/0000-0002-3730-2892

Introduction

Fine needle aspiration (FNA) is a well–established investigative procedure useful in the diagnosis of palpable and impalpable breast lesions. The results are reported using a five-tier categorical system based on probability of malignancy. This system complements the clinical and radiological BIRAD reporting systems and these modalities formulate thetriple test. The cytology categories include; C1 (non-diagnostic), C2(benign), C3(atypical), C4(suspicious) and C5 (malignant). C1 category usually reflects technical difficulties and provides little informative value. C2 and C5 have high predictive values as shown by many studies.The suspicious C4 category is not as certain as C5 but still suggests a high probability of malignancy. The C3 category is an equivocal category commonly used when there is some diagnostic uncertainty regarding the true nature of the lesion (1-5). There are no standard clinical management strategies for C3 cytology. Patients with C3 cytology and with low levels of suspicion clinically and on imaging are often observed and the FNA is sometimes repeated 3 to 6 months later. Immediate repeat FNA, core biopsy or even surgical excision is performed in C3 instances with higher levels of clinical or radiological concern (6, 7). Therefore determining the risk of malignancy within the C3 category may improve the clinical utility of C3 FNA (8).

We, like others have shown a heterogeneous mix of pathological outcomes during the follow up of C3(5, 9-14). In our earlier study, of 230 atypical (C3) cases, we found 37.4% to be malignant on follow-up and the remaining 62.6% yielded benign outcomes. We grouped the final outcomes into three mutually exclusive histological entities (malignant, benign proliferative and benign non-proliferative). These outcomes were established by correlating the histology

in 87%, repeat FNA in 3.5% and clinical and radiological follow-up 2 years post index FNA episode in 9.5%. The main reasons for placing these cases into the C3 category were suboptimal in 34.8% of the cases and the remaining 65.2% were diagnostically challenging, making interpretation difficult (9,15).

To further our investigations, we tested the ability of a number of cytomorphological criteria to predict malignant, benign proliferative or benign non-proliferative pathological outcomes from the training set of C3 cases (16).

Using the data from the training set, we now aim to create and validate a cytomorphological approach based upon the probability of cancer to stratify the final outcomes of C3. The use of a probability calculator to deliver a practical approach to help in the diagnosis of difficult breast lesions and their appropriate allocation into the C3 category will be tested.

Material and Methods

Ethical approval was granted for this study by the Hunter New England Human Research Ethic Committee and the University of Newcastle, N.S.W., Australia.

A search of the laboratory archive to identify C3 cases was conducted for both the training set and the validation set. Search terms used were reports containing any of the following phrases; "atypical cytological pattern", "indeterminate cytological pattern", "suggestive/consistent with papillary lesion", "Code 3", "Category 3" or "C3". The training set contained 230 FNA sourced C3 cases comprising 73 screen detected lesions and 157 lesions from symptomatic patients. All specimens were from women aged between 16-85 years of age, with an average age of 56.3 years. In comparison, the validation set

contained228 C3 cases consisting of 94 screen detected lesions and 134 lesions from symptomatic patients with an average age of 53. 1years and age range from 23 to 89 years.

The observations from a blind rescreen of the training set were previously analysed using a logistic regression and receiver operating characteristic (ROC) curve to determine which cytomorphological criteria showed a statistically significant association with the pathological outcomes of C3. The outcomes were classified into three groups; malignant, benign proliferative and benign non-proliferative. The malignant outcome included ductal carcinoma in situ (DCIS), invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC). Benign papilloma, fibroadenoma (FA), proliferative fibrocystic change (FCC), complex sclerosing lesions (CSL)/radial scar (RS), sclerosing adenosis, epithelial hyperplasia and uncommon specific tumours such as hamartoma, adenoma and benign phyllodes were grouped together under the umbrella of benign proliferative outcome. The remaining conditions comprised the benign non-proliferative outcome and included non-proliferative FCC, lactation associated change, radiation-induced changes, fat necrosis and clinically or radiologically stable but cytologically atypical lesions, not otherwise specified, with no histological follow up.

Five common key criteria with high predictive power for the final outcomes (malignant, benign proliferative or benign non-proliferative) were identified from this previous study (16). These were:
1. cystic background defined as having foamy macrophages within a proteinaceous wash
2. cohesiveness which was assessed as being predominantly cohesive if more than 95% of material was present in groups or discohesive if more than 5% of cells presented as single intact cells

3. myoepithelial cells or bare bipolar nuclei which included the presence or absence of small dark elliptical nuclei overlying sheets of ductal epithelial cells or bipoler nuclei found In the background of smear
4. papillary fragments which were noted to be present if there was peripheral palisading or anatomical edges with a fibrovascular core
5. tubular structures which were described as angulated hose-like structures with parallel sides, sharp ends and no obvious myoepithelial cells.

