Assessing the Clinical Practicability of Artificial Intelligence Assisted Reporting Using the Paris System for Urinary Cytology

PRESENTED BY

Tien-Jen Liu, M.D.

USCAP 113TH ANNUAL MEETING

BRINGING EDUCATION TO LIFE

USCAP

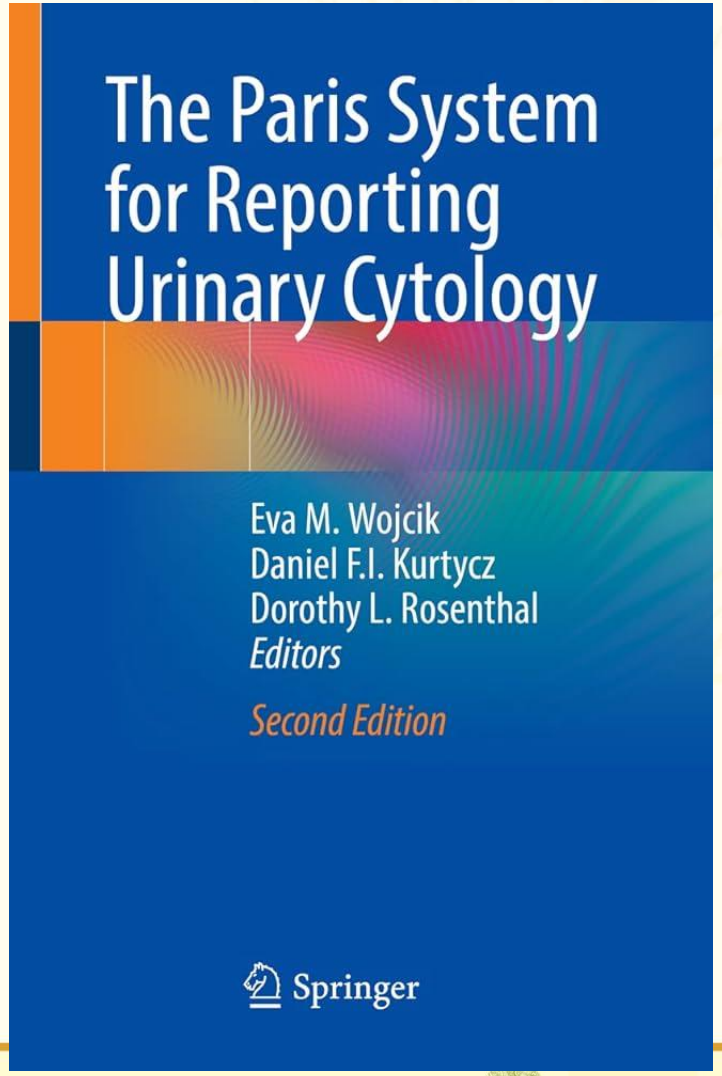
UNITED STATES AND CANADIAN ACADEMY OF PATHOLOGY

