Journal of the American Society of Cytopathology (2019) 8, 18–26

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.jascyto.org/

ORIGINAL ARTICLE

Variability among observers utilizing the CellSolutions BestCyte Cell Sorter imaging system for the assessment of urinary tract cytology specimens

Elise Gelwan, MD[a], M. Lisa Zhang, MD[b], Derek B. Allison, MD[a], Morgan L. Cowan, MD[a], Juliana DeLuca, MS, CT (ASCP)[a], J. Judd Fite, MD, MBA[a], Sintawat Wangsiricharoen, MD[a], Bonnie Williamson, MS, CT (ASCP)[a], Amy Zhou, MD[a], Christopher J. VandenBussche, MD, PhD[a,c,*]

[a] *Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore, Maryland*
[b] *Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts*
[c] *Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, Maryland*

Received 26 April 2018; received in revised form 29 September 2018; accepted 8 October 2018

KEYWORDS
Urine;
Urothelial carcinoma;
Urothelial neoplasia;
Bladder cancer;
Digital imaging

Introduction Image analysis systems are not currently commonly used for evaluating urinary cytology specimens. We evaluated whether the BestCyte Cell Sorter (CellSolutions, Greensboro, NC) imaging system can reliably identify atypical cells in urinary cytology specimens.

Methods Fifty-three consecutive urine cytology specimens underwent 2 preparations: one slide using SurePath (SP; BD Diagnostics, Sparks, MD)™ for routine clinical evaluation, and a second slide using the CellSolutions F50 system for analysis by the BestCyte Cell Sorter (BCCS) scanning system. Eight observers reviewed atypical cells flagged by BCCS and assigned a BCCS diagnosis to each of the 53 specimens. The observers also blindly reviewed the SP preparation (when available) and assigned an SP diagnosis. The SP diagnoses given by one "expert" observer was considered as a reference diagnosis.

Results There was fair-to-moderate agreement among observers for identifying any atypia and high-grade atypia (Fleiss kappa: 0.417 and 0.338, respectively) using BCCS. Review of SP preparations had slightly better agreement (Fleiss kappa: 0.558 and 0.564, respectively). Intraobserver agreement between the two methods varied greatly between individuals (Cohen's kappa range: 0.260 to 0.647). When a consensus

*Corresponding author: Christopher J. VandenBussche, MD, PhD, Department of Pathology, The Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, MD 21287; Tel.: (410) 955-1180; Fax: (410) 614-9556.
E-mail address: cjvand@jhmi.edu (C.J. VandenBussche).

https://doi.org/10.1016/j.jasc.2018.10.001

diagnosis could be reached among the observers for cases with surgical follow-up, the consensus diagnosis was concordant in 11 of 12 instances, with one instance being a one-step discrepancy.

Conclusions Specimen review by BCCS resulted in slightly greater interobserver variability than review of routine SP preparations. This may have been due to variations in observer experience and comfort with the use of a digital imaging system, which is further suggested by the wide range of intraobserver agreement among individuals.

Introduction

Urinary tract cytology (UTC) is commonly used to screen patients who present with unexplained hematuria and to monitor patients with a history of urothelial carcinoma.[1] Despite its relatively high sensitivity for the detection of high-grade urothelial carcinoma (HGUC), however, urinary cytology has suffered from multiple limitations: low sensitivity for the detection of low-grade urothelial neoplasms, significant inter- and intraobserver diagnostic variability, and difficulty in defining the criteria for and clinical significance of the equivocal "atypical" category.[2] To address these issues, a standardized set of guidelines for classifying urinary cytology specimens was created: The Paris System (TPS) for Reporting Urinary Cytology.[3] This system utilizes the nuclear-to-cytoplasmic (N:C) ratio as a prominent feature to distinguish between the 4 diagnostic categories, along with several additional nuclear features (nuclear hyperchromasia, irregular nuclear membrane contour, and coarse/clumped chromatin).[2] Despite the development of more standardized cytomorphologic criteria, subjectivity continues to be an issue when evaluating urinary cytology specimens. For instance, using quantitative image analysis to precisely measure N:C ratios of urothelial cells, several studies have shown that morphologists have difficulty visually estimating the N:C ratio, particularly at ratios close to the cutoff values used in TPS (0.5 and 0.7).[4-6] This limitation is especially problematic given the findings of a separate study that showed that small changes in the N:C ratio between 0.4 and 0.6, as determined by image analysis, significantly affect the sensitivity and specificity of predicting HGUC on follow-up biopsy.[7] These findings raise the possibility that computer-aided image analysis (either as a primary screening device or as an adjunct to visual screening) may be diagnostically beneficial in urinary cytology.

In this study, we evaluate the use of the BestCyte Cell Sorter (BCCS; CellSolutions, Greensboro, NC) imaging system on UTC specimens.[8] BCCS was developed as an imaging system for Papanicolaou test specimens; thus, the purpose of our study was to determine: (1) whether the BCCS algorithm could identify clinically relevant atypia in a given UTC specimen; (2) the interobserver variability in the interpretation of BCCS data; and (3) the performance of diagnoses based on BCCS in relation to follow-up biopsy results.

Materials and methods

Case selection and slide preparation

Approval to conduct this study with waived consent was obtained from the institutional review board. Fifty-three consecutive urine cytology specimens from 50 patients collected at the Johns Hopkins Hospital were used for the study. Only specimens with sufficient volume to be split into a clinical and experimental specimen were used. One slide was prepared using SurePath (SP) liquid-based preparation for clinical evaluation, and a second de-identified slide was prepared using the CellSolutions F50 liquid-based cytology slide preparation system. These de-identified slides were subsequently digitally scanned and analyzed using the BCCS scanning system.

