

ORIGINAL ARTICLE

Automated classification of cancer from fine needle aspiration cytological image use neural networks: A meta-analysis

Jian Huang MS  | Dongcun Wang MD | Jiping Da MD

Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital and Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China

Correspondence

Jiping Da, No. 113 Baohe Road, Longgang District, Shenzhen 518116, China.
Email: jiping_da@aliyun.com;
m13524885941@163.com

Funding information

Supported by Sanming Project of Medicine in Shenzhen, Grant/Award Number: SZSM20181207

Abstract

Background: The role of retrospective analysis has been evolved greatly in cancer research. We undertook this meta-analysis to evaluate the diagnostic value of Neural networks (NNs) in Fine needle aspiration cytological (FNAC) image of cancer.

Methods: We systematically retrieved 396 literatures on cytodagnosis of NNs from Cochrane, PubMed, and EMBASE. After screening, only six studies were included in meta-analysis finally. Data was comprehensively analyzed by RevMan and meta-Disc software.

Results: A total of 1165 cases were extracted from six articles. Among them, 593 cases were in the abnormal/positive group and 572 cases in the normal/negative group. The pooled estimates for the NNs cytology were Area under ROC curve (AUC): 0.99, Sensitivity: 0.85 (95% CI:0.82-0.88), Specificity: 0.96 (95% CI:0.94-0.97), Positive Likelihood Ratio (LR):18.43 (95% CI:6.83-49.74), Negative Likelihood Ratio (LR): 0.06 (95% CI:0.001-0.58), and Diagnostic odds ratio (DOR): 343.21 (34.41-3422.77).

Conclusions: This meta-analysis confirms that NNs Automated Classification algorithm can facilitate to some extent the FNCA diagnosis of cancer.

KEYWORDS

FNCA, meta-analysis, neural networks

1 | INTRODUCTION

Fine needle aspiration cytological (FNAC) is a quick, cheap, minimally invasive and widely available method of examination of cells collected from patients that can characterize the majority of lesions, the effective use of FNAC has helped us to reduce many unwanted surgeries.¹ In practice, The FNAC pathology report was signed out by a cytopathologist after reviewing slices/smears microscopically. However, the interpretation of report is an elaborating and time consuming task, even for a senior cytopathologist, which requires the visual inspection and evaluation of subtle change in nuclei morphology,^{1,2} this is a great challenge.

Artificial intelligence (AI) has experienced the shallow learning represented by k-nearest neighbor algorithm (KNN) and support vector machine (SVM) in 1950s, until the deep learning represented by neural

network (NN) algorithm at now. NNs, also known as Artificial Neural Networks (ANNs), is mathematical algorithm Model that simulates the behavior characteristics of animal NNs and carries out distributed parallel information processing. Compared with single-layer shallow learning, the model structure of NNs contains multi-layer hidden layer network, which focuses on the characteristics of feature level. In recent years, ANNs-powered devices have transformed our daily life, for example: smartphones, self-driving cars, and intelligent home appliances. Similarly, ANN has made significant progress in image analysis and object identification.^{3,4} More and more cytopathologists try use NNs to assist cytological diagnosis, and have been getting brilliant advances,⁵ In view of most of these studies was from small samples, and its reliability of results was not consistent, further research is necessary. Meta-analysis is kind of statistical method for quantitative synthesis of many studies on the same subject with

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Diagnostic Cytopathology* published by Wiley Periodicals LLC.