Age was also factored into the calculations because of its significant predictive power for each of the final outcomes and was grouped into 2 categories, less than 50 years old and 50 years or older.

Cellularity was used as a filter for inclusion in the study, hence only cases with more than 30 groups of epithelial cells (a group of cells was defined as more than 3 epithelial cells in a cluster) were included. This left 180 cellular cases from the original 230 training set from which the statistical analysis was performed.

A blind rescreen of the validation set was conducted by one author (JW), to assess for the presence or absence of the selected cytomorphological criteria only. These observations were recorded and the probability of the combination of identified key criteria to predict each of the final outcomes was calculated. All statistical analyses were completed using Stata 11.2 (17).

In our previous study, we developed an algorithm using the key criteria to predict malignant, benign proliferative and benign non-proliferative outcomes. We now apply this algorithm to the validation set but use the coefficients derived from the training set. Table 1 lists the coefficient values for each key criterion. A ROC curve and Area Under the Curve (AUC) was then used to assess the

accuracy to predict the outcome in the validation set (16,18). ROC curves plot sensitivity against one minus specificity (1-specificity) and the AUC is used to measure the accuracy and discriminating power of the combined criteria.

Table 1. The coefficient values calculated from the training set which were subsequently applied to the validation set are shown.

	Malignant	Proliferative	Benign nonproliferative
Intercept	-3.835	+3.567	-2.461
Myoepithelial/bare bipolar nuclei	+1.812	-2.087	+1.333
Cohesiveness	+0.564	-0.640	-0.113
Cystic background	+1.565	-1.789	+1.465
Age group	+1.729	-1.606	+0.060
Papillary	-0.736	+1.157	n/a
Tubules	+2.711	-2.284	n/a

An AUC of 1.0 indicates a perfect test and an AUC below 0.75 is considered not to be clinically useful (19). The outcomes were established by correlating the histological specimen or obtaining clinical and radiological findings in the subsequent 2 years if histology was not available. The resulting AUCs were then compared with the AUCs from the training set to find any statistically significant differences

The algorithm derived from the ROC curve from the malignant statistical analysis of the training set was used to create a probability calculator.

Observations were entered into the calculator which then computed the probability. The calculated probability of malignancy for each C3 case was ordered from lowest to highest and was used to produce ROC curve. Lower and upper cut-points were established from the ROC curve.

The cut-points were selected at a value where the greatest number of cases was correctly classified as benign or malignant. The calculator was then tested against the validation set. The results of the probability calculator were ordered from lowest to highest probability and the cut-points applied. Those cases below the lower cut-point were down-graded into C2 and those above the upper cut-point were up-graded into C4. The re-graded cases were reviewed to identify probable reasons for under or over-interpretation. The probability calculator has the potential to remove cases wrongfully placed into the C3 category without compromising the integrity of these categories.

Results

The training set consisted of 180 C3 cases from which five statistically significant key cytological criteria were identified for the outcomes (malignant, benign proliferative and benign non-proliferative). Age was found to have significant predictive value and was therefore included with the key cytological criteria. The same criteria were applied to the validation set which comprised 182 subsequent cellular C3 FNA cases with follow-up. Figure1 displays the type of follow-up sourced to establish the pathological outcomes for both the training set and the validation set. Table 2 demonstrates the percentage breakdown for each pathological outcome in the training set and validation set.

The ROC curve analyses produced from the calculated probability of the combined selected key criteria and age group for all outcomes for both the training set and the validation set are shown in Figure 2. There were lower AUCs in the validation set for all

outcomes when compared to the training set, however, the differences were only
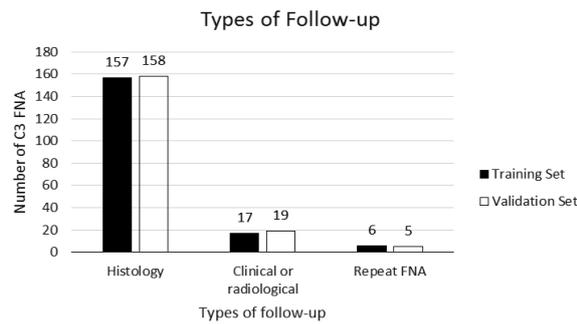

