Peer Review File
Article information: https://dx.doi.org/10.21037/jtd-23-81

Reviewer A

This study investigated the feasibility of a new automated system for EUS-FNA specimen processing by comparing it with conventional processing methods. Automation is highly important, and the realization of such system would be very useful in clinical practice. On the other hand, the methods and results of this study lack scientific credibility due to many statements that could be considered contradictory. I recommend the authors that the manuscript be resubmitted in a more polished form.

Major comments
Comment 1: Although this study uses diagnostic accuracy as its primary endpoint, it does not provide its detailed definition and the rationale for setting the sample size, despite the fact that it is positioned as a prospective study. In addition, it appears that two pathologists made the decision, but it is not clear whether the decision was made independently or collegially, and in the former case, the details regarding the degree of agreement are also unclear. If they only accumulated cases prospectively in the first place, and the analysis was done retrospectively and exploratory, it should be clearly stated as such.
Reply 1: Thank you for this comment. The description of the study as a pilot study has been further refined (ln 185-190, 197-199). Since this was a pilot, feasibility study, we did not go through a rigorous 'sample size calculation' to arrive at an optimal sample size. Further evaluation of the slide quality will be pursued with a large sample size to adequately assess superiority vs. inferiority. The pathologists reviewed the slides independently without cross over evaluation (ln 263-264). Additional description of pathology assessment were added to address this comment as no "degree of agreement" could be calculated.

Comment 2: In the "Procedural Description" section, it is stated that "one lymph node per patient was selected to undergo each sample preparation pathway," but at the beginning of the Results section, it is stated that "One to three lymph nodes were sampled for each patient for study purposes for a total of 72 paired samples collected." Furthermore, 66 samples were actually evaluated for ASP and 60 samples for SOC, but there is no explanation for these discrepancies, including the above.
Reply 2: This discrepancies in sample numbers have been adjusted in description. 72 paired sample were collected. For evaluation, any sample that was nondiagnostic was removed. If only one of the pairs was nondiagnostic, the other remained in the calculation of sample assessment for that modality. Only the pairs of samples were used to assess diagnostic equivalency. This has been more explicitly stated in the manuscript (ln 268-275).

Comment 3: It seems that after the first 24 samples, the ASP was adjusted due to the discordant material, is this something that was stipulated in the protocol or was it properly revised? In addition, it seems unreasonable to evaluate cases in the pre- and post-stage together, but how is this handled in the protocol?
Reply 3: This change was properly revised due to the finding of the discordant material. This has been reworded to provide better clarity. Additional analysis was performed on the post-stage only and the results reported. We feel that analysis of all samples, regardless of stage, warrants consideration since slide quality was assessed and not equivalency. This information was kept in the manuscript for completeness (ln 268-275).

Comment 4: While it is understandable that various evaluation items were listed and that each was rated on a three-point scale, a detailed explanation should be provided. Unless details are provided not only by item name, but also by what scale was used for what indicators, or unless the references are clearly indicated, it is difficult for first-time readers to correctly interpret the results.
Reply 4: Thank you. This three-point scale was used since what makes a slide "good" vs "poor" may be defined differently by different pathologists. However, in order to account for this issue, the pathologist reviewed both the SOC and the instrument slide for each pair (ln 264-265).

Comment 5: I understand that SOC stands for standard of care, but does ASP refer to the name of the system (or company)? Especially in the Abstract, it is impossible for first-time readers to understand if ASP is mentioned without any indication. Statements that may evoke the name of a specific company should be avoided, and what they stand for (is that automated sample preparation?) should be clearly stated.
Reply 5: Thank you. We have removed the designation and have called the study slide previously "ASP" to "instrument" (changed throughout the manuscript).

Minor comments
Is the RedCap electronic data captured? Even at a minimum, an explanation should be appended.
Reply: Thank you. The REDCap was used for data collection and a description added to the manuscript consistent with projectredcap.org (ln 204-209).

Reviewer B

The reviewer is honored to review an article about an evaluation of automated sample preparation system for lymph node sampling in EBUS-FNA. The paper was a prospective single-center study. The paper is well written and easy to understand, and so this paper is potentially acceptable; however, there are several points to be revised, as follows:

Comment 1: In the text, Stage 1 and stage 2 were written in the study slides section, but Stage I and Stage II were written in Figure 2. Please unify these descriptions.
Reply 1: Thank you. Changes in the text were made to maintain consistency (ln 234, 242).

Comment 2: According to the text, we found that ASP slides were much better than SOC slides in their qualities. The authors should comment on the cost-benefit when this machine is used.
Reply 2: We have added comments about the cost of the machine and the need for additional analysis of cost-benefit (ln 380).

Reviewer C

In this study, the authors conclude that slides prepared by the automated system are of noninferior quality for adequacy assessment with diagnostic concordance compared to SOC slides. Simplifying ROSE and uniformity of slide quality in EBUS-TBNA will reduce the burden on surgeons during the procedure and benefit those judging cytology. To improve the quality of this paper, we would like to ask the authors a few questions, as indicated below.

Major comments
Comment 1: It would like the authors to add the detailed diagnostic results of the punctured lymph node. Furthermore, please add whether there was a difference in staining results of ASP between malignant and benign diseases.
Reply 1: Thank you. The differences in staining results for malignant vs benign diseases were added (ln 303-307)

Comment 2: What could be the reason why only Monolayer had better staining results with APS?
Reply 2: The instrument utilizes a unique method of spraying the specimen on the glass slide which creates a thin monolayer on the glass slide. This is different from a traditional smearing process which applies pressure between the two glass slides to make a smear. Hence, this is likely the reason for the samples prepared by the instrument yielding a higher percentage of monolayer cells compared to the traditional smearing process (ln 357-362).

Comment 3: What needle size was used in the EBUS-TBNA? Also, was the EBUS-TBNA performed using the negative pressure method or the slow pull method?
Reply 3: Thank you. We have adjusted the methods description to include the information (22G via the slow pull method) (ln 216-217).

Comment 4: Was the doctor the one who performed the ROSE of the SOC? Was it a cytology technician? I suggest adding this information together with approximate years of experience.

Reply 4: The cytotechnologist performs the ROSE in the bronchoscopy with further evaluation of the study slides by two pathologists. The cytotechnologists are part of a large group of different years of experience and are assigned to the bronchoscopy suite based on a predetermined schedule (ln 230-231).

Minor comments
Comment 1: How much does the ASP device cost? $15000.
Reply 1: This was added to line 380.

Comment 2: Is staining with the ASP device also possible for stamp cytology with specimens collected by forceps biopsy?
Reply 2: Thank you. This information was added to the background section (ln 178-180).

Reviewer D

Comment 1: Diagnostic equivalency and diagnostic accuracy are not the same things to me. Diagnostic equivalency means SOC and ASP slide results match. Diagnostic accuracy may need more.
Reply 1: We appreciate this comment and want to focus on diagnostic equivalency. We have updated the manuscript accordingly (ln 205 switched “accuracy” to “equivalency”).

Comment 2: After a few days of procedures, how long does it take to replenish material in machine? Is it possible to let give us an idea of the maintenance needs of machine?
Reply 2: A discussion of maintenance of the machine was added to the introductions (ln 178-182).