Can Fine Needle Aspiration Cytology be Used as a “Proxy Gold Standard” to Diagnose Tuberculous Mastitis?

Meherbano M. Kamal, Hemant R. Kulkarni, Manjiri M. Makde, Radha Munje

Department of Pathology, Government Medical College, 2Department of Tuberculosis and Chest Diseases, Indira Gandhi Government Medical College, Nagpur, Maharashtra, India, 1M and H Research, LLC, San Antonio, Texas, USA

Abstract

Objective: To assess the performance of fine needle aspiration cytology (FNAC) in the diagnosis of tuberculosis mastitis. Materials and Methods: Diagnostic test performance evaluation using two methods—as compared to an alloyed gold standard as well as in the absence of a gold standard. Alloyed gold standard combined the results of acid fast bacilli in cytology smears, histopathological confirmation, and response to treatment. Bayesian estimation of test parameters was done in the absence of the gold standard. Results: FNAC was carried out in 6,496 consecutive cases of breast lump and 104 cases of granulomatous mastitis were detected. Both methods of test parameter estimation identified a high specificity of FNAC for the diagnosis of tuberculosis mastitis (98.9% and 98.4%, respectively). Estimation of sensitivity was falsely high (100%) using the alloyed gold standard because of a workup bias and falsely low (8.41%) using the Bayesian estimation because of low prevalence. Likelihood ratios by both methods suggested that FNAC has good discriminatory capability. Conclusion: In situations where prevalence of tuberculosis is high and where facilities for histopathological evaluation do not exist, FNAC can offer an optional alternative to base the therapeutic decision for starting antitubercular treatment.

Keywords: Bayesian estimation, FNAC, gold standard, granulomatous mastitis, tuberculosis breast

Introduction

Tuberculosis of the breast is considered to be a rare condition in surgical practice with reported prevalence rates varying from 0.5% to 4.5% of all breast lesions.1-8 This is true even for countries where tuberculosis is still a common disease.3-6,9 Among others, the factors contributing to this rarity include underdiagnosis,10 under-responding,11 and insensitive methods to demonstrate the presence of acid fast bacilli.12 Histopathology is considered diagnostic for tuberculous mastitis only if the acid fast bacilli can be demonstrated.12,13 Moreover, the restrictive indications and costs make this investigation infeasible in most cases.13 In situations of relatively high prevalence of tuberculosis it becomes imperative to search for a more optimal alternative.

Currently, the diagnostic approach to a breast lump includes the use of fine needle aspiration cytology (FNAC) in addition to the detailed clinical examination and mammography. However, it is still needed to demonstrate the presence of acid fast bacilli to label tuberculosis. It is well established that FNAC is a rapid and reasonably inexpensive investigation.6,14-17 However, its diagnostic performance in tuberculous mastitis is not precisely known owing to the problems in gold standard.

We conducted this study to assess if FNAC can be used as the next best method to diagnose tuberculous mastitis in the absence of a gold standard method. Specifically, we were interested in quantitating the performance of FNAC to diagnose tuberculous mastitis in patients reporting with a breast lump.

Materials and Methods

All incident cases of lump in breast reporting to or referred to the study center were included in the present study. All the cases of lump in breast underwent FNAC for the diagnosis of tuberculosis breast. To document the presence of tuberculosis, data on other investigation like acid fast bacilli, histopathology, montoux test, and mass miniature were used as and when available.

Address for correspondence: Dr. Manjiri M. Makde, 206/6 “Vasanti,” Near GPO Square, Civil Lines, Nagpur, Maharashtra, India. E-mail: majiri0288@gmail.com