Fifty-one SP specimens were available for review at the time of this study. One board-certified cytopathologist with expertise in UTC (CJV) reviewed the specimens and assigned each one a diagnosis; this was considered the reference diagnosis for each specimen. The 7 other observers also reviewed the 51 SP specimens and rendered an "SP diagnosis" for each specimen. The 7 observers included 3 cytopathology fellows, 2 cytotechnologists, and 2 pathology residents who had finished at least 2 months of cytopathology rotations.

The digital images of the scanned slides were prepared using the CellSolutions method and were analyzed by the BCCS imaging analysis software, which categorizes cells based on various cytomorphologic criteria using proprietary technology. The software places the cells into panels based on the type of atypia identified, which are then available for subsequent review by a pathologist. The 8 observers independently reviewed the panels of atypical cells for each of the 53 de-identified cases without time limitation. Based on their review of the digital images of the flagged cells, and without knowledge of the SP diagnosis, the observers recorded a "BCCS diagnosis" for each case.

During the study, our laboratory formally transitioned from our previous nomenclature system to The Paris System for Reporting Urinary Cytology, both of which utilize 2 indeterminate diagnostic categories.[9,10] Because of our institution's involvement in the development of TPS since 2013, our faculty have integrated TPS concepts into their practice and instruction and thus the terminology between the nomenclature systems became interchangeable during this period.[11] For simplicity, any diagnoses rendered using the previous nomenclature system have been converted to TPS terminology in this study. Although no special instructional or educational material was provided to observers prior to the study, trainees and cytotechnologists at our institution are well trained regarding TPS.

Data analysis

For each specimen, a consensus BCCS diagnosis was determined using the diagnoses of 7 observers. The eighth observer (CJV), whose diagnosis was used as the reference diagnosis, did not contribute to the consensus BCCS diagnosis. A consensus BCCS diagnosis was reached if at least 4 of the 7 observers agreed on one of the following categories: Negative for High Grade Urothelial Carcinoma (NHGUC), Atypical Urothelial Cells (AUC), and Suspicious for High Grade Urothelial Carcinoma (SHGUC)/High Grade Urothelial Carcinoma (HGUC). When available, follow-up biopsy results and other clinical data (eg, cystoscopic findings) were extracted from the pathology database and electronic medical record.

To determine variations between observer diagnoses, a Fleiss generalized kappa was calculated. To compare intrarobserver variation between BCCS and SP diagnoses, a Cohen's kappa was calculated for each observer. Statistical analyses were performed using R version 3.2.2.

Results

Interobserver variability in the assessment of digitized slides

A total of 415 BCCS diagnoses by 8 different observers were recorded for 53 specimens based on examination of the atypical cells selected by the BCCS system from scanned slides (Figs. 1-4). These BCCS diagnoses were distributed as follows: NHGUC, 225 observations (55.7%); AUC, 90 observations (21.7%); SHGUC, 47 observations (11.3%); and HGUC, 47 observations (11.3%). There were 8 instances in which an observer felt that the digital images were

