J Pathol Inform

Editor-in-Chief:
Anil V. Parwani , Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA

OPEN ACCESS
HTML format

For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

Original Article

Investigation of scanning parameters for thyroid fine needle aspiration cytology specimens: A pilot study

Maheswari S. Mukherjee[1], Amber D. Donnelly[1], Elizabeth R. Lyden[2], Whitney R. Wedel[3], Mary F. McGaughey[4], John J. Baker[3], Stanley J. Radio[1,3]

[1]Cytotechnology Education, College of Allied Health Professions, University of Nebraska Medical Center, [2]Department of Biostatistics, College of Public Health, [3]Department of Pathology and Microbiology, College of Medicine, University of Nebraska Medical Center, [4]Department of Anatomic Pathology, Nebraska Medicine, Omaha, NE, USA

E-mail: *Maheswari S. Mukherjee - mmukherj@unmc.edu
*Corresponding author

Received: 27 January 2015 Accepted: 21 May 2015 Published: 28 July 2015

Abstract

Background: Interest in developing more feasible and affordable applications of virtual microscopy in the field of cytology continues to grow. **Aims:** The aim of this study was to investigate the scanning parameters for the thyroid fine needle aspiration (FNA) cytology specimens. **Subjects and Methods:** A total of twelve glass slides from thyroid FNA cytology specimens were digitized at ×40 with 1 micron (μ) interval using seven focal plane (FP) levels (Group 1), five FP levels (Group 2), and three FP levels (Group 3) using iScan Coreo Au scanner (Ventana, AZ, USA) producing 36 virtual images (VI). With an average wash out period of 2 days, three participants diagnosed the preannotated cells of Groups 1, 2, and 3 using BioImagene's Image Viewer (version 3.1) (Ventana, Inc., Tucson, AZ, USA), and the corresponding 12 glass slides (Group 4) using conventional light microscopy. **Results:** All three raters correctly identified and showed complete agreement on the glass and VI for: 86% of the cases at FP Level 3, 83% of the cases at both the FP Levels 5 and 7. The intra-observer concordance between the glass slides and VI for all three raters was highest (97%) for Level 3 and glass, same (94%) for Level 5 and glass; and Level 7 and glass. The inter-rater reliability was found to be highest for the glass slides, and three FP levels (77%), followed by five FP levels (69.5%), and seven FP levels (69.1%). **Conclusions:** This pilot study found that among the three different FP levels, the VI digitized using three FP levels had slightly higher concordance, intra-observer concordance, and inter-rater reliability. Scanning additional levels above three FP levels did not improve concordance. We believe that there is no added benefit of acquiring five FP levels or more especially when considering the file size, and storage costs. Hence, this study reports that FP level three and 1 μ could be the potential scanning parameters for the thyroid FNA cytology specimens.

Key words: Digital imaging, scanning parameters, thyroid fine needle aspiration, virtual microscopy

Access this article online

Website: www.jpathinformatics.org

DOI: 10.4103/2153-3539.161610

Quick Response Code:

INTRODUCTION

Virtual microscopy (VM), a digital imaging technology is being successfully used in cytology education, proficiency testing, and telecytopathology.[1-10] With an effort to make use of digital images in routine diagnostic cytopathology, the diagnostic accuracy has been compared between digital images and traditional glass slides in previous studies.[11-13] One of these studies[13]

How to cite this article: Mukherjee MS, Donnelly AD, Lyden ER, Wedel WR, McGaughey MF, Baker JJ, *et al.* Investigation of scanning parameters for thyroid fine needle aspiration cytology specimens: A pilot study. J Pathol Inform 2015;6:43.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2015/6/1/43/161610

found that the diagnostic accuracy was higher with glass slides. One of the possible reasons for this result could be due to the fact that the digital images used in this study were scanned using a single focal plane (FP) level. Because of the thick and three-dimensional (3D) nature of the cytopathology specimens, the single FP level is not always sufficient to obtain adequate focus of the cell clusters. While these focusing issues could possibly be overcome using z-axis level scanning methods, the disadvantages (one being large file size that increases with the number of FPs used) of z-axis scanning are still limiting the usage of digital images in routine diagnostic cytopathology.[1,6,7,11,14-18] Hence, finding optimal scanning parameters to digitize the cytopathology specimens while maintaining a reasonably smaller file size is needed.