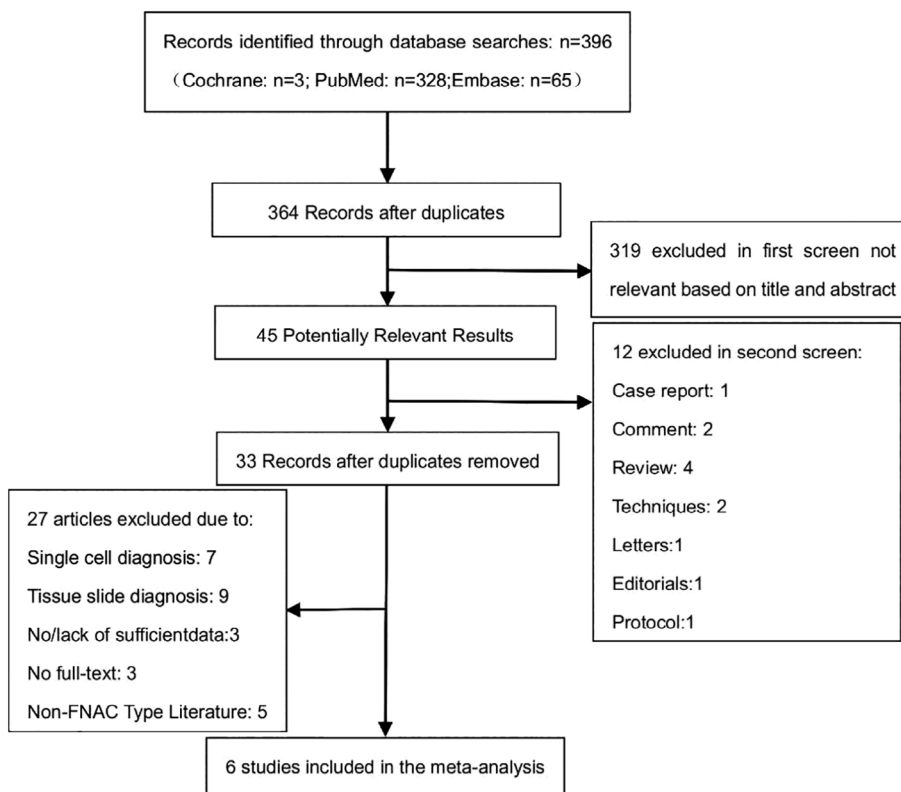


FIGURE 1 Flow diagram for selection of studies included in this meta-analysis

TABLE 1 Characteristics of all studies included in the meta-analysis

Author/ published time	Tumor type	Sample type	Diagnostic procedures	TP	FP	FN	TN	Total
Anagnostopoulos, 2006	Benign/breast cancer	FNAC	DiagNN	209	9	3	348	569
Dey, 2013	Benign/breast cancer	FNAC	ANN	40	0	4	20	64
Momeni-Boroujeni, 2017	Benign/pancreatic exocrine carcinoma	FNAC	MNN	32	4	3	36	75
Savala, 2018	Follicular adenoma (FA)/follicular carcinoma (FC)	FNAC	ANN	30	1	1	25	57
Subbaiah, 2014	Fibroadenomas/infiltrating ductal carcinoma	FNAC	ANN	54	0	1	47	102
Teramoto, 2017	Adenocarcinoma/squamous cell carcinoma/small cell carcinoma	FNAC	DCNN	139	9	77	73	298

Abbreviations: ANN, artificial neural network; DCNN, deep convolutional neural network; DiagNN, Diagnosis neural network; FN, false negative; FNAC, Fine needle aspiration cytology; FP, false positive; MNN, multilayer perceptron neural network; TN, true negative; TP, true positive.

specific conditions, which improves the credibility of the results by increasing the number of samples to solve the inconsistency among the results of the study.⁶ This article applying meta-analysis method systematically evaluated the diagnostic value of NNs in FNAC, in order to find a reliable basis for assist cytopathologist to recognition, and reduce false negative diagnosis.

2 | MATERIALS AND METHODS

2.1 | Search strategy and study selection

Electronic databases, including Embase, PubMed, and Cochrane, were searched for studies published from the established of each database to September, 2019 that could be included after screening. Two investigators