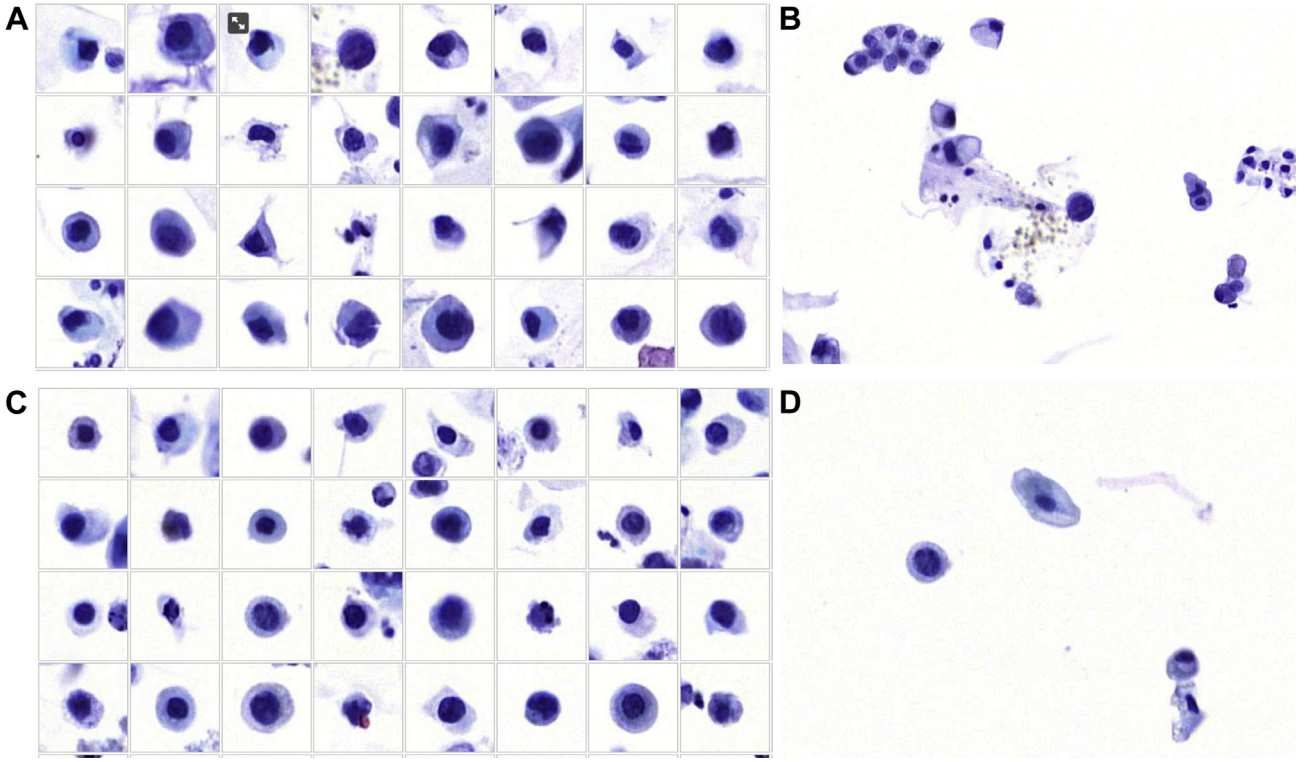

Figure 1 A and C, Two subsets of cells identified by BCCS based on increased nuclear-to-cytoplasmic (N:C) ratios in 2 urinary tract specimens from two different patients, both diagnosed as high-grade urothelial carcinoma (HGUC) on conventional SurePath preparations. The consensus digital diagnosis among observers was also HGUC for both cases. B and D, Expanded fields of interest containing a malignant cell from the panels in A and C, respectively. The cells are small but contain dark nuclei with irregular borders and N:C ratios above 0.7.

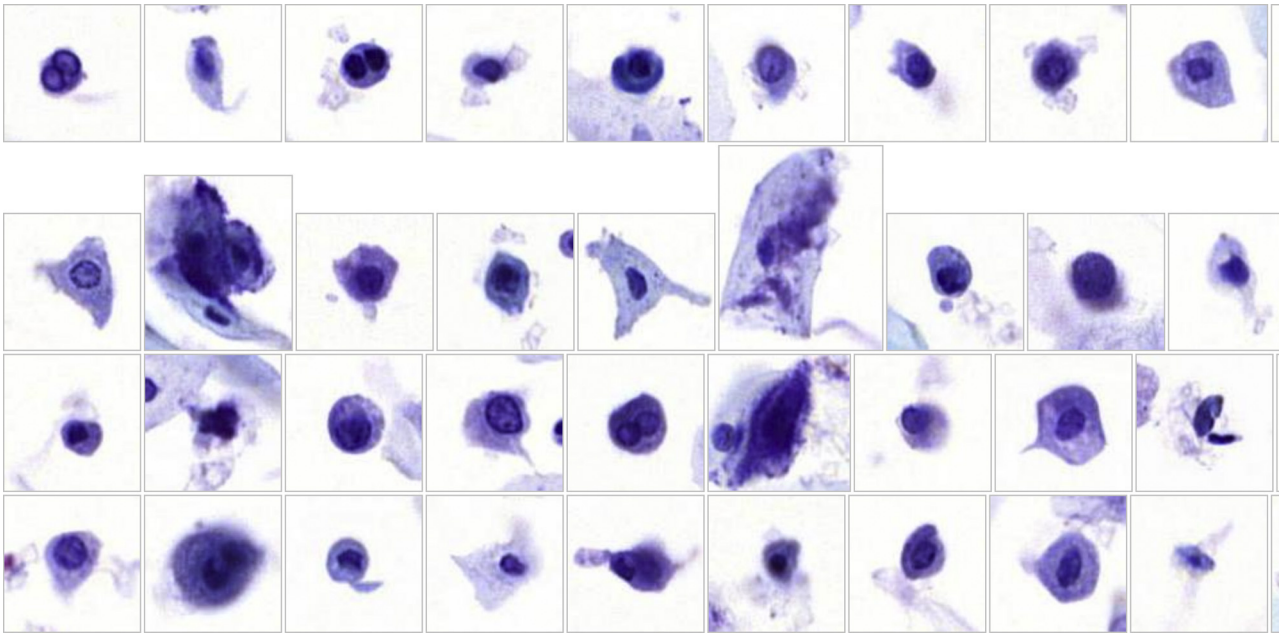

Figure 2 A subset of cells identified by BestCyte Cell Sorter (BCCS) based on increased nuclear-to-cytoplasmic (N:C) ratios in a urinary tract specimen diagnosed as Atypical Urothelial Cells (AUC) on a conventional SurePath specimen. The consensus digital diagnosis among observers was Negative for High-Grade Urothelial Carcinoma (NHGUC). The patient had a low-grade urothelial carcinoma on follow-up.

of inadequate quality to confidently make a diagnosis; this was distributed over 2 slides/cases.

Interobserver variability for BCCS diagnosis was determined using a Fleiss kappa for instances in which concordance required agreement on the exact diagnostic category ($K = 0.273$), agreement on whether a specimen contained atypical cells (NHGUC versus AUC/SHGUC/HGUC; $K = 0.417$), and agreement on whether a specimen was suspicious or diagnostic of HGUC (NHGUC/AUC versus SHGUC/HGUC; $K = 0.338$).