Figure 1 Type of follow up for the training set and the validation

statistically significant in the benign non-proliferative ROC curve analysis (see Table 3). The ROC curve produced from the malignant analysis was used to select lower and upper cut-points as shown in Figure 3. A lower cut-point was chosen at the probability value of 0.093. From this point 61/182 (33.5%) could be downgraded to C2. However, this included 4 encysted papillary cancers and 1 DCIS arising in a papilloma, which was incorrectly classified. An upper cut-point was chosen at the probability value of 0.75. Above this point 33/182 (18.1%) cases could be up-graded into a higher reporting category. Malignant histology was seen in 25 (76%) of the up-graded cases but 8 (24%) produced a proliferative non-malignant outcomes which included 1 fibroadenoma, 4 complex sclerosing lesions/radial scars, 1 papilloma, 1 sclerosing adenosis with epithelial hyperplasia and 1 clinically and radiologically stable case.

Discussion

The heterogeneity of pathological outcomes resulting from the atypical (C3) reporting category is wide and varied. Although C3 is a legitimate category, it is not very useful clinically and the diagnostic value of C3 is therefore limited because it contains a

Table 2 Comparison of the pathological outcomes from the training set and validation

	Training set (n=180)	Validation Set (n=182)
Benign non-proliferative	**24 (13.3%)**	**23 (12.6%)**
Benign – clinical or Radiologicalfollow up only	15	13
Benign specific (lactation changes, fat necrosis, scar, mastitis)	4	7
Non-proliferative FCC	5	3
Benign proliferative	**89 (49.4%)**	**109 (59.9%)**
Papilloma	44	50
CSL/RS/SA	9	22
FA	22	23
Proliferative FCC	8	9
Proliferative specific	6	5
Malignant	**67 (37.2%)**	**50 (27.5%)**
Ducal carcinoma in-situ	11	8
Invasive ductal carcinoma grade 1	31	22
Invasive ductal carcinoma grade 2	7	4
Invasive ductal carcinoma grade 3	9	3
Invasive lobular carcinoma	4	10
Cancer other (phyllodes, metastatic & unsuitable surgical candidates with linical/radiological diagnoses)	5	3

mix of benign and malignant entities (5, 9-12, 14, 20, 21). This grey zone causes diagnostic
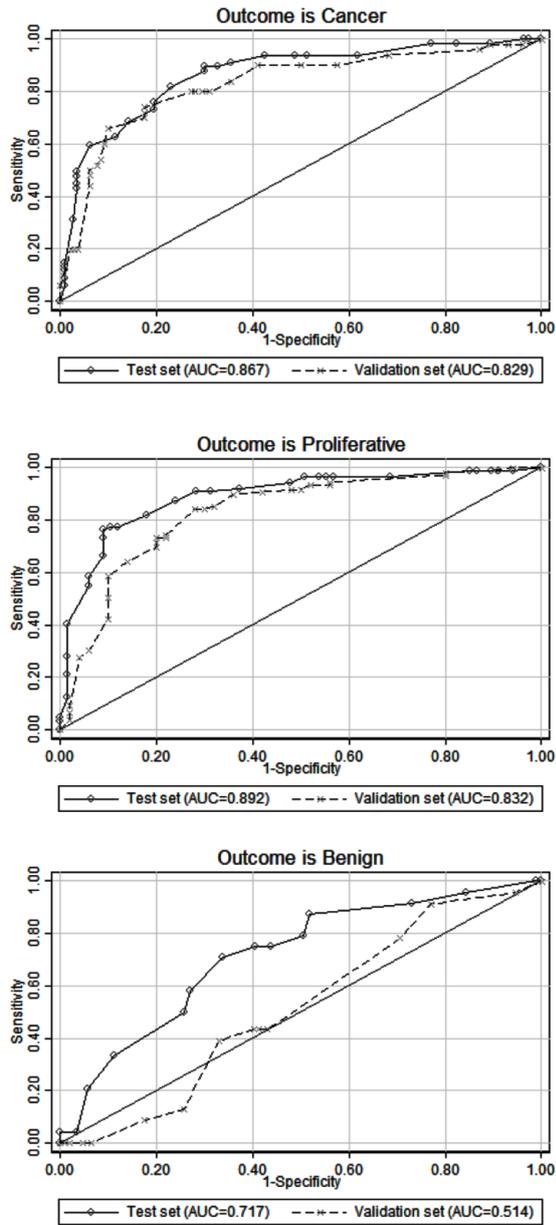

Figure 2. Comparison of the ROC curves with AUC for the test set and validation set for each of the pathological outcomes is presented.

dilemmas which result in unclear management directions. Therefore, the production of an evidence-based probability calculator for difficult breast FNA cases is a novel approach to stratifying the risk of malignancy when allocating FNA cases to a diagnostic category.