A previous study determined that while considering the file size and intra-rater diagnostic concordance, the optimal scanning parameters for SurePath™ prepared gynecological (GYN) cytology specimens are three FP levels at a 1 micron (μ) interval.[19] In this study, we aimed to investigate the scanning parameters for thyroid fine needle aspiration (FNA) cytology specimens.

SUBJECTS AND METHODS

For this Institutional Review Board approved study, we used the CoPathPlus computer application, to retrieve 10 thyroid FNA cytology cases for each of the following categories: Benign-benign thyroid nodule (BTN)/negative for malignancy (including colloid nodule); neoplasm-cytologic features consistent with hurthle cell neoplasm (HCN); neoplasm-cytologic features consistent with follicular cell neoplasm (FCN); and malignant papillary thyroid carcinoma (PTC).

An experienced cytotechnologist screened all the cases and selected a set of 12 cases among those 40 cases. These 12 cases consisted of: Three BTN, three HCN, three FCN, and three PTC. Each of the selected 12 cases had many slides, with an average number of eight sides per case (range: 4–13). Among those 12 cases, one slide per case that was: Prepared with smear technique; Papanicolaou stained; and best represented the diagnostic criteria was selected for the study purpose. These slides had dots which were placed by the pathologist and cytotechnologist who initially diagnosed the case. These selected slides with previous dots were further screened and approved by an experienced pathologist.

Scanning

Scanning the whole slide requires more time and results in bigger file sizes. Therefore, for the slides which had a smear that covered more than 75% of the slide was marked no <40% of the smeared portion for scanning. For the slides which had a smear covering less than half of the slide, the entire smeared area was scanned. The average area scanned per slide was approximately 83% (range: 40–100%).

After photocopying these slides for future reference, the ink dots made by the pathologists and cytotechnologists who originally diagnosed these cases were removed. These slides were subsequently de-identified and relabeled. Using iScan Coreo Au scanner (Ventana, AZ, USA), each of the 12 slides were digitized at ×40 (numerical aperture = 0.75) with 1 μ interval using seven FP levels (Group 1), five FP levels (Group 2), three FP levels (Group 3). Thus, 36 virtual images (VI) were produced; with an average file size of 16 GB, 12 GB, and 9 GB for the image scanned using seven, five, and three FP levels, respectively [Table 1]. The output image files in Bioimagene format, were saved in a password protected encrypted external hard drive.

Once digitized, the original ink dots were replaced on the glass slides using the previously taken photocopy. From these dotted glass slides, an experienced cytotechnologist selected a few dots which represented the diagnostic criteria of the cases. The dots which were not selected were erased. The selected dots were digitally annotated in the corresponding VI, using the software Image Viewer (version 3.1) (Ventana, Inc., Tucson, AZ, USA) a web-based application by BioImagene.

Data Collection

After a brief training session on accessing the VI saved on the encrypted external hard drive and using the Image Viewer Software (Ventana, Inc., Tucson, AZ, USA) to screen the VI, a cytopathologist, a pathology resident, and a cytotechnologist diagnosed 36 preannotated VI. Subsequently, they diagnosed the corresponding dotted glass slides (Group 4) using light microscopy (LM). The participants were asked to provide their diagnoses by choosing the following choices: (a) BTN; (b) HCN; (c) FCN; or (d) PTC. An additional choice, (e) unable to diagnose was also added and participants were asked to select this choice, if they found that they were unable to diagnose the VI or the glass slides due to the quality of the images/slides or any other reasons. The participants were told that all the cases contained an adequate sampling of cells for diagnosis. The participants were encouraged to review other cells, in addition to the annotated cells on the VI and glass slides if they needed. In addition, they were told that if they felt it necessary they could use the image enhancement feature of the software, which helps to increase/decrease the brightness

Table 1: Average file sizes of the VI scanned using 7, 5, and 3 FP levels

Images scanned with z-axis levels	Average file size in GB
7 FP levels	16
5 FP levels	12
3 FP levels	9

FP: Focal plane, VI: Virtual images

and contrast of the VI. Furthermore, the participants were requested to record the screening times for each virtual/glass in the respective diagnosis log sheet.