(Jian Huang and Dongcun Wang) independently searched the databases. The following search strategy were used:(((Cytodiagnosis[MeSH Terms]) OR cytodiagnosis)) AND (((((((((((((((Neural Networks[MeSH Terms]) OR Network, Neural) OR Networks, Neural) OR Neural Network) OR Perceptrons) OR Perceptron) OR Neural Network Models) OR Connectionist Models) OR Connectionist Model) OR Model, Connectionist) OR Models, Connectionist) OR Models, Neural Network) OR Model, Neural Network) OR Network Model, Neural) OR Network Models, Neural) OR Neural Network Model) in PubMed, and ((Cytodiagnosis (all field)) AND neural networks (all field)) in Embase and Cochrane. The search was restricted to human studies and there was no restriction in terms of publication time. Study selection was based on the Preferred Reporting Items for Systematic Reviews and meta-Analyses (PRISMA) statement.⁷

FIGURE 2 Methodological quality graph and summary: review authors' judgments about each methodological quality item for each included study [Color figure can be viewed at wileyonlinelibrary.com]

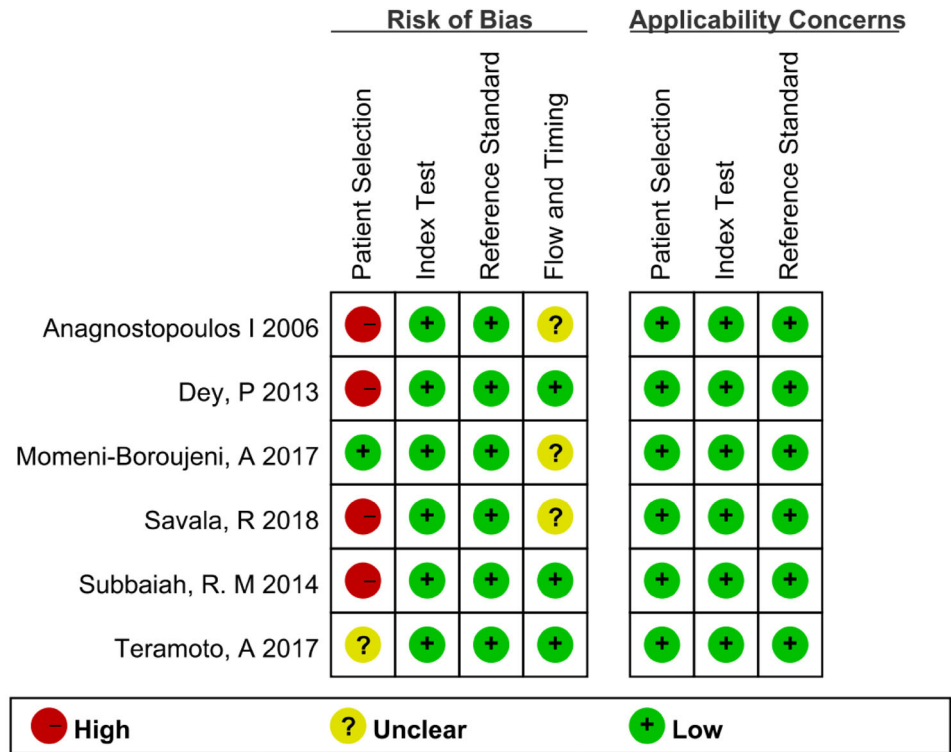
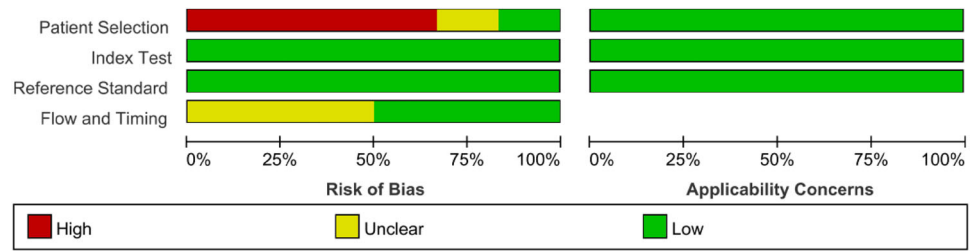
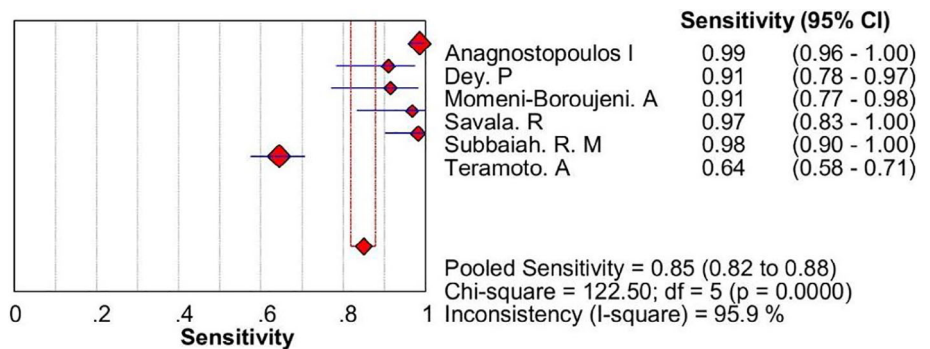