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Kamal MM, Kulkarni HR, Makde MM, Munje R. Can fine needle aspiration cytology be used as a “Proxy Gold Standard” to Diagnose tuberculous mastitis? J Cytol 2018;35:159-62.
FNA was carried out by using a 10 ml syringe attached to a Cameco’s syringe holder. Aspiration was repeated in doubtful cases or when there was a paucity of the aspirate. As a protocol, any associated lymph nodes were also aspirated for cytological examination. Cytologically, a case was considered positive for tuberculous mastitis when the smears showed epitheloid granulomas, acute or chronic inflammatory cells, and Langhans type giant cells with or without caseous necrosis. For evaluating the performance of FNA in the diagnosis of tuberculosis of breast two different approaches were used. First, we defined an alloyed gold standard for diagnosis of tuberculosis. Using this strategy, a case was defined as having tuberculosis if one or more of the following features were observed: smears positive for acid fast bacilli, histopathological features diagnostic of tuberculosis and response to antitubercular drug therapy. Patient was monitored clinically every 2 months till 18 months. Progressive regression in the size of the breast lump and axillary lymph node whenever present was considered as response to antitubercular treatment. The performance of FNA was then compared with this alloyed gold standard using the standard two by two contingency table analysis. In the second approach for evaluation of the diagnostic performance, we used the method of Bayesian estimation of test parameters in the absence of a gold standard as suggested by Joseph et al.\[18]\). Bayesian estimation of the test parameters require a priori assumptions about the distribution of each expected parameter. We derived these estimates from the results obtained using the alloyed gold standard. We used the Perl program BayesDiagnosticTests.pl (available at http://www.medicine.mcgill.ca/epidemiology/joseph/Bayesian-Software-Diagnostic-Testing.html) to implement the Bayesian analyses. The program generated 20,000 random samples based on the distributional assumptions and then estimated the mean and interquartile range (IQR) of the test characteristics. Using both the approaches we estimated the following parameters for FNA: sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratio of the positive test, and likelihood ratio of negative test. We also estimated the prevalence of tuberculosis in our study population by both the approaches. Finally, we calculated the 95% confidence intervals around all the parameters.

**RESULTS**

A total of 6,496 consecutive cases of lump in breast reporting to the study center were recruited into the study. Of these, evidence of granulomatous mastitis by FNAC was observed in 104 cases (1.6%). There were three cases who had concomitant pulmonary tuberculosis. Two had taken treatment and one was a defaulter. Data on testing for acid fast bacilli were available in 33 (31.73%) of these cases whereas histopathological evaluation was done in 17 cases (16.35%). Both these investigations detected three cases each as positive for tuberculosis. Antitubercular treatment was started to all the 104 cases, however, complete response to treatment could be demonstrated in 30 cases only (28.85%). Using the alloyed gold standard, therefore, presence of tuberculosis was seen in 36 cases only (34.62% of the total number of cases positive for granulomatous mastitis by FNAC). Table 1 classifies the complete dataset by the tuberculosis status and needle aspiration results.

It can be observed from Table 2 that the use of alloyed gold standard for estimating the test parameters was associated with very high sensitivity and specificity values, low positive predictive value, and high negative predictive value. The likelihood ratios for positive and negative tests also were convincing.

From these results, we made following prior assumptions regarding the α and β distributional parameters used by the BayesDiagnosticTests program: prevalence: 0.95 and 37.05, respectively; sensitivity and specificity: 17.1 and 0.9, respectively. Results of the Bayesian estimation of test parameters, fully corroborated those obtained using the alloyed gold standard [Table 3]. Specifically, the results indicated a very high sensitivity and specificity of FNAC to detect tuberculous mastitis in spite of an estimated low prevalence (0.8%) of this condition in breast lumps.

**DISCUSSION**

Tuberculous mastitis is difficult to diagnose.\[19\] As a result, we had to devise an alloyed gold standard for comparison of the performance of FNAC in diagnosis of tuberculous mastitis. However, there can be several problems with the alloyed gold standard that we used. First, as already mentioned, the difficulties in demonstration of acid fast bacilli in smears

| Table 1: Distribution of the study participants according to disease status and FNAC results |
|-----------------------------------------------|
| Tuberculosis of breast | Total |
| Proved | Not proved |
| FNAC | |
| Positive | 36 | 68 | 104 |
| Negative | 0 | 6,392 | 6,392 |
| Total | 36 | 6,460 | 6,496 |

| Table 2: Diagnostic performance of FNAC using gold standard comparison and Bayesian estimation |
|-----------------------------------------------|
| Parameter | Gold standard comparison | Bayesian estimate |
|-----------------------------------------------|
| Estimate | 95% CI | Median | IQR |
| Prevalence | 0.0055 | 0.0037-0.0073 | 0.0080 | 0.0000-0.0178 |
| Sensitivity | 1 | - | 0.9641 | 0.8083-0.9999 |
| Specificity | 0.9895 | 0.9870-0.9920 | 0.9915 | 0.9833-0.9997 |
| PPV\(^1\) | 0.3462 | 0.2547-0.4376 | 0.4727 | 0.0212-0.9786 |
| NPV\(^2\) | 0.9999 | 0.9997-1.0000 | 0.9998 | 0.9977-1.0000 |
| LR\(^+\) | 93.7 | 73.7-119.1 | 111.1 | 55.4-2772.0 |
| LR\(^-\) | 0.01 | 0.001-0.022 | 0.04 | 0.001-0.19 |