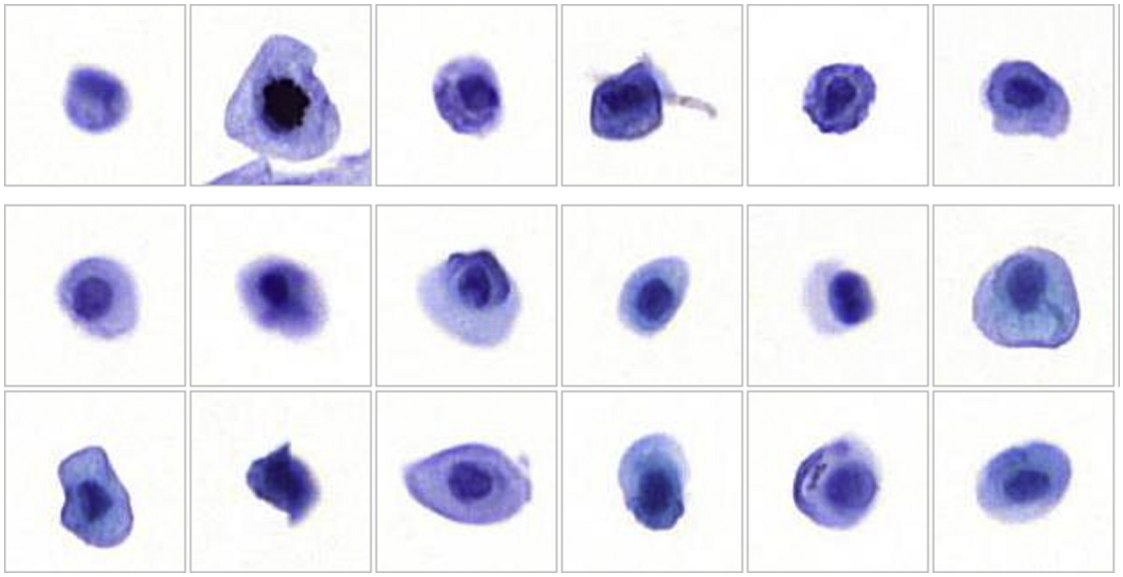

Figure 3 A subpanel of cells identified by BestCyte Cell Sorter (BCCS) based on increased nuclear-to-cytoplasmic (N:C) ratios in a urinary tract specimen diagnosed as Negative for High-Grade Urothelial Carcinoma (NHGUC) on a conventional SurePath specimen. The consensus digital diagnosis among observers was Negative for High-Grade Urothelial Carcinoma (NHGUC). The patient did not have disease over a 2-year follow-up period.

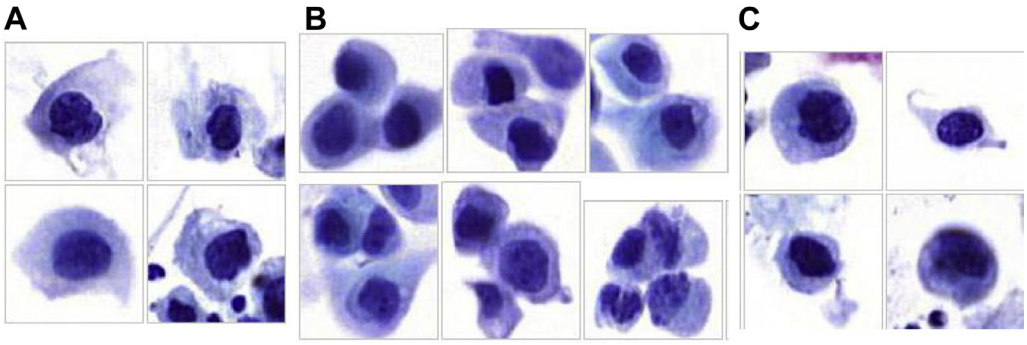

Figure 4 Examples of high-grade urothelial cells identified by BestCyte Cell Sorter (BCCS) in a malignant specimen based on different algorithms: A, Examples of cells identified as "atypical" by the digital image analysis algorithm. B, Examples of cells in clusters as identified by the digital image analysis algorithm. C, Examples of cells with high N:C ratios as identified by the digital image analysis algorithm.

Intraobserver variability in the assessment of digitized slides compared with conventional light microscopy

Of the 53 UTC specimens prepared for BCCS, 51 had corresponding SP glass slides available for review at the time of this study (Table 1). Interobserver variability for the SP diagnosis was determined using a Fleiss kappa for instances in which concordance required agreement on the exact diagnostic category (K = 0.397), agreement on whether a specimen contained atypical cells (NHGUC versus AUC/SHGUC/HGUC; K = 0.558), and agreement on whether a specimen was suspicious or diagnostic of HGUC (NHGUC/AUC versus SHGUC/HGUC; K = 0.564). These values were compared with Fleiss kappa values calculated for BCCS diagnoses (Table 2). Intraobserver variability between an observer's assessment of the digital slides (BCCS diagnosis) and glass slides (SP diagnosis) was determined using Cohen's kappa, which was calculated for agreement on the exact diagnostic category as well as for whether a specimen was suspicious or diagnostic of HGUC (NHGUC/AUC versus SHGUC/HGUC; Table 3).

Comparison between consensus BCCS diagnoses and reference diagnosis

To determine the performance of the BCCS diagnoses, the SP diagnosis provided by one of the study's observers (CJV; board certified in cytopathology and with expertise in UTC) was considered the "reference diagnosis" for each specimen with which the BCCS diagnosis could be compared. The distribution of reference diagnoses was as follows: NHGUC, 28 (55%); AUC, 13 (25%); SHGUC, 4 (7.8%); HGUC, 6 (12%).