FIGURE 3 Forest plot of sensitivity [Color figure can be viewed at wileyonlinelibrary.com]



Analysis Options:
Add 1/2 to all cells of the studies with zero
Filter OFF

2.2 | Inclusion criteria and exclusion criteria

Inclusion criteria: (a) The purpose of study was to explore the value of NNs in FNAC; (b) The included article was a case-control study; (c) The diagnostic standard was postoperative pathological

examination or Diagnosis by experienced cytopathologist; (d) The critical reference value is clear; and (e) The results of study are clear and capable perform statistical analysis.

Exclusion criteria: (a) article theme or type inconsistency; (b) Non-primary study; (c) Result indicators cannot be statistically analyzed;

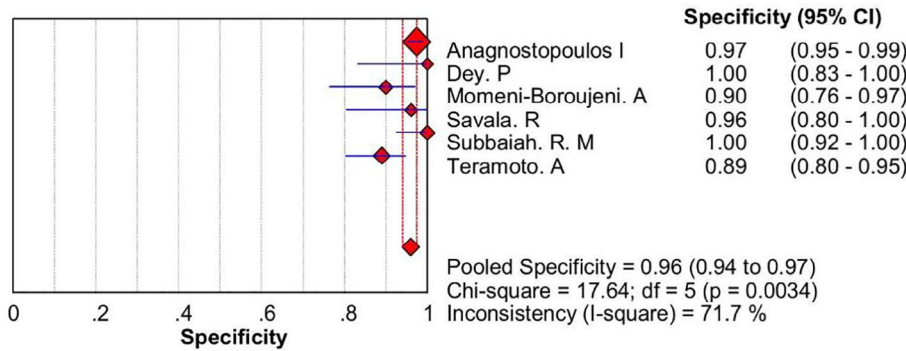


FIGURE 4 Forest plot of specificity [Color figure can be viewed at wileyonlinelibrary.com]