\(^1\)Positive predictive value; \(^2\)negative predictive value; \(^+\)likelihood ratio of a positive test; \(^-\)likelihood ratio of a negative test
are well documented. Second and consequently, the same difficulties can also be experienced if histopathology is used to demonstrate the acid fast bacilli. Third, the strategy of using response to antitubercular treatment as a confirmation of existing tuberculous disease is, essentially, empirical. Fourth, all the three methods—acid fast bacilli in smears, histopathological confirmation, and response to treatment—are very specific to tuberculosis but rather insensitive. Fifth, the alloyed gold standard—as it were—does not fulfill all of the Koch’s postulates that are, indeed, necessary for diagnosing tuberculosis. Although our strategy of combining the diagnostic methods by the OR logical operator might slightly alleviate the problem of insensitivity of the individual components, the other limitations to the alloyed gold standard remain. There are at least two ways in which we could improve upon this alloyed gold standard.

First, we could have used the information contained in clinical examination into the final diagnosis making the diagnostic strategy more sensitive. For our dataset, however, this strategy did not have a great improvement in the sensitivity. Table 3 summarizes the provisional clinical diagnosis in relation to the diagnosis of tuberculous mastitis in the 104 cases diagnosed to have granulomatous mastitis by FNAC. It can be seen from Table 3 that the positivity rates for tuberculous mastitis were not significantly different for the various provisional diagnosis. Moreover, there were 24 cases (23%) in whom a provisional diagnosis was not specified. Using the provisional clinical diagnosis as a summary measure of the clinical information could not, thus have improved the performance of alloyed gold standard was independent of provisional clinical diagnosis.

The other strategy to counter the problem of gold standard was to handle it statistically. Various statistical techniques are available to estimate the test parameters in the absence of a gold standard. For example, the latent class analysis estimates the test sensitivities and specificities along with prevalence of the disease. However, this analysis needs simultaneous information on all the tests for all the individuals. As already stated our data were limited in that respect. In fact, there were only three cases of the 104 granulomatous mastitis cases where information on acid fast bacilli in smears, histopathology, and response to treatment was simultaneously available. Obviously these three cases do not provide enough degrees of freedom for using the latent class analysis. Also, the latent class analysis cannot be used if only two tests are being compared to each other without making at least two assumptions out of the five parameters being estimated (two sensitivities, two specificities, and prevalence). Therefore, we could not use latent class analysis.

We used the Bayesian estimation procedure for the one test situation. Here, the assumption is that we have results of only one test on all the study subjects (in our case, the results of FNAC on all the study subjects). Given some starting values, the estimation program then proceeds iteratively, using the Gibbs sampling algorithm to come up with a stable solution for the three parameters—sensitivity, specificity, and prevalence. The ease of interpretation of results of the Bayesian estimation of test parameters comes at the cost of distributional assumptions of these parameters. While interpreting the results, therefore, one must bear in mind these assumptions.

Our results indicated that comparison with the alloyed gold standard and the Bayesian estimation procedure gave very similar results despite the fact that there was a workup bias in the study design. According to the clinical protocol in the study center, all cases diagnosed to have granulomatous mastitis are given a trial of antitubercular treatment whereas those negative for granulomatous mastitis do not. Consequently, the cell in Table 1 that represents those who have tuberculosis (using the alloyed gold standard) but who have a negative FNAC result will always be zero. When using the Bayesian estimation for sensitivity one needs to consider another limitation of the method of estimation. It is known that if the number of test positives is very low, then the data contain more information on specificity and less information on sensitivity. As a result, the Bayesian estimation could result in an underestimation of the test sensitivity. For our dataset it appears that there was overestimation of sensitivity using the alloyed gold standard because of the workup bias whereas there was an underestimation of sensitivity using Bayesian estimation because of low prevalence. The specificity was very high indicating that a negative FNAC can be used as a very good rule-out criterion. Finally, the likelihood ratios (using both statistical approaches) showed that the FNAC is a very good discriminatory test.

In developing countries like India, the problem of tuberculosis is compounded by the burden of disease as well as availability of diagnostic facilities. As already stated, the clinical protocol at the study center is to start a trial of the antitubercular drugs should a case be positive for granulomatous mastitis by FNAC. From the therapeutic perspective at least it seems that the decision to treat can solely be based on the results of

| Provisional diagnosis         | Tuberculosis proved (n) | Tuberculosis not proved (n) | Total (n) | % Positives |
|-------------------------------|-------------------------|-----------------------------|-----------|-------------|
| Benign breast lesion          | 11                      | 21                          | 32        | 34.38       |
| Pyogenic breast abscess       | 7                       | 17                          | 24        | 29.17       |
| Carcinoma of breast          | 5                       | 12                          | 17        | 29.41       |
| Tuberculosis of breast        | 2                       | 9                           | 11        | 18.18       |
| Not specified                 | 13                      | 11                          | 24        | 54.17       |

*Total number of diagnoses are 108 because four cases had two provisional diagnoses each*
FNAC. Given the high specificity of FNAC it seems that an over treatment of tuberculosis can result in only 1–1.5% of nontuberculous mastitis.