Our previous study of 230 C3 cases showed 34.8% of the C3 cases to be suboptimal due to low cellularity or poorly preserved material, contributing to their allocation into the C3 category. Previous statistical testing showed less reliable results when only limited material was available for assessment (16). By excluding this suboptimal group (less than 30 well- preserved epithelial groups per case) we could conduct a more meaningful statistical analysis. This enabled us to focus on the remainder 180 (65.2%) cellular but challenging C3 cases.

Our previous work (16) produced a set of statistically tested criteria with discriminating power for the stratification of atypical breast cytology. In this current study, we applied five key cytomorphological criteria and age to a validation set of closely matched cases. The comparison between the two sets of data showed similar predictive ability for all outcomes except for the benign non-proliferative conditions. We note that the validation set produced lower AUCs than the training set. The training set utilised the best predictive model for the stated pathological outcomes. This may have introduced a selection bias. By applying this best predictive model to the validation set, which was sourced from the subsequent C3 cases and hence contained a different mix of pathological outcomes than the training set, we eliminated this bias. The unbiased validation set results still produced clinically relevant results with the exception of the benign non-proliferative conditions. This analysis contained the least number of cases and mostly included clinically and radiologically stable but cytologically atypical cases without histological evaluation. The small number of cases and the inability to specifically identify histological outcomes may have impacted upon the statistical analysis producing low discretionary values. This

Table 3 Comparison of the calculated AUC from the outcomes of the training set and validation set (note statistically different analysis is shown in bold).

|		AUC	Standard error	95% CI
Malignant				
	Training set	0.8668	0.0285	0.81103-0.92256
	Validation Set	0.8289	0.0370	0.75644-0.90144
	Prob>chi2	0.4173		
Benign Proliferative				
	Training set	0.8918	0.0268	0.83925-0.94442
	Validation Set	0.8321	0.0369	0.75984-0.90438
	Prob>chi2	0.1903		
Benign Non–proliferative				
	Training set	0.7165	0.0576	0.60356-0.82950
	Validation Set	0.5140	0.0584	0.39941-0.62851
	Prob>chi2	**0.0136**		
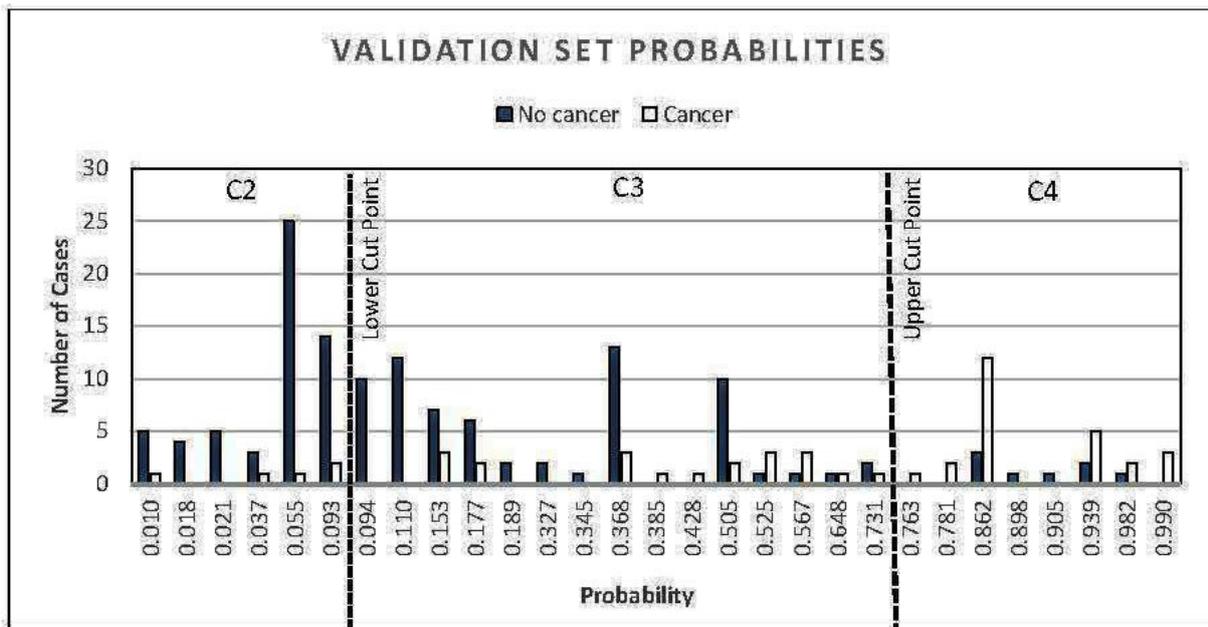
Figure 3. The calculated probabilities of the validation set with the lower and upper cut points in place, displaying the number of cases beyond the cut-points are shown