Participants were encouraged to comment on the quality of the VI on their diagnosis sheet. Because of the participants' time constraints the washout period was an average of 2 days. All participants independently diagnosed the VI/glass slides using their personal computer monitors/microscopes without any major interruptions and distractions.

Statistical Analysis

Concordance

Concordance was summarized using a percentage and a 95% confidence interval. Concordance with the reference slide was attained when the reviewer's diagnosis of the glass slide and VM (at a particular FP level) agreed with the reference diagnosis. Intra-observer concordance was attained when the reviewer's diagnosis based on the glass slide agreed with his/her diagnosis on VM (at a particular FP level).

Inter-rater reliability

Inter-rater reliability was examined for glass, FP Levels 7, 5, and 3 separately. For interpretation of kappa statistics, a kappa statistic below 0.00 was considered as "poor agreement;" 0.00–0.20 was considered as "slight agreement;" 0.21–0.40 was considered as "fair agreement;" 0.41–0.60 was considered as "moderate agreement;" 0.61–0.80 was considered as "substantial agreement;" and 0.81–1.00 was considered as "almost perfect agreement."[20]

Screening time

Means, standard deviations (SDs), medians, minimums, and maximums were used to describe minutes of viewing. A linear mixed effects model with a random effect for specimen slide and a fixed effect for type of slide (i.e., glass, VM Level 3, VM Level 5, and VM Level 7) was used to compare the minutes of slide viewing based on type of slide for each participant. The mean minutes for each level of VM were compared to the mean minutes for glass using Tukey's test. A contrast statement that compared all levels of VM to glass was also included for each participant. Residual plots were done to verify the assumptions of the model. Two-sided tests were used and a $P < 0.05$ was considered statistically significant.

RESULTS

Case Details

A set of 12 thyroid FNA cytology specimen glass slides/cases consisting of three BTN, three HCN, three FCN, 3PTC were scanned at ×40, 1 μ interval using seven FP levels, five FP levels, and three FP levels producing 36 VI. Three participants (one pathologist, one pathology resident, and one cytotechnologist) diagnosed these 36 VI and the corresponding 12 glass slides and categorized their interpretations as: BTN, HCN, FCN, PTC, or, unable to diagnose.

Concordance

All three participants correctly identified and showed complete agreement on the glass slides and VI for: 86% of the cases at FP Level 3, and 83% of the cases at both the FP Levels 5 and 7 [Figure 1].

Concordance by Participant

The concordance by participant is summarized in Table 2:

- Participant one showed overall complete agreement on glass slides and VM for 86% of the cases. Among the VI scanned at three different FP levels, participant one showed highest agreement for the VI scanned at three FP level (92%), followed by VI scanned at five and seven (83%)
- Participant two showed overall complete agreement on glass slides and VM for 92% of the cases; and had the same level of agreement (92%) for the VI scanned at three, five, and seven FP levels
- Participant three showed overall complete agreement on glass slides and VM for 75% of the cases; and had the same level of agreement (75%) for the VI scanned at three, five, and seven FP levels.

Intra-observer Concordance

The intra-observer concordance between the glass slides and VI for all three raters was highest (97%) for Level 3 and glass [Table 3]. The concordance rate was same (94%) for Level 5 and glass; and Level 7 and glass.

- Participant one had 100% concordance rate for Level 3 and glass; and 92% concordance rate for both Level 5 and glass and Level 7 and glass
- Participant two had 100% concordance rate for Level 3 and glass, Level 5 and glass; and Level 7 and glass
- Participant three had 92% concordance rate for Level 3 and glass, Level 5 and glass; and Level 7 and glass.

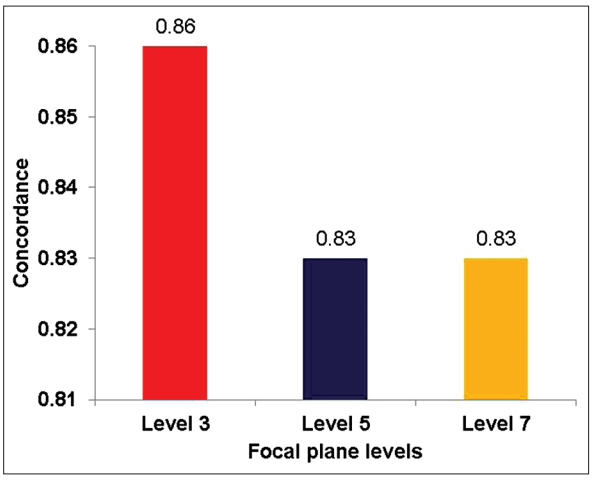

Figure 1: The complete agreement of all three participants on the glass slides and virtual images (digitized using focal plane Levels: 3, 5, and 7)