In order to summarize the BCCS diagnoses from multiple observers, a consensus BCCS diagnosis was determined from the BCCS diagnoses of the other 7 observers. To prevent bias, the eighth observer providing the reference diagnosis was not included in the consensus diagnosis. The distribution of consensus diagnoses was as follows: NGUC, 24 (47%); AUC, 6 (12%); SHGUC, 4 (7.8%); HGUC, 2 (3.9%). For 10 (20%) specimens, a consensus was not reached among the 7 observers (Table 3).

For the cases in which the reference diagnosis was SHGUC or HGUC, the majority (6 out of 10 cases) had a consensus BCCS diagnosis of SHGUC/HGUC, with 3 cases

Table 1 Comparison of BestCyte CellSorter diagnosis to SurePath diagnosis made by the 8 observers.

BCCS diagnosis	SurePath diagnosis, n (%)				Total, n (%)
	NHGUC	AUC	SHGUC	HGUC	
NHGUC	165 (41.8)	50 (12.7)	6 (1.5)	4 (1.0)	225 (57.0)
AUC	37 (9.4)	28 (7.1)	12 (3.0)	8 (2.0)	85 (21.5)
SHGUC	8 (2.0)	15 (3.8)	14 (3.5)	9 (2.3)	46 (11.6)
HGUC	3 (0.8)	6 (1.5)	7 (1.8)	23 (5.8)	39 (9.9)
Total	213 (53.9)	99 (25.1)	39 (9.9)	44 (11.1)	395 (100)

Abbreviations: BCCS, BestCyte Cell Sorter; NHGUC, negative for high-grade urothelial carcinoma; AUC, atypical urothelial cells; SHGUC, suspicious for high-grade urothelial carcinoma; HGUC, high-grade urothelial carcinoma.
A total of 51 SurePath slides were available for review, making the total possible observation pairs 408; pairs in which one observation was "non-diagnostic" or not provided by the observer are excluded from this table.

Table 2 Interobserver variation among the 8 observers.

Diagnosis	Fleiss kappa (BCCS)	Fleiss kappa (SP)
NHGUC versus AUC versus SHGUC versus HGUC	0.273	0.397
NHGUC/AUC versus SHGUC/HGUC	0.338	0.564
NHGUC versus AUC/SHGUC/HGUC	0.417	0.558

Abbreviations: BCCS, BestCyte CellSorter; SP, SurePath; NHGUC, negative for high-grade urothelial carcinoma; AUC, atypical urothelial cells; SHGUC, suspicious for high-grade urothelial carcinoma; HGUC, high-grade urothelial carcinoma.
The 8 observers reviewed urinary tract specimens using BCCS (53 specimens) and SP preparations by conventional microscopy (51 specimens).

failing to reach a consensus. The greatest number of discrepancies were one-step discrepancies between AUC and NHGUC, with 6 of 13 (46%) AUC reference diagnoses being downgraded to NHGUC on the consensus BCCS diagnosis.

Comparison of cytologic diagnosis with surgical pathology and clinical follow-up

To determine whether any of the discrepancies between the reference diagnoses and the consensus BCCS diagnoses could be clinically significant, the results of follow-up biopsies were reviewed (Table 4). Over a period of 23 months after collection of the cytology specimens, biopsy results were available for 19 out of 53 cases (35.8%) (Table 5). Two specimens with surgical pathology follow-up did not have glass slides available for review and thus were not assigned a reference diagnosis.

There were 7 biopsy-proven cases of HGUC; when a consensus could be reached, the consensus diagnosis was similar to the reference diagnosis, with the exception of one NHGUC diagnosis that corresponded to a reference diagnosis of AUC. For the remaining 12 cases with LGUC or benign follow-up, consensus could be reached for 9 cases, and corresponded with the reference diagnosis in all instances. This included 2 instances in which the reference diagnosis was HGUC and the subsequent biopsies contained denudation, a finding sometimes associated with undersampled carcinoma in situ.

Table 3 Intraobserver variation between the BestCyte CellSorter and SurePath diagnoses for each of the 8 observers.

Observer	Cohen's kappa (NHGUC versus AUC versus SHGUC versus HGUC)	Cohen's kappa (NGHUC/AUC versus SHGUC/HGUC)
1	0.220	0.299
2	0.344	0.583
3	0.307	0.580
4	0.477	0.647
5	0.220	0.260
6	0.176	0.459
7	0.293	0.447
8	0.400	0.609

Abbreviations: BCCS, BestCyte CellSorter; SP, SurePath; NHGUC, negative for high-grade urothelial carcinoma; AUC, atypical urothelial cells; SHGUC, suspicious for high-grade urothelial carcinoma; HGUC, high-grade urothelial carcinoma.
The 8 observers reviewed urinary tract specimens of 51 urinary tract specimens reviewed using BCCS and SP preparations by conventional microscopy.

Discussion

The practice of urinary cytology has historically suffered from poor sensitivity for detecting low-grade urothelial neoplasms and poor inter- and intraobserver reproducibility. These limitations have lead several groups in the past to pursue the development of computer-based image analysis systems to increase the objectivity of diagnoses.[12-23] The earliest work in this area was pioneered by Dr Leopold Koss, who published several papers between 1975 and 1989 on a computer image analysis system for Papanicolaou-stained, voided urine specimens.[12-17,23] The system involved a microscope with a mounted video scanning system and a computer; morphologic features were extracted from scanned images of a "training" set of cells with established cytologic diagnoses, and these features were then utilized by the computer software to construct cell (and whole specimen) classification rules/algorithms.[14] Several

Table 4 Comparison of consensus BCCS diagnosis with the reference diagnosis.