analysis was also affected by the absence of architectural features. Although the presence of papillary and tubular structures are highly predictive of papillary lesions and malignancy respectively, the benign non-proliferative cohort in the training set did not contain any of these features thus producing an errant

ROC curve when applied to the validation set as seen in Figure 2.

Application of the probability calculator produced from the malignant analysis down-graded 61/182 (33.5%) C3 to C2 and up-graded 33/182 cases (18%) from C3 to C4 in the validation set. The rationale behind

choosing the lower cut-point was to find a point which included as many benign outcomes below the selected probability value but not contain false negative cases. Similarly, the upper cut-point needed to identify malignant lesions beyond the upper probability value but limit those cases of high probability but with non-malignant outcomes. These cut-points met with limited success.

When considering the cases below the lower cut-point, 5 cases from 61 (8.2%) were found to be malignant, including 4 encysted papillary carcinomas and 1 DCIS associated with a papilloma. Clearly, the calculator under-interpreted these abnormal papillary lesions, highlighting a failure of the calculator. The reason for low probability can be attributed to the presence of papillary fragments in all of these cases. This criterion has a protective effect when predicting cancer, thereby reducing the probability of cancer. However, in these 5 cases, identification of a papillary lesion was cytologically possible due to the presence of papillary fragments and discerning cytologists could override the calculator findings. This inherent failure of the calculator suggests a lower cut-point should not be used due to the risk of wrongful allocation of malignant cases into the benign category. These cases should all remain in the C3 category with specific mention of papillary features.

Of the 33 cases found beyond the upper cut-point, 25 (75.8%) resulted in malignancy however, 8 (24.2%) were proliferative but not malignant. Further analysis of this group revealed 1 fibroadenoma, 4 complex sclerosing lesions/radial scars, 1 papilloma, 1 sclerosing adenosis with epithelial hyperplasia and 1 clinically and radiologically stable case. These overcalls are understandable as the above cases can be challenging both histologically and cytologically. Placement into the C4 (suspicious) category for these non-malignant lesions would not have changed

treatment, as all these lesions in our institution undergo further histological evaluation before definitive treatment.

Five of the eight over-interpreted cases contained tubules which would have significantly increased their probability of cancer. Tubular structures were an uncommon finding and highly predictive of malignancy as shown by our previous study (16). The tubular structures were found in the fibroadenoma, the clinically and radiologically stable case, one case of sclerosing adenosis and 2 of the complex sclerosing lesions/radial scars. On review, the tubular structures were misinterpreted. We defined these structures as angulated tubular structures with parallel sides, sharp ends and no obvious myoepithelial cells. Shabb et. al. and Simsir et. al. note the finding of tubular structures should be interpreted with caution as they can also be found in fibroadenomas, sclerosing adenosis and complex sclerosing lesions but are more frequently observed in tubular carcinoma (9, 22). Kundu also warns against over- interpreting tubular structures as they can mimic malignancy in cases of sclerosing adenosis but states that these should also contain myoepithelial cells. Tubular carcinoma usually has greater cellularity with more rigid or acutely angled configurations when compared to sclerosing adenosis(23). These observations have also been reported by Bozanini et. al. (24). The review of the specific cases containing tubules, showed round rather than sharp ends and obvious myoepithelial cells. Figure 4 illustrates some of the misinterpreted tube-like structures seen in fibroadenomas and complex sclerosing lesions. The failure of the calculator in this instance is due to human interpretative error.

The other 3 non-malignant cases found beyond the upper cut-point all displayed lack of benign features including cohesion, myoepithelial cells or bare bipolar nuclei and

cystic background. The absence of these benign features produced higher probability hence their up-grade to C4. These included 2 complex sclerosing lesions and 1 papilloma. Orell when examining false positives caused by complex sclerosing lesions, comments that these cases often contain some proportion of co-existing benign epithelium and bare bipolar nuclei (25). Mak and Field also note the importance of not over-diagnosing cytological material as malignant in cases of complex sclerosing lesions/radial scars (26). Simsir & Cangiarella, Saad et. al. and Agnihotri et. al. all report that papillomas can display loss of cohesion and this may contribute to false positive diagnoses (22, 27, 28). Certainly, the combination of diminished numbers of myoepithelial cells or bare bipolar nuclei in the background, loss of cohesion and absence of a cystic background amounts to a higher probabilistic malignant conclusion.