Analysis Options:
Add 1/2 to all cells of the studies with zero
Filter OFF

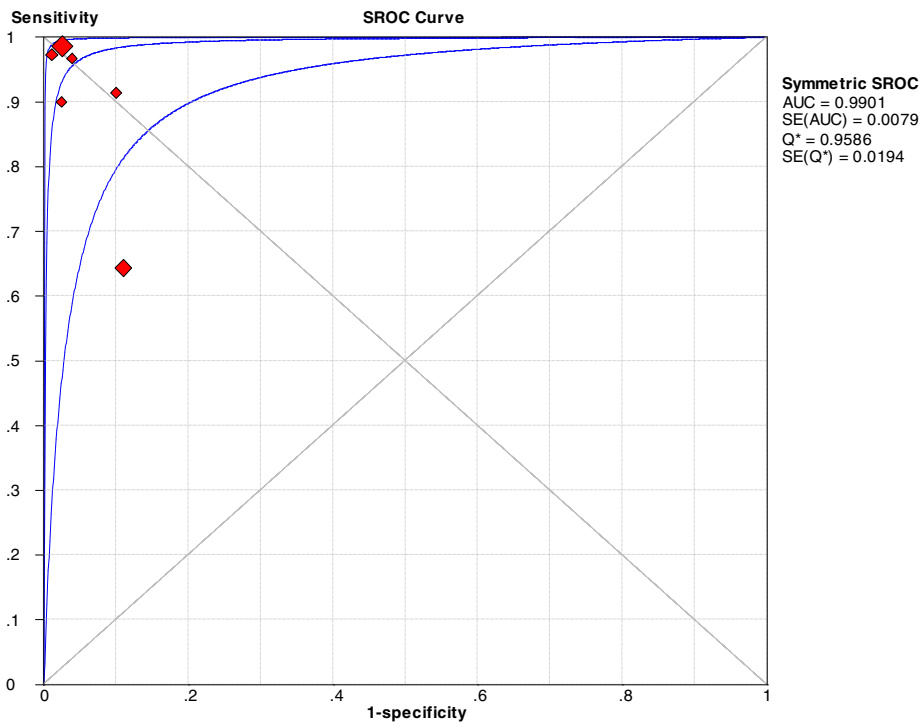


FIGURE 5 Forest plot of summary ROC curve [Color figure can be viewed at wileyonlinelibrary.com]

Analysis Options:
Add 1/2 to all cells of the studies with zero
Filter OFF
Symmetric SROC curve fitted using Moses' Model [Weighted regression (Inverse Variance)]
Defined relevant region: All ROC spac

(d) Review or Non-research article; (e) Single cell identify and tissue diagnosis; (f) No definite critical reference value; and (g) No full-text and Non-FNAC Type Literature.

A total of 396 papers were searched and six articles were selected according to inclusion and exclusion criteria by two investigators Independent screened.

2.3 | Analysis method and data processing

For diagnostic studies, QUADAS-2 (A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies) was used to assess the

methodological quality of eligible studies. The 14 requirements was evaluated according to the “yes, no, or unclear”, and the corresponding evaluation was carried out.⁸ We obtained data directly/indirectly from the study or compute it to made a four-grid table: true positive (TP) and false positive (FP), false negative (FN), and true negative (TN). Statistical analysis was calculated using Revman (St Albans, London) and meta-Disc software (Clinical Biostatistics Unit, Ramón y Cajal Hospital, Spain).^{9,10} This study used risk difference as a diagnostic effect to calculate the 95% CI. In a diagnostic test, one of the important reasons for heterogeneity is threshold effect: when threshold effect exists, the spearman of correlation coefficient of the logarithm of sensitivity and (1-specificity) is strongly positive.¹¹

3 | RESULTS

3.1 | Study selection and characteristics

The screening process for study inclusion was summarized in Figure 1. A total of 328 paper were retrieved from PubMed, 65 paper were retrieved from Cochrane and three paper were retrieved from Embase. Among them, 32 duplicated article were excluded first; 319 excluded was non-relevant paper according to title and abstract; 12 excluded in second screening because of article type; and 27 articles was excluded because of Single cell identify or Tissue pathology diagnosis, did not provide the relevant data, Non-FNAC Type Literature, even non-full text. Finally, only six studies were included in this studies, The characteristics of included studies are shown in Table 1 (the data were combined in some study according to Cochrane handbook).^{12,13}

3.2 | Quality assessment

The quality of diagnostic accuracy studies-2 (QUADAS-2) tools was adopted in order to evaluate the selected studies. In this assessment,

Analysis of Diagnostic Threshold

Spearman correlation coefficient: -0.543 p-value= 0.266
(Logit(TPR) vs Logit(FPR))

Moses' model (D = a + bS)

Weighted regression (Inverse Variance)

Var	Coeff.	Std. Erro	T	p-value
a	6.286	0.980	6.417	0.0030
b (1)	1.181	0.985	1.199	0.2968

Tau-squared estimate = 3.2890 (Convergence is achieved after 8 iterations)

Restricted Maximum Likelihood estimation (REML)

No. studies = 6

Filter OFF

Add 1/2 to all cells of the studies with zero

FIGURE 6 Analysis of diagnostic threshold

14 items were evaluated and each was either valued as “yes,” when positive or “no,” when unsupported or “unclear” due to unavailable and/or insufficient information. The QUADAS-2 evaluate result were shown in Figure 2. The majority of studies applied the ideal reference standard were of high quality.