Consensus BCCS diagnosis	Reference diagnosis[a], n (%)				Total
	NHGUC	AUC	SHGUC	HGUC	
NHGUC	18 (35)	6 (12)	0 (0.0)	0 (0.0)	24 (47)
AUC	2 (3.9)	3 (5.9)	1 (2.0)	0 (0.0)	6 (12)
SHGUC/HGUC	1 (2.0)	2 (3.9)	0 (0.0)	6 (12)	9 (18)
Inadequate	2 (3.9)	0 (0.0)	0 (0.0)	0 (0.0)	2 (3.9)
No consensus	5 (9.8)	2 (3.9)	3 (5.9)	0 (0.0)	10 (20)
Total	28 (55)	13 (25)	4 (7.8)	6 (12)	51 (100)

Abbreviations: BCCS, BestCyte CellSorter; NHGUC, negative for high-grade urothelial carcinoma; AUC, atypical urothelial cells; SHGUC, suspicious for high-grade urothelial carcinoma; HGUC, high-grade urothelial carcinoma.
[a]The reference diagnosis was the diagnosis rendered on re-review of the original SurePath slides by a single board-certified cytopathologist for the 51 specimens available for immediate review.

Table 5 Consensus digital diagnoses and glass slide diagnoses for 19 cases with available follow-up biopsies.

Follow-up biopsy result	Reference diagnosis	Consensus BCCS diagnosis	Majority digital diagnosis
HGUC	AUC	NHGUC	
HGUC	HGUC	SHGUC/HGUC	
HGUC	NHGUC	No consensus	NHGUC (3)
HGUC	HGUC	SHGUC/HGUC	
HGUC	SHGUC	No consensus	SHGUC/HGUC (3)
CIS	N/A	SHGUC/HGUC	
CIS	AUC	No consensus	NHGUC (3)
LGUC	AUC	AUC	
LGUC	NHGUC	NHGUC	
LGUC	AUC	AUC	
LGUC	NHGUC	No consensus	AUC (3)
LGUC	NHGUC	No consensus	AUC (3)
Benign/denuded	NHGUC	NHGUC	
Benign/denuded	HGUC	SHGUC/HGUC	
Benign/denuded	HGUC	SHGUC/HGUC	
Benign	N/A	No consensus	AUC (3), NHGUC (3)
Benign	AUC	AUC	
Benign	NHGUC	NHGUC	
Benign	NHGUC	NHGUC	

Abbreviations: NHGUC, negative for high-grade urothelial carcinoma; AUC, atypical urothelial cells; SHGUC, suspicious for high-grade urothelial carcinoma; HGUC, high-grade urothelial carcinoma; LGUC, low-grade urothelial carcinoma; CIS, carcinoma in situ; N/A, not available (SurePath slide unavailable for review).

more recent studies have combined extraction of morphometric properties of urothelial cells (by image analysis) with neural networks to allow for discrimination of benign from malignant urinary specimens.[20,24] In a large prospective study of 1455 UTC specimens, van Hemel et al. utilized the ThinPrep Imaging System (TIS) and found substantial agreement (kappa score: 0.77) between conventional screening and TIS.[25] Unfortunately, these efforts have not yet resulted in an accepted method for computer-aided examination of UTC specimens.

There was greater agreement among our observers for SP diagnoses than for BCCS diagnoses (Table 2). Although one would expect increased agreement using the BCCS system, which presents the same images of atypical cells to each observer, a decrease in agreement may have been caused by differences in experience and comfort among observers for using a digital system to review cells, as well as a different preparation method for the creation of slides. This is further suggested by the fact that some observers (for instance, observers 1 and 5 in Table 3) had significantly less agreement between diagnoses made on SP specimens versus BCCS. Interobserver agreement regarding the identification of atypia (NHGUC versus AUC/SHGUC/HGUC) or high-grade atypia (NHGUC/AUC versus SHGUC/HGUC) was greater than agreement regarding the exact diagnostic category (NHGUC versus AUC versus SHGUC versus HGUC). For the BCCS system, however, agreement regarding the identification of any atypia (K: 0.417; Table 2) was better than agreement regarding the identification of high-grade atypia (K: 0.338); this was not true for diagnoses made on SP (K: 0.558 versus 0.564). This indicates that whereas observers tended to agree on the identification of atypia using BCCS, there was more variation in determining the degree of atypia. Because in many instances urologists treat low-risk indeterminate diagnoses (eg, AUC) similarly to benign diagnoses (eg, NHGUC), such variation has the potential to impact patient management.[26,27]

When the BCCS diagnoses of all observers were summarized as a consensus diagnosis and compared with a reference SurePath diagnosis, the greatest number of discrepancies were instances in which the reference diagnosis was called AUC and the consensus BCCS diagnosis was NHGUC. In fact, approximately half of the AUC reference diagnoses were downgraded to NHGUC on the BCCS consensus diagnosis, indicating that either the atypical cells were not found in the BCCS preparation, did not appear as atypical when presented by the software, or were not detected as atypical by the software. Another contributing factor is likely the continued variability between observers regarding the threshold for an atypical diagnosis, which may be reduced but not completely eliminated by TPS.[28]

It is likely that the BCCS algorithms to identify high grade squamous intraepithelial lesions (HSIL) were the most useful in detecting HGUC cells, given the cytomorphologic similarity between these two entities (as opposed to low grade intraepithelial lesions). Typically, software algorithms are designed with a low threshold for identifying atypical cells, thus allowing the cytotechnologist or pathologist to make the final determination. In this case, however, the higher threshold for atypia used by BCCS may have been appropriate, since only one of the specimens that was downgraded from AUC to NHGUC on consensus BCCS

diagnosis came from patients with HGUC on follow-up biopsy and the clinical follow-up (including cystoscopy and imaging findings) was largely unremarkable in the vast majority of these downgraded cases, explaining the lack of follow-up biopsies for most of them.