Inter-rater Reliability

The results of inter-rater reliability are summarized in Figure 2. For the interpretations of BTN, HCN, FCN, PTC: The overall kappa among all participants using glass slides was 0.77; the overall kappa among all participants using VI scanned at three FP level was 0.77; the overall kappa among all participants using VI scanned at five FP level was 0.70; the overall kappa among all participants using VI scanned at seven FP level was 0.69.

Screening Time

- For each of the three participants, there was a statistically significant difference in the mean number of minutes to read at least two of the four types of slides presented ($P < 0.0001$)
- For participant one, there was a statistically significant difference in the mean number of minutes for reading, Level 7 compared to glass and Level 5 compared to glass ($P < 0.0001$ and $P = 0.0003$, respectively). Furthermore, when all VM levels were combined and compared to the mean number of minutes for reading glass slides, there was a statistically significant difference ($P < 0.0001$)
- For participant two, the mean number of minutes for reading Level 5 differed significantly compared to glass ($P = 0.0002$). Comparing the mean number of minutes for reading all VM levels and to reading glass slides showed no significant difference ($P = 0.23$)
- For participant three, there was a statistically significant difference in the mean number of minutes for reading Level 3 and Level 7 compared to glass ($P < 0.0001$ and $P = 0.04$, respectively). Furthermore, when all VM levels were combined and compared to the mean number of minutes for reading glass slides, there was statistically significant difference ($P < 0.0001$).

Table 2: Concordance and 95% CI by participant

Overall (*n*=36) (%)	FP level 3 (*n*=12) (%)	FP level 5 (*n*=12) (%)	FP level 7 (*n*=12) (%)
86 (75, 97)	92 (76, 100)	83 (62, 100)	83 (62, 100)
92 (82, 100)	92 (76, 100)	92 (76, 100)	92 (76, 100)
75 (61, 89)	75 (51, 100)	75 (51, 100)	75 (51, 100)

CI: Confidence interval, FP: Focal plane

Table 3: The intra-observer concordance and 95% CI between the glass slides and VI (*n*=12 for each level) for all three participants

Participants	Level 7 and glass (%)	Level 5 and glass (%)	Level 3 and glass (%)
All	94 (87, 100)	94 (87, 100)	97 (92, 100)
1	92 (76, 100)	92 (76, 100)	100
2	100	100	100
3	92 (76, 100)	92 (76, 100)	92 (76, 100)

VI: Virtual images, CI: Confidence interval

Participant one

The mean (SD) screening times were 2.1 (1.2) min for glass; 6.9 min (1.8) for Level 7; 5.3 (3.0) min for Level 5; and 3.3 (0.9) min for Level 3.

Participant two

The mean (SD) screening times were 2 (2) min for glass; 1.6 (0.8) min for Level 7; 4.8 (2) min for Level 5; and 1.3 (0.7) min for Level 3.

Participant three

The mean (SD) screening times were, 2.3 (1.4) min for glass; 4.8 (4) min for Level 7; 4.1 (2.2) min for five levels; and 7.9 (3.9) min for Level 3.

DISCUSSION

The purpose of this study was to investigate the scanning parameters for thyroid FNA cytology specimens.

A previous study determined that while considering the file size and intra-rater diagnostic concordance, the optimal scanning parameters for SurePath™ prepared GYN cytology specimens are three FP levels at a one μ interval.[19] For this study, we chose to start with three FP levels and 1 μ interval and added five and seven FP levels, since we knew that some FNA specimens might be more cellular and 3D in nature (PTC for instance).