3.3 | Meta-analysis results

After pooling all included studies, we performed by pooled sensitivity and specificity at first, as shown in Figures 3 and 4. We could find that the pooled sensitivity and specificity are 0.85 (95% CI:0.82-0.88; Figure 3) and 0.96 (95% CI:0.94-0.97; Figure 4), respectively. I^2 for both sensitivity (95.9%) and specificity (71.7%) is >50%. As such, we conducted a more cautious analysis using a random effect model.¹¹ The symmetrical sROC curve does not displaying a “shoulder arm” (Figure 5), and Spearman correlation coefficient is -0.543 ($P = .266$; Figure 6), suggesting that no threshold effect was present. The area under ROC curve (AUC) is 0.99, The positive LR is 18.43 (95% CI: 6.83-49.74), the negative LR is 0.06 (95% CI:0.001-0.58), and the diagnostic OR is 343.21 (34.41-3422.77) were shown in Figures 5 and 7-9.

4 | DISCUSSION

AI progress can be divided into three stages; its concept was first proposed at 1950s; machine learning appeared in 1980s, using mechanical thinking and logical knowledge to solve problems.¹⁴ Until recently, based of NNs technology of AI was applied in pathologic diagnosis, for example: the early tumor screening, disease preliminary classifying and between benign and malignancy recognizing.¹⁵⁻²² In the field of cytopathology, the PAPNET computer-assisted diagnosis system progressed based on “brain neural network” in 1992 greatly reduced the work of cytopathologists.²³⁻²⁶ At the same time, NNs diagnosis have been applied to FNAC. The results were listed Table 1, containing screen six literatures meta-analysis to evaluate the diagnostic value of NNs in FNAC.

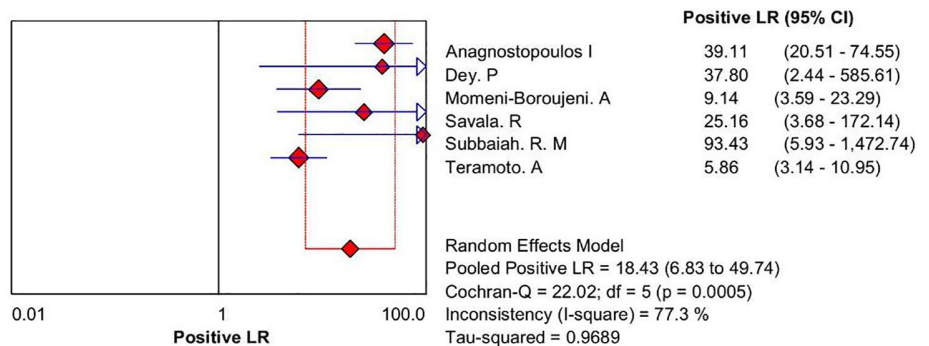


FIGURE 7 Forest plot of positive LR [Color figure can be viewed at wileyonlinelibrary.com]

Analysis Options:

Add 1/2 to all cells of the studies with zero
Filter OFF

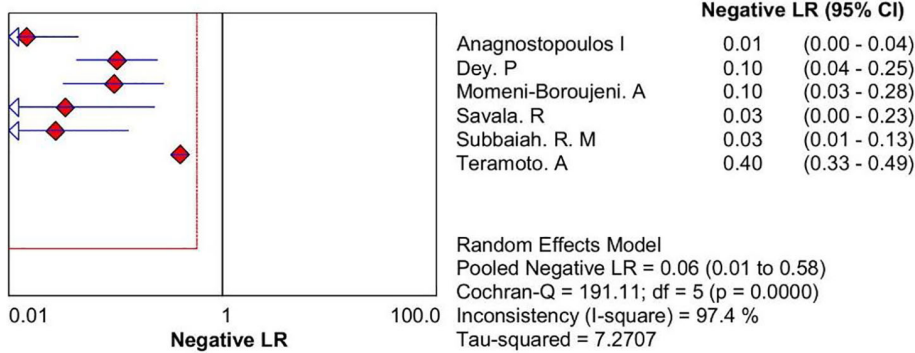


FIGURE 8 Forest plot of negative LR [Color figure can be viewed at wileyonlinelibrary.com]