One important limitation to this study is that the slides analyzed by BCCS were not exactly the same as the SP slides analyzed by conventional microscopy, as the specimens were split for each preparation. It is therefore possible that atypical urothelial cells could have been present on only one preparation. This limitation could theoretically account for some of the cases that showed a diagnostic discrepancy of NHGUC versus AUC when comparing the 2 methods.

The preparation required by BCCS differs from that used for SurePath, which is the method used in our laboratory to prepare routine urinary tract specimens. This may have caused additional discrepancies, as our faculty, trainees, and cytotechnologists are less familiar with interpreting the preparation used by BCCS. This obstacle may also explain the variability among observers for certain cases. Although it is difficult to measure the impact of this limitation on the current study, in practice it could be overcome with additional training and experience.

Another limitation is that the BCCS software was specifically designed to analyze images of Papanicolaou test specimens rather than UTC specimens. Clearly, the cytomorphologic criteria used by pathologists to diagnose squamous lesions of the uterine cervix are different from those used to screen urinary cytology specimens for high-grade urothelial carcinoma. It is interesting, however, that even despite these differences, the system performed reasonably well overall. A future area of investigation could involve making modifications to the image analysis software to incorporate cytomorphologic criteria that are considered important in urinary cytology, such as those recommended by TPS, to determine whether this results in improved performance. We recently demonstrated that digital calculations of N:C ratio in urine cytology specimens were useful in predicting the likelihood of HGUC on follow-up. Very subtle changes in the N:C ratio between 0.4 and 0.6 (as calculated by digital image analysis) significantly affected the sensitivity and specificity of predicting HGUC.[7] As these changes in N:C ratio would likely be difficult to assess subjectively with the human eye, incorporation of these data into digital image analysis algorithms could be very helpful and could potentially improve performance.

Certain factors may have additionally contributed to a disadvantage in the BCCS arm. Because our faculty and trainees do not use digital imaging systems on a regular basis, inexperience may have contributed to suboptimal diagnoses. The dots placed by cytotechnologist on the SurePath preparations could not be removed because of our laboratory policy, thus the glass slides were pre-screened by trained personnel.

This study is the first to assess the BCCS system to analyze UTC specimens. Considering that the software is designed to analyze Papanicolaou test specimens, the system performed well. In instances in which consensus could be reached, the system did not fail to identify atypical cells in patients who would subsequently be diagnosed with HGUC except in one instance in which reference diagnosis was AUC, a diagnosis that may not have altered the patient's clinical management. BCCS also reduced the number of overall atypical diagnoses in patients who were not diagnosed with HGUC over the 2-year follow-up period. This finding is likely due to the algorithm used, which, although designed to have a low threshold for identifying atypical cells in Papanicolaou test specimens, appears to be more restrictive than our practice of identifying atypical urothelial cells in UTC specimens. Modifications of such a system utilizing cytomorphologic criteria as defined by TPS could potentially reduce specimen screening time and reduce irrelevant atypical diagnoses by only identifying cells with relevant features.

Disclosures

Dr. VandenBussche serves as a consultant for Personal Genome Diagnostics. A CellSolutions representative prepared and scanned the slides used for digital image analysis in this study. The authors have no other relevant financial disclosures or conflicts of interest.

Funding sources

None.

Author contributions

Elise Gelwan: Data curation, formal analysis, investigation, project administration, writing − original draft, writing − review and editing. M. Lisa Zhang: Data curation, formal analysis. Derek B. Allison: Investigation, writing − review and editing. Morgan L. Cowan: Investigation, writing − review and editing. Juliana DeLuca: Investigation. J. Judd Fite: Investigation. Sintawat Wangsiricharoen: Investigation. Bonnie Williamson: Investigation. Amy Zhou: Investigation. Christopher J. VandenBussche: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, visualization, writing − review and editing.