On review of these three non-malignant cases, poor presentation of the cytological material was evident including distortion and air-drying artefact, contributing to the misinterpretation of the lesions thereby limiting the use of the calculator.

This statistically derived probability calculator was created to enhance decision making when interpreting difficult breast FNA and to reduce sources of bias(29, 30).The methodology used for this project was intended to reduce subjectivity by providing an evidence based tool which could calculate the risk of malignancy of the combined predictive criteria thereby assist in allocating a lesion into a suitable diagnostic category. However, this calculator has limitations, as it focuses only on the microscopic features and does not include the clinical and imaging findings of the triple test. Although some advancement in diagnosis can be made by using the calculator in some instances, it is not a substitute for experience and teamwork. The atypical category by its very nature is a

subjective category with biologically variable lesions and often limited by poor sampling and pre-analytical issues.

Overall, the calculator was found to be of limited value in re-assigning cases to either lower or higher categories. However, further evaluation of the calculator for inter-observer and intra-observer agreement, is required. The probability calculator may be of value in cytologically difficult cases but other diagnostic modalities such core biopsy or surgical excision is likely to be required in these circumstances.

In summary, the atypical category (C3) breast FNA cytology encompasses histologically benign and malignant conditions. Using an algorithm based upon five key cytomorphological criteria and age, developed from a training set of 180 cases and applied to a validation set of 182 subsequent C3 cases, we were able to predict malignant and benign proliferative outcomes with reasonable accuracy. Benign non-proliferative changes could not be reliably diagnosed using the selected criteria.

The ability of the probability calculator to aid in the allocation of cases into the C3 reporting category was limited. Cut-points based on the probability of malignancy were chosen to re-grade cases from C3 into lower or higher categories. At the lower cut-point, the calculator re-allocated 61 (33.5%) cases into C2, however 5 (8.2%) papillary malignancies were under-interpreted (false-negatives), highlighting the limitation of the calculator,

which was considered unacceptable. Upgrading C3 to a higher category (C4) yielded 33/182 (18.1%) of the cases. However, this up-graded group contained 8 benign proliferative lesions (24.2%), signifying over interpretation and wrongful allocation into C4.

The C3 category remains a heterogeneous mix of pathological lesions. Attempts to minimise inappropriate allocation of C3 cases into this category have met with limited success.
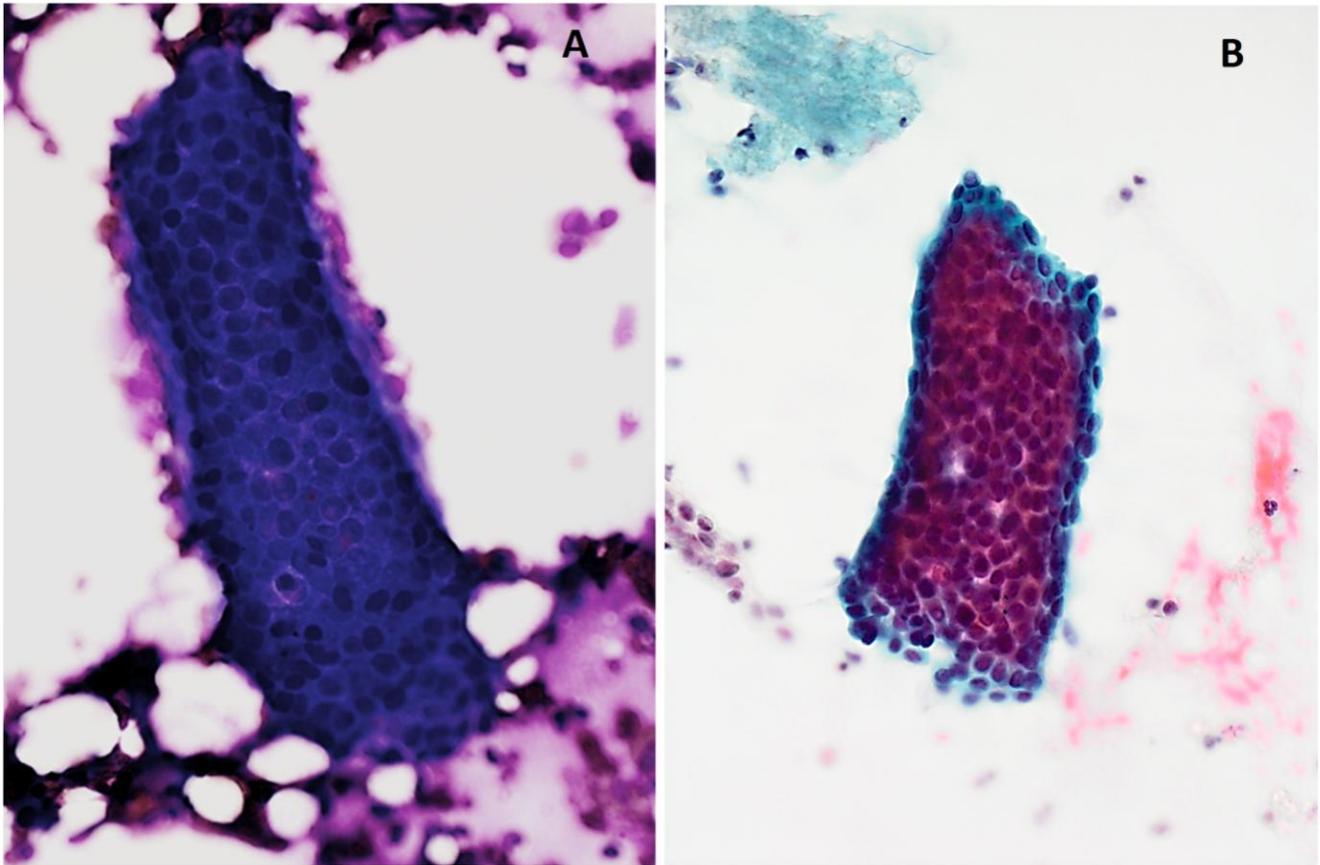