As the results of this study indicate, three FP levels had the slightly higher concordance with the glass slides (86%), and the intra-observer concordance between the glass slides and VI for all three raters was highest (97%) for three FP levels and glass slides. Furthermore, the inter-rater reliability was the same for glass slides and three FP levels (0.77); however, screening time was varied. However, this study did not use the "nondiagnostic or unsatisfactory," "atypia of undetermined significance or follicular lesion of undetermined significance," and

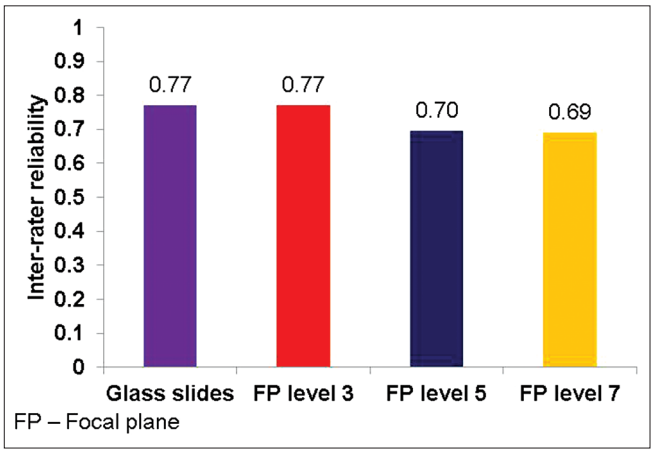

Figure 2: The summary of the inter-rater reliability for glass slides, virtual images (digitized using focal plane Levels: 3, 5, and 7)

"suspicious for malignancy" categories. The concordance rate and inter-rater reliability may have decreased with the inclusion of these "harder" categories.

Case Discrepancies

One FCN case was diagnosed as BTN by participants two and three; however, both participants diagnosed this case as BTN on the glass slide as well as the VI scanned at seven, five, and three FP levels. Hence, the participants were consistent with the diagnoses on both the glass slide and all VI. No comments were given on this image by the participants.

Another FCN case was diagnosed as BTN by participant three. Again, this participant diagnosed this case as BTN on the glass slide as well as the VI scanned at seven, five, and three FP levels. Hence, the participant was consistent with the diagnoses on both the glass slide and all VI. A general comment by the participant on this case on the VI scanned using seven, five and three was: "Hard to see if there is any colloid present; and did not see any colloid."

One HCN case was diagnosed as BTN on the VI scanned using seven FP levels by participant one. The comment on this case was: "Reasonably confident, colloid visualized well."

One BTN case was diagnosed by participant one as PTC on the VI scanned using five FP levels. The comment given by the participant for this case was that: "Low confidence, several groups would not focus."

Another HCN case was diagnosed by participant one as PTC on both the glass slide and all the VI scanned at seven, five, and three FP levels. The participant was consistent with diagnoses between the glass slide and VI. The same HCN case was diagnosed as PTC by participant three on the VI scanned using seven, five, three FP levels, but as HCN on the glass slide. Even though the diagnosis differed between the glass slide and VI, it was consistent among all the VI. We believe that the reason behind the diagnosis of this HCN case as PTC could be the presence of intranuclear cytoplasmic inclusions.

All participants diagnosed the VI using their personal computers. Participant one commented in the additional comment section that there were significant system problems and the software frequently "locks up," necessitating shutting down and reopening the program. However, during such times, this participant noted the screening time for image that was reopened. We experienced this "locking up" of the software in our previous studies, as well as when annotating the VI. We believe that one of the potential reasons for this problem could be clicking on the images (to change the magnification, change the FP level or to move to the next annotation) too fast without giving enough time for the software to process. There were no comments regarding the software lock-up issues from the other two participants.

Overall, the participants used the image enhancement feature of the software to change the brightness and contrast of the VI for many cases. Participant three felt some of the groups were dark even after using the image enhancement feature. Participant one mentioned in the open comment section that the VI were recognizable and thus were biased; however, in spite of the participant's comment, the varying diagnoses of two out of three cases were not consistent among all three sets of VI and the glass slides demonstrating that the participant did not fully recognize the cases. The third case was misinterpreted consistently on the glass slides as well as the VI scanned using seven, five, and three FP levels. One of the limitations of this study was the short washout period between each group. The washout period varied depending on the time and schedules of the participants. Even though there was a washout period of 4 days in between some reads, the average of the washout period was 2 days which was similar to a previous study,[13] which had a short washout period of 3 days.