References

1. Owens CL, Vandenbussche CJ, Burroughs FH, Rosenthal DL. A review of reporting systems and terminology for urine cytology. *Cancer Cytopathol*. 2013;121:9−14.
2. Barkan GA, Wojcik EM, Nayar R, et al. The Paris System for Reporting Urinary Cytology: the quest to develop a standardized terminology. *Acta Cytol*. 2016;60:185−197.
3. VandenBussche CJ. A review of the Paris System for Reporting Urinary Cytology. *Cytopathology*. 2016;27:153−156.
4. Zhang ML, Guo AX, VandenBussche CJ. Morphologists overestimate the nuclear-to-cytoplasmic ratio. *Cancer Cytopathol*. 2016;124:669−677.
5. Layfield LJ, Esebua M, Frazier SR, et al. Accuracy and reproducibility of nuclear/cytoplasmic ratio assessments in urinary cytology specimens. *Diagn Cytopathol*. 2017;45:107−112.
6. Vaickus LJ, Tambouret RH. Young investigator challenge: the accuracy of the nuclear-to-cytoplasmic ratio estimation among trained morphologists. *Cancer Cytopathol*. 2015;123:524−530.
7. Hang JF, Charu V, Zhang ML, VandenBussche CJ. Digital image analysis supports a nuclear-to-cytoplasmic ratio cutoff value of 0.5 for atypical urothelial cells. *Cancer*. 2017;125:710−716.
8. Delga A, Goffin F, Kridelka F, Maree R, Lambert C, Delvenne P. Evaluation of CellSolutions BestPrep(R) automated thin-layer liquid-based cytology Papanicolaou slide preparation and BestCyte(R) cell sorter imaging system. *Acta Cytol*. 2014;58:469−477.
9. VandenBussche CJ, Sathiyamoorthy S, Owens CL, Burroughs FH, Rosenthal DL, Guan H. The Johns Hopkins Hospital template for urologic cytology samples: parts II and III: improving the predictability of indeterminate results in urinary cytologic samples: an outcomes and cytomorphologic study. *Cancer Cytopathol*. 2013;121:21−28.
10. Rosenthal DL, Vandenbussche CJ, Burroughs FH, Sathiyamoorthy S, Guan H, Owens C. The Johns Hopkins Hospital template for urologic cytology samples: part I-creating the template. *Cancer Cytopathol*. 2013;121:15−20.
11. VandenBussche CJ, Allison DB, Gupta M, Ali SZ, Rosenthal DL. A 20-year and 46,000-specimen journey to Paris reveals the influence of reporting systems and passive peer feedback on pathologist practice patterns. *Cancer Cytopathol*. 2018;126:381−389.
12. Sherman A, Koss LG, Adams S, et al. Bladder cancer diagnosis by image analysis of cells in voided urine using a small computer. *Anal Quant Cytol*. 1981;3:239−249.
13. Sherman AB, Koss LG, Adams SE. Interobserver and intraobserver differences in the diagnosis of urothelial cells. Comparison with classification by computer. *Anal Quant Cytol*. 1984;6:112−120.
14. Sherman AB, Koss LG, Wyschogrod D, Melder KH, Eppich EM, Bales CE. Bladder cancer diagnosis by computer image analysis of cells in the sediment of voided urine using a video scanning system. *Anal Quant Cytol Histol*. 1986;8:177−186.
15. Koss LG, Bartels PH, Bibbo M, et al. Computer analysis of atypical urothelial cells. I. Classification by supervised learning algorithms. *Acta Cytol*. 1977;21:247−260.
16. Melder KK, Koss LG. Automated image analysis in the diagnosis of bladder cancer. *Appl Opt*. 1987;26:3367−3372.
17. Wong EK, Liang EH, Lin EK, Simmons DA, Koss LG. A selective mapping algorithm for computer analysis of voided urine cell images. *Anal Quant Cytol Histol*. 1989;11:203−210.
18. van der Poel HG, Witjes JA, Schalken JA, Debruyne FM. Automated image analysis for bladder cancer. *Urol Res*. 1998;26:1−5.
19. van der Poel HG, Witjes JA, van Stratum P, Boon ME, Debruyne FM, Schalken JA. Quanticyt: karyometric analysis of bladder washing for patients with superficial bladder cancer. *Urology*. 1996;48:357−364.
20. Pantazopoulos D, Karakitsos P, Iokim-Liossi A, Pouliakis A, Dimopoulos K. Comparing neural networks in the discrimination of benign from malignant lower urinary tract lesions. *Br J Urol*. 1998;81:574−579.
21. Borkowski T, Monika Dulewicz A, Borkowski A, Pietka D, Radziszewski P. Clinical value of digital image analysis in the diagnosis of urinary bladder cancer, particularly in aggressive tumors: a preliminary report. *Pol J Pathol*. 2016;67:122−129.
22. Dulewicz A, Pietka BD, Jaszczak P, et al. Computer identification of neoplastic urothelial nuclei from the bladder. *Anal Quant Cytol Histol*. 2001;23:321−329.
23. Koss LG, Eppich EM, Melder KH, Wersto R. DNA cytophotometry of voided urine sediment. Comparison with results of cytologic diagnosis and image analysis. *Anal Quant Cytol Histol*. 1987;9:398−404.
24. Vriesema JL, van der Poel HG, Debruyne FM, Schalken JA, Kok LP, Boon ME. Neural network-based digitized cell image diagnosis of bladder wash cytology. *Diagn Cytopathol*. 2000;23:171−179.
25. van Hemel BM, Haarsma JG, Ruitenbeek T, Groen HJ, Suurmeijer AJ. Application of the ThinPrep imaging system in urine cytology: a prospective study. *Cancer Cytopathol*. 2013;121:410−414.
26. Gupta M, VandenBussche CJ, Bivalacqua TJ. Urinary cytology and The Paris System for Reporting Urinary Cytology: implications for urological management. *Cytopathology*. 2018;29:368−370.
27. Fite JJ, Rosenthal DL, VandenBussche CJ. When words matter: a "suspicious" urinary tract cytology diagnosis improves patient follow-up among nonurologists. *Cancer Cytopathol*. 2018;126:282−288.
28. Kurtycz DF, Barkan GA, Pavelec DM, et al. Paris Interobserver Reproducibility Study(PIRST). *J Am Soc Cytopath*. 2018;7:174−184.