Analysis Options:

Add 1/2 to all cells of the studies with zero
 Filter OFF

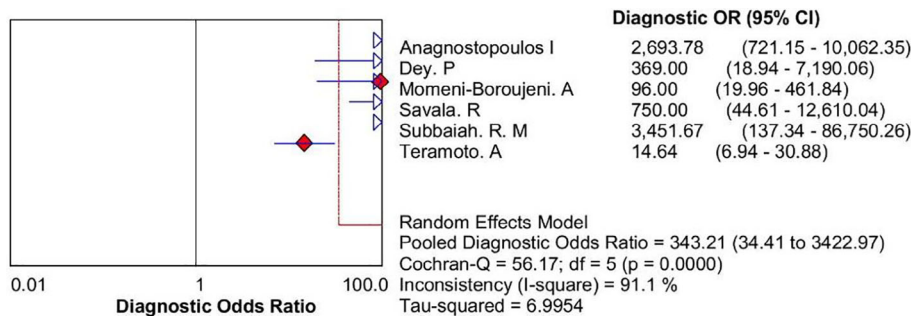


FIGURE 9 Forest plot of diagnostic OR [Color figure can be viewed at wileyonlinelibrary.com]

Analysis Options:

Add 1/2 to all cells of the studies with zero
 Filter OFF

The statistical analysis was performed by the Revman software. Meta-analysis showed that the sensitivity, specificity, positive LR, negative LR, and DOR of the cytology of AI in cancer were 0.85, 0.96, 18.43, 0.06, and 343.21, respectively (Figures 3, 4, and 7-9). According to the theory of medical statistics, if the diagnostic accuracy of NNs in cytology reaches 79%, and the sensitivity can reach >80%, Meanwhile, the area under the AUC curve reach 0.9, which indicates that this method has a higher diagnostic value.^{27,28} Obviously, in this study, the sensitivity of the NNs diagnosis was 85%, the specificity was 96%, and the AUC was 0.99 (Figure 5), so that we can find significance of the NNs in the FNAC of cancer. In future, the NNs can be applied to an unknown of qualitative data of FNAC to accurately identify the nature of the lesion. This will be a great assistance of the cytologist in cancer screening and other tedious work.

However, the shortcomings of this meta-analysis should be noted, for example: the number of studies included in this analysis was insufficient, the number of cases in the study sample is relatively small, the standard of exclusion and inclusion is not perfect and part of the literature cannot obtain. But meta-analysis as a new literature research method can play a positive role in enhancing the credibility of the results by quantitative synthesis of many studies of the same subject with

specific conditions. If we can reduce the source of heterogeneity and raise the inclusion standard of article, its conclusions are more reliable.

The larger sample size, the more reliable result, it will be according to the statistical principle. Due to the differences in sample quantity among the studies, the reliability of the results is not consistent, thus, we would like to have a multi-center, large-scale randomized controlled study, so as to improve the reliability of meta-analysis conclusions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Jian Huang  <https://orcid.org/0000-0002-0363-2412>

REFERENCES

- Meinkoth JH, Cowell RL. Sample collection and preparation in cytology: increasing diagnostic yield. *Vet Clin North Am Small Anim Pract.* 2002;32(6):1187-1207.
- Christopher MM, Hotz CS. Cytologic diagnosis: expression of probability by clinical pathologists. *Vet Clin Pathol.* 2004;33(2):84-95.
- Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng.* 2018;2(10):719-731.