Self-reporting the screening time could also be a limitation of this study. The diagnosis log sheet provided to participants had columns with "starting time," "end time" for recording their screening time. Participants used their computer/wrist watch to enter the time when they started and ended screening the image/glass slides. Participant one decreased the time with a decrease in FP levels (mean screening times: 6.9 min for seven FP; 5.3 min for five FP; and 3.3 min for three FP), and overall, the VM took longer than the glass slides (mean screening time for glass slides: 2.1 min), which is what one would expect based on previous studies. Participant two took the longest to screen five FP (mean screening time: 4.8 min) and the seven FP (mean screening time: 1.6 min) and three FP (mean screening time: 1.3 min) took a shorter amount of time than the glass slide (mean screening time: 2.0 min). The reasons for shorter screening time for seven FP and three FP levels could be either due to the participant's miscalculation of time, shorter washout period between the sets where the participant remembered the images and the diagnoses even before taking time to look at all the annotations, or the participant had less experience with LM, and hence diagnosed the cases faster using VM. Participant two had no comments on the quality of the images, except mentioning the usage of image enhancement for some of the images. Participant three took the longest to screen three FP (mean screening time: 7.9 min). This participant noted in the open comment section that difficulty seeing dark groups, even after using the image enhancement feature of the software, was the reason for

the longer screening time of the VI scanned using three FP levels. This could be due to the participant's personal computer problems.

The participants were asked to circle the FP level on the diagnosis log sheet they felt confident to make the diagnosis of the VI. When given three FP and five FP levels, the participants used all levels to render their diagnoses, but when given seven FP levels, only five FP levels were used [Table 4]. We believe this is because the 3D groupings did not have a depth great enough for seven FP levels. After five FP levels the cells, even in 3D groupings, went out of focus. Perhaps a smaller interval between levels (we used 1 μ) would create a need for more FP levels. The reason 1 μ was chosen in this study was because a previous GYN study demonstrated that 1 μ interval was potentially the best interval to scan GYN specimens prepared with SurePath™. Since no alternate intervals were investigated in this study, however, we are unsure if this made a difference. Regarding the screening time, even though the participants circled -2 or 2 levels [Table 4] and felt confident to give the diagnosis using five FP levels, we believe that they did look at all seven levels to make that decision. Hence, it is understood that the time reported for seven FP level was for seven FP levels.

Another limitation of the study could be the experience of the participants. Participant one has practiced cytology for over 30 years, participant two was in the 3rd year of training, and participant three has practiced cytology for 4 years. All participants have different levels of experience with VM as well. Participants two and three have volunteered for many VM studies and have the most experience, while this study was the first for participant one.

CONCLUSION

This pilot study found that among the three different FP levels, the VI digitized using three FP levels had slightly higher concordance, intra-observer concordance, and inter-rater reliability. Furthermore, considering the inter-rater reliability, although all three FP levels (seven, five, and three) had substantial agreements, the inter-rater reliability was found to be the same for glass slides and the VI digitized using FP Level three.

In addition, although a few of the diagnoses made by the participants differed from the reference diagnoses, all diagnoses made using with three FP were the same as the glass slide diagnoses except one VI in the HCN category. For these discrepant cases, the technology used (LM vs. VM) did not appear to play a major role since the participants were consistent in their diagnosis on both glass slides and VI.

Based on the results of this pilot study, we believe that there is no added benefit of acquiring five FP levels or more especially when considering the additional scanning time, file size, and storage costs. Hence, this study reports that FP Level 3 and 1 μ could be the potential scanning parameters for the thyroid FNA cytology specimens. However, the sample size in this study was small, and there were only three participants in this pilot study. Hence, the generalizability of our conclusions is restricted. Hence, the scanning parameters investigated from this pilot study (three FP, 1 μ interval) should be investigated with a larger sample size, to determine if three FP level and 1 μ interval is optimal for thyroid FNA cytology specimens in order to use in clinical situation. However, we believe that these scanning parameters could be sufficient enough to scan the thyroid cytology specimens for creating annotated teaching modules for educational purposes.

Table 4: FP levels used by the participants to interpret the VI scanned using FP levels: Seven, five, and three

	7 FPs	5 FPs	3 FPs	Overall
BTN	−2,2	−2,2	−1,0	−2,2
HCN	−1,1	−2,0	−1,0	−2,1
FCN	−2,2	−1,2	−1,1	−2,2
PTC	−2,2	−1,2	−1,0	−2,2
Overall	−2,2	−2,2	−1,1	

VI: Virtual images, BTN: Benign-benign thyroid nodule/negative for malignancy (including colloid nodule), HCN: Neoplasm-cytologic features consistent with hurthle cell neoplasm, FCN: Neoplasm-cytologic features consistent with follicular cell neoplasm, PTC: Malignant-papillary carcinoma, FPs: Focal planes

Financial Support and Sponsorship
Nil.