Figure 4. Illustration of some misidentified tube like structures. Panel A depicts a tubular structure from a fibroadenoma. Note the rounded ends and overlying myoepithelial cells (QD stain, original magnification x60). Panel B is from a complex sclerosing lesion and shows a rigid type tubular structure with elongated peripheral nuclei and sharp ends but obvious overlying myoepithelial cells (Papanicolaou stain, original magnification x60).

References

1. National Breast and Ovarian Cancer Centre Breast fine needle aspiration cytology and core biospy: a guide for practice. In. Edited by Department of Health and Ageing, First edn. Camperdown, NSW: Australian Federal Government; 2004.
2. National Cancer Institute Fine-Needle Aspiration of Breast Workshop S: The uniform approach to breast fine-needle aspiration biopsy. *Diagnostic Cytopathology* 1997, 16(4):295-311

3. National Health Service Breast Screening Programme: Guidelines for Non-operative diagnostic procedures and reporting in breast cancer screening. In. Edited by Programme NHSCS, vol. NHSBSP Publication No 50. Sheffield; 2001.

4. Moyes C, Dunne B. Predictive power of cytomorphological features in equivocal (C3, C4) breast FNAC. *Cytopathology* 2004; 15(6):305-310.

5. Deb RA, Matthews P, Elston CW, Ellis IO, Pinder SE. An audit of 'equivocal' (C3) and 'suspicious' (C4) categories in fine needle aspiration cytology of the breast. *Cytopathology* 2001; 12(4):219-226.

6. Mander BJ, Beresford PA, Tildsley G, Qureshi T, Wishart GC. Management of patients with intermediate (C3) cytology and a solitary breast lump. *The Breast* 2001; 10(2):163-165.

7. Kocjan G, Bourgain C, Fassina A, Hagmar B, Herbert A, Kapila K, Kardum-Skelin I, Kloboves-Prevodnik V, Krishnamurthy S, Koutselini H *et al*. The role of breast FNAC in diagnosis and clinical management: a survey of current practice. *Cytopathology* 2008; 19(5):271-278.

8. Pambuccian SE. What is atypia? Use, misuse and overuse of the term atypia in diagnostic cytopathology. *Journal of the American Society of Cytopathology* 2015; 4(1):44-52.

9. Shabb NS, Boulos FI, Abdul-Karim FW. Indeterminate and Erroneous Fine-Needle Aspirates of Breast with Focus on the 'True Gray Zone': A Review. *Acta cytologica* 2013; 57(4):316-331.

10. Tran PV, Lui PC, Yu AM, Vinh PT, Chau HH, Ma TK, Tan P, Tse GM. Atypia in fine needle aspirates of breast lesions. *Journal of Clinical Pathology* 2010; 63(7):585-591.