4. El-Hassoun O, Maruscakova L, Valaskova Z, et al. Artificial intelligence in service of medicine. *Bratisl Lek Listy*. 2019;120(3):218-222.
5. Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. *Cancer Lett*. 2020; 471:61-71.
6. Delgado-Rodriguez M, Sillero-Arenas M. Systematic review and meta-analysis. *Med Intensiva*. 2018;42(7):444-453.
7. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
8. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
9. Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006;6:31.
10. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011; 343:d5928.
11. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-560.
12. Elamin MB, Flynn DN, Bassler D, et al. Choice of data extraction tools for systematic reviews depends on resources and review complexity. *J Clin Epidemiol*. 2009;62(5):506-510.
13. Furlan AD, Pennick V, Bombardier C, van Tulder M, Editorial Board, Cochrane Back Review Group. 2009 updated method guidelines for systematic reviews in the Cochrane Back review group. *Spine*. 2009; 34(18):1929-1941.
14. Mosquera-Lopez C, Agaian S, Velez-Hoyos A, Thompson I. Computer-aided prostate cancer diagnosis from digitized histopathology: a review on texture-based systems. *IEEE Rev Biomed Eng*. 2015;8: 98-113.
15. Dey P, Logasundaram R, Joshi K. Artificial neural network in diagnosis of lobular carcinoma of breast in fine-needle aspiration cytology. *Diagn Cytopathol*. 2013;41(2):102-106.
16. Muralidaran C, Dey P, Nijhawan R, Kakkar N. Artificial neural network in diagnosis of urothelial cell carcinoma in urine cytology. *Diagn Cytopathol*. 2015;43(6):443-449.
17. Teramoto A, Tsukamoto T, Kiriya Y, et al. Automated classification of lung cancer types from cytological images using deep convolutional neural networks. *Biomed Res Int*. 2017;2017:4067832.
18. Delibasis KK, Asvestas PA, Matsopoulos GK, Zoulias E, Tseleni-Balafouta S. Computer-aided diagnosis of thyroid malignancy using an artificial immune system classification algorithm. *IEEE Trans Inf Technol Biomed*. 2009;13(5):680-686.
19. Yoshida H, Shimazu T, Kiyuna T, et al. Automated histological classification of whole-slide images of gastric biopsy specimens. *Gastric Cancer*. 2018;21(2):249-257.
20. Mishra R, Daescu O, Leavey P, Rakheja D, Sengupta A. Convolutional neural network for histopathological analysis of osteosarcoma. *J Comput Biol*. 2018;25(3):313-325.
21. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-1567.
22. Kok MR, Boon ME, Schreiner-Kok PG, Koss LG. Cytological recognition of invasive squamous cancer of the uterine cervix: comparison of conventional light-microscopical screening and neural network-based screening. *Hum Pathol*. 2000;31(1):23-28.
23. Ashfaq R, Liang Y, Saboorian MH. Evaluation of PAPNET system for rescreening of negative cervical smears. *Diagn Cytopathol*. 1995;13 (1):31-36.
24. Sherman ME, Mango LJ, Kelly D, et al. PAPNET analysis of reportedly negative smears preceding the diagnosis of a high-grade squamous intraepithelial lesion or carcinoma. *Mod Pathol*. 1994;7(5):578-581.
25. Ouwerkerk-Noordam E, Boon ME, Beck S. Computer-assisted primary screening of cervical smears using the PAPNET method: comparison with conventional screening and evaluation of the role of the cytologist. *Cytopathology*. 1994;5(4):211-218.
26. Koss LG, Lin E, Schreiber K, Elgert P, Mango L. Evaluation of the PAPNET cytologic screening system for quality control of cervical smears. *Am J Clin Pathol*. 1994;101(2):220-229.
27. Moore RG, McMeekin DS, Brown AK, et al. A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol*. 2009;112(1):40-46.
28. Cui R, Wang Y, Li Y, et al. Clinical value of ROMA index in diagnosis of ovarian cancer: meta-analysis. *Am J Clin Pathol*. 2019;11:2545-2551.

How to cite this article: Huang J, Wang D, Da J. Automated classification of cancer from fine needle aspiration cytological image use neural networks: A meta-analysis. *Diagnostic Cytopathology*. 2020;48:1027-1033. <https://doi.org/10.1002/dc.24520>