Conflicts of Interest
There are no conflicts of interest.

REFERENCES

1. Steinberg DM, Ali SZ. Application of virtual microscopy in clinical cytopathology. Diagn Cytopathol 2001;25:389-96.
2. Neel JA, Grindem CB, Bristol DG. Introduction and evaluation of virtual microscopy in teaching veterinary cytopathology. J Vet Med Educ 2007;34:437-44.
3. Goldberg HR, Dintzis R. The positive impact of team-based virtual microscopy on student learning in physiology and histology. Adv Physiol Educ 2007;31:261-5.
4. Chen YK, Hsue SS, Lin DC, Wang WC, Chen JY, Lin CC, *et al.* An application of virtual microscopy in the teaching of an oral and maxillofacial pathology laboratory course. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2008;105:342-7.
5. Kalinski T, Zwönitzer R, Sel S, Evert M, Guenther T, Hofmann H, *et al.* Virtual 3D microscopy using multiplane whole slide images in diagnostic pathology. Am J Clin Pathol 2008;130:259-64.
6. Mori I, Nunobiki O, Ozaki T, Taniguchi E, Kakudo K. Issues for application of virtual microscopy to cytoscreening, perspectives based on questionnaire to Japanese cytotechnologists. Diagn Pathol 2008;3 Suppl 1:S15.
7. Dee FR. Virtual microscopy in pathology education. Hum Pathol 2009;40:1112-21.
8. Nielsen PS, Lindebjerg J, Rasmussen J, Starklint H, Waldstrøm M, Nielsen B. Virtual microscopy: An evaluation of its validity and diagnostic performance in routine histologic diagnosis of skin tumors. Hum Pathol

2010;41:1770-6.
9. Merk M, Knuechel R, Perez-Bouza A. Web-based virtual microscopy at the RWTH Aachen University: Didactic concept, methods and analysis of acceptance by the students. Ann Anat 2010;192:383-7.
10. Szymas J, Lundin M. Five years of experience teaching pathology to dental students using the WebMicroscope. Diagn Pathol 2011;6 Suppl 1:S13.
11. Evered A, Dudding N. Accuracy and perceptions of virtual microscopy compared with glass slide microscopy in cervical cytology. Cytopathology 2011;22:82-7.
12. Wright AM, Smith D, Dhurandhar B, Fairley T, Scheiber-Pacht M, Chakraborty S, *et al.* Digital slide imaging in cervicovaginal cytology: A pilot study. Arch Pathol Lab Med 2013;137:618-24.
13. House JC, Henderson-Jackson EB, Johnson JO, Lloyd MC, Dhillon J, Ahmad N, *et al.* Diagnostic digital cytopathology: Are we ready yet? J Pathol Inform 2013;4:28.
14. Dee FR, Donnelly A, Radio S, Leaven T, Zaleski MS, Kreiter C. Utility of 2-D and 3-D virtual microscopy in cervical cytology education and testing. Acta Cytol 2007;51:523-9.
15. Stewart J 3rd, Miyazaki K, Bevans-Wilkins K, Ye C, Kurtycz DF, Selvaggi SM. Virtual microscopy for cytology proficiency testing: Are we there yet? Cancer 2007;111:203-9.
16. Huisman A, Looijen A, van den Brink SM, van Diest PJ. Creation of a fully digital pathology slide archive by high-volume tissue slide scanning. Hum Pathol 2010;41:751-7.
17. Wilbur DC. Digital cytology: Current state of the art and prospects for the future. Acta Cytol 2011;55:227-38.
18. Khalbuss WE, Pantanowitz L, Parwani AV. Digital imaging in cytopathology. Patholog Res Int 2011;2011:264683.
19. Donnelly AD, Mukherjee MS, Lyden ER, Bridge JA, Lele SM, Wright N, *et al.* Optimal z-axis scanning parameters for gynecologic cytology specimens. J Pathol Inform 2013;4:38.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.