11. Kanhoush R, Jorda M, Gomez-Fernandez C, Wang H, Mirzabeig M, Ghorab Z, Ganjei-Azar P. Atypical and Suspicious diagnoses in breast aspiration cytology. *Cancer Cytopathology* 2004; 102(3):164-167.

12. Lim JC, Al-Masri H, Salhadar A, Xie HB, Gabram S, Wojcik EM. The significance of the diagnosis of atypia in breast fine-needle aspiration. *Diagnostic Cytopathology* 2004; 31(5):285-288.

13. Özkara SK, Üstün MÖ, Paksoy N. The gray zone in breast fine needle aspiration cytology: How to report on it? *Acta cytologica* 2002; 46(3):513-518.

14. al-Kaisi N. The spectrum of the "gray zone" in breast cytology. A review of 186 cases of atypical and suspicious cytology. *Acta cytologica* 1994; 38(6):898-908.

15. Weigner. J, Zardawi. I, Braye. S. The True Nature of Atypical Breast Cytology. *Acta cytologica* 2013; 57(5):464-472.

16. Weigner. J, Zardawi. I, Braye. S, McElduff. P. The Microscopic Complexities of C3 in Breast Cytology. *Acta cytologica* 2014; 58(4):335-346.

17. StataCorp: Stata Release II. In: *Statistical Software College Station, TX.* StataCorp; 2009.

18. Weigner J., Zardawi I., Braye S., McElduff P. The Conundrum of Papillary Breast Lesions within the C3 Category. *Acta cytologica* 2015; 59(4):289-297.

19. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *CJEM : Journal of the Canadian Association of Emergency Physicians* 2006; 8(1):19-20.

20. Bofin AM, Lydersen S, Isaksen C, Hagmar BM. Interpretation of fine needle aspiration cytology of the breast: a comparison of cytological, frozen section, and final histological diagnoses. *Cytopathology* 2004; 15(6):297-304.

21. Chaiwun B, Sukhamwang N, Lekawanvijit S, Sukapan K, Rangdaeng S, Muttarak M, Thorner PS. Atypical and suspicious categories in fine needle aspiration cytology of the breast: histological and mammographical correlation and clinical significance. *Singapore medical journal* 2005; 46(12):706-709.

22. Simsir A, Cangiarella J. Challenging breast lesions: Pitfalls and limitations of fine-needle aspiration and the role of core biopsy in specific lesions. *Diagnostic Cytopathology* 2012; 40(3):262-272.

23. Kundu UR, Guo M, Landon G, Wu Y, Sneige N, Gong Y. Fine-needle aspiration cytology of sclerosing adenosis of the breast: a retrospective review of cytologic features in conjunction with corresponding histologic features and radiologic findings. *American journal of clinical pathology* 2012; 138(1):96-102.

24. Bonzanini M, Gilioli E, Brancato B, Pellegrini M, Mauri MF, Dalla Palma P. Cytologic features of 22 radial scar/complex sclerosing lesions of the breast, three of which associated with carcinoma: Clinical, mammographic, and histologic correlation. *Diagnostic Cytopathology* 1997; 17(5):353-362.

25. Orell SR. Radial scar/complex sclerosing lesion—a problem in the diagnostic work-up of screen-detected breast lesions. *Cytopathology* 1999; 10(4):250-258.

26. Mak A, Field AS. Positive predictive value of the breast FNAB diagnoses of epithelial hyperplasia with atypia, papilloma, and radial scar. *Diagnostic Cytopathology* 2006; 34(12):818-823.

27. Saad RS, Kanbour-Shakir A, Syed A, Kanbour A. Sclerosing papillary lesion of the breast: A diagnostic pitfall for malignancy in fine needle aspiration biopsy. *Diagnostic Cytopathology* 2006; 34(2):114-118.

28. Agnihotri M, Naik L, Kothari K, Fernandes G, Ojha S. Fine-Needle Aspiration Cytology of Breast Lesions with Spontaneous Infarction: A Five-Year Study. *Acta cytologica* 2013; 57(4):413-417.

29. Schmidt RL, Factor RE, Witt BL, Layfield LJ. Quality Appraisal of Diagnostic Accuracy Studies in Fine-Needle Aspiration Cytology. *Archives of Pathology & Laboratory Medicine* 2013; 137(4):566-575.

30. Schmidt RL, Factor RE. Understanding Sources of Bias in Diagnostic Accuracy Studies. *Archives of Pathology & Laboratory Medicine* 2013; 137(4):558-565.