**Conclusion**

Where facilities for increasing the yield of acid fast bacilli in smears by cyto-centrifugation and sedimentation[24] and histopathology are inadequate and where the prevalence of tuberculosis infection is high, it seems reasonable to base the therapeutic decisions on the results of FNAC alone. Within the constraints of limitations, therefore, we recommend that FNAC can be used as a proxy gold standard in situations of high prevalence of tuberculosis infection and nonavailability of diagnostic histopathology for tuberculous mastitis.

**Financial support and sponsorship**

Nil.

**Conflicts of interest**

There are no conflicts of interest.

**References**

1. Morgan M. Tuberculosis of the breast. Surg Gynecol Obstet 1931;53:593-605.
2. Miller RE, Solomon PF, West JP. The co-existence of carcinoma and tuberculosis of the breast and axillary lymph nodes. Am J Surg 1971;121:335-40.
3. Dubey MM, Agarwal S. Tuberculosis of the breast. J Indian Med Assoc 1968;51:358-9.
4. Alagaratnam TT, Ong GB. Tuberculosis of the breast. Br J Surg 1980;67:125-6.
5. Banerjee SN, Ananthakrishnan N, Mehta RB, Prakash S. Tuberculous mastitis: A continuing problem. World J Surg 1987;11:105-9.
6. Gupta D, Rajiwanshi A, Gupta SK, Nijhawan R, Saran RK, Singh R. Fine needle aspiration cytology in the diagnosis of tuberculous mastitis. Acta Cytol 1999;43:191-4.
7. Akcakaya A, Eryilmaz R, Sahin M, Ozkan O. Tuberculosis of the breast. Breast J 2005;11:85-6.
8. Tewari M, Shukla HS. Breast tuberculosis: Diagnosis, clinical features and management. Indian J Med Res 2005;122:103-10.
9. Manoj DK, Smitha Nair, Rajani M, Kumar C. Breast TB: Clinical Profile and Treatment Outcome. Int J Med Res Prot 2016;2:200-3.
10. Helmer M, Pokieser L, Salomonowitz E. Tuberculosis of the female breast: Diagnostic clarification. Rontgenblatter 1986;39:357-9.
11. Dent DM, Webber BL. Tuberculosis of the breast. S Afr Med J 1977;51:611-4.
12. Hale JA, Peters GN, Cheek JH. Tuberculosis of the breast: Rare but still extant. Review of the literature and report of an additional case. Am J Surg 1985;150:620-4.
13. Wilson TS, MacGregor JW. The diagnosis and treatment of tuberculosis of the breast. Can Med Assoc J 1963;89:1118-24.
14. Nayar M, Saxena HMK. Tuberculosis of the breast: A cytomorphic study of needle aspirates and nipple discharges. Acta Cytol 1984;28:325-8.
15. Kumarasinghe MP. Cytology of granulomatous mastitis. Acta Cytol 1997;41:727-30.
16. Sharma AK, Sree S, Mishra SK. Tubercular mastitis: A pragmatic approach to its management. Aust N Z J Surg 1993;63:263-5.
17. Jayaram G. Cytomorphology of tuberculous mastitis: A report of nine cases with fine needle aspiration cytology. Acta Cytol 1985;29:974-8.
18. Joseph L, Gyoroks TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of gold standard. Am J Epidemiol 1995;141:263-72.
19. Talantov VA. Difficulties in the differential diagnosis of tuberculosis of the breast. Vestn Khir 1979;123:15-8.
20. Van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, De Groot JA. Latent Class Models in Diagnostic Studies When There is No Reference Standard—A Systematic Review. Am J Epidemiol 2014;179:423-31.
21. Walter SD, Irwig LM. Estimation of test error rates, disease prevalence and relative risk from misclassified data: A review. J Clin Epidemiol 1988;41:923-37.
22. Ranshoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Eng J Med 1978;299:926-30.
23. De Groot JAH, Dendukuri N, Janssen KJM, Reitsma JB, Brophy J, Joseph L, et al. Adjusting for partial verification or workup bias in meta-analyses of diagnostic accuracy studies. Am J Epidemiol 2012;175:947-53.
24. Fodor T. Detection of mycobacteria in sputum smears prepared by cyto centrifugation and sedimentation. Tubercle Lung Dis 1995;76:273-4.