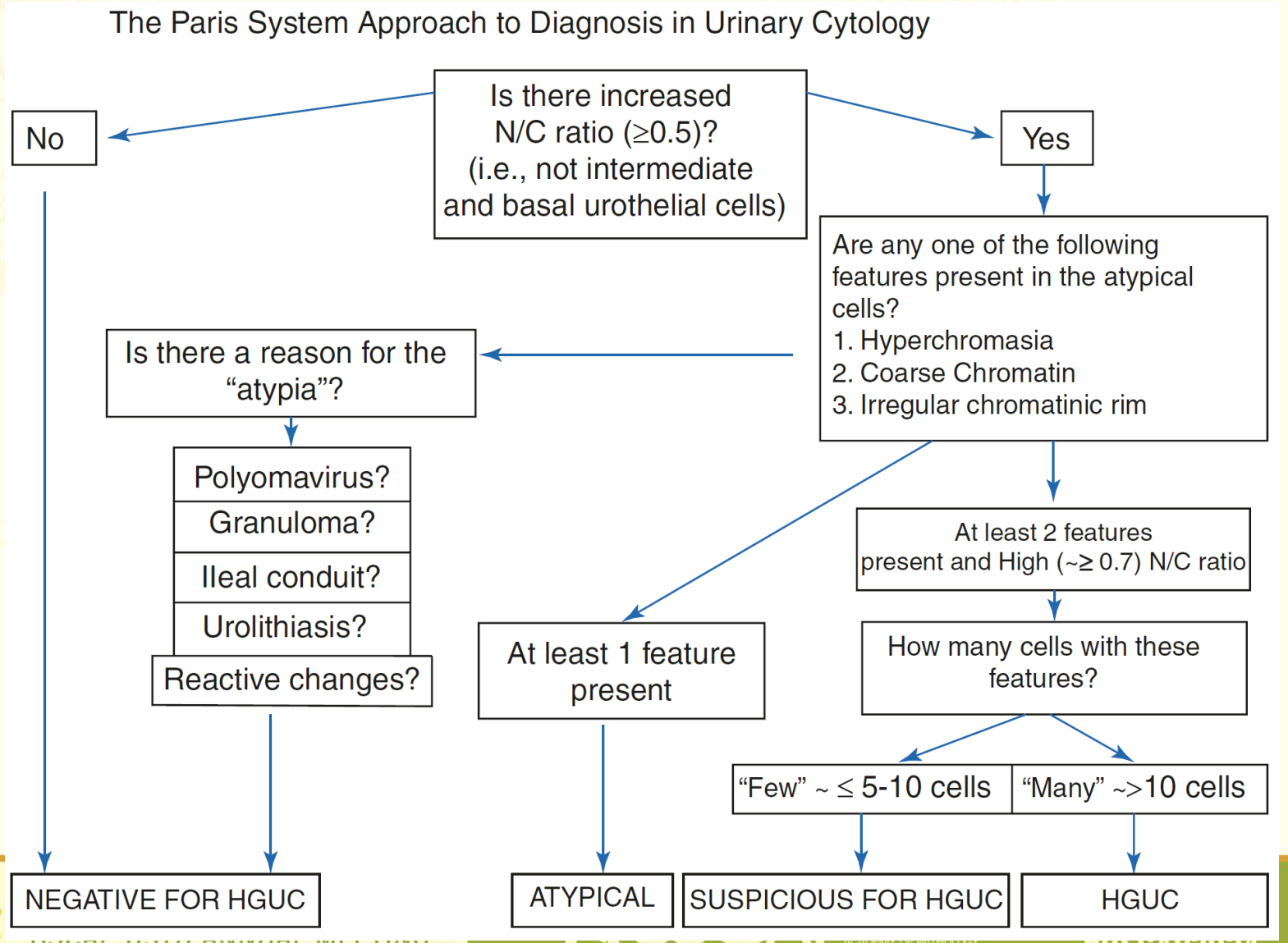
Creating a Better Pathologist

Disclosure

I serve as the Co-Founder and Chief Operating Officer at AIxMed, Inc.

Introduction





Introduction

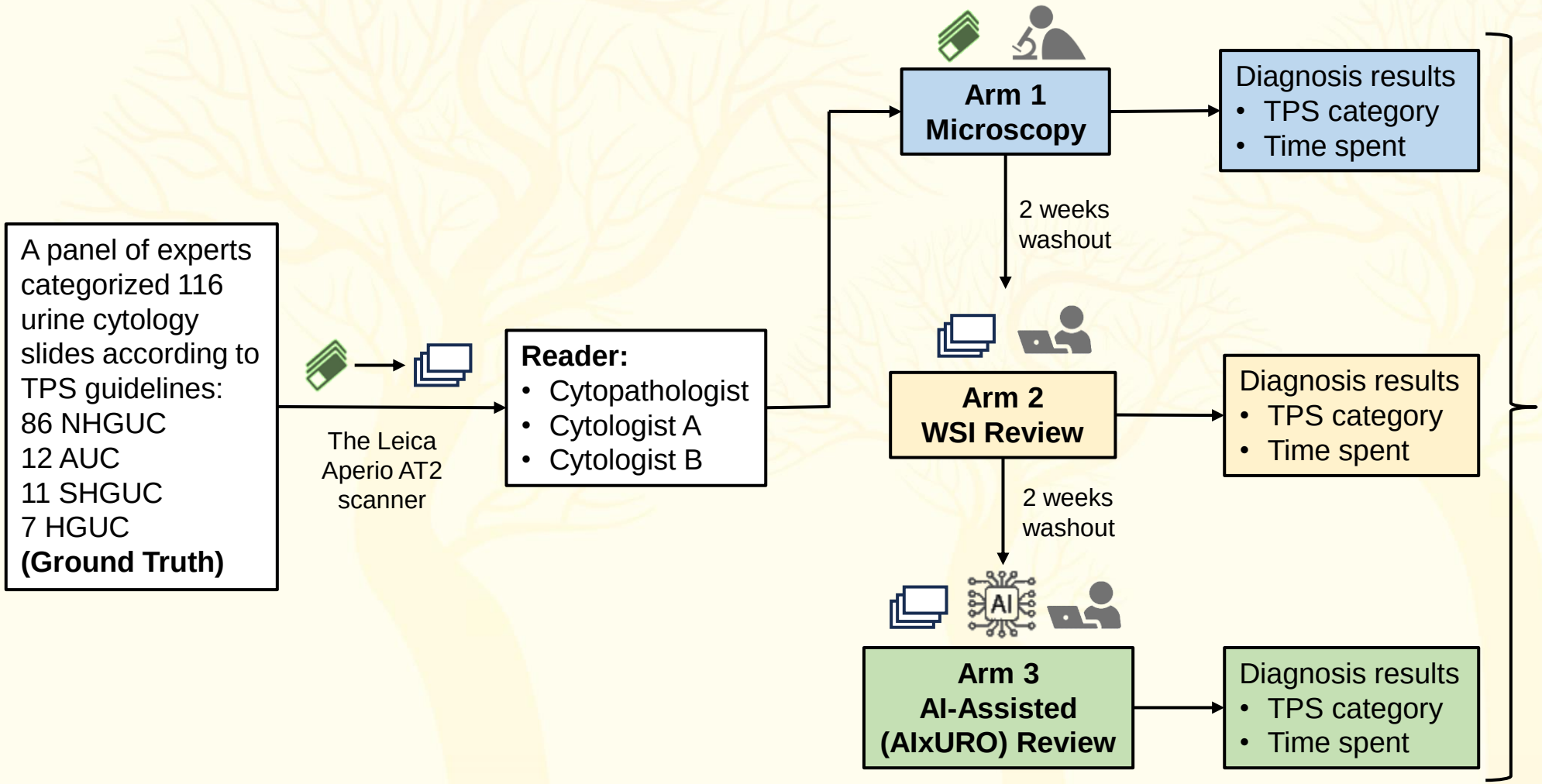
- Despite adherence to guidelines of The Paris System (TPS), variability in microscopic interpretation persists, demanding objective diagnostic methods.
- AIxURO, a deep-learning tool, enhances digital urine cytology by accurately identifying atypical cells in whole-slide images (WSIs), demonstrating clinical validity for bladder cancer diagnosis.
- Incorporating AIxURO into the clinical urine cytology workflow is expected:

  - To streamline diagnostic efficiency and accuracy

  - To reduce subjectivity in TPS categorization

  - To surpass the capabilities of conventional microscopy

Materials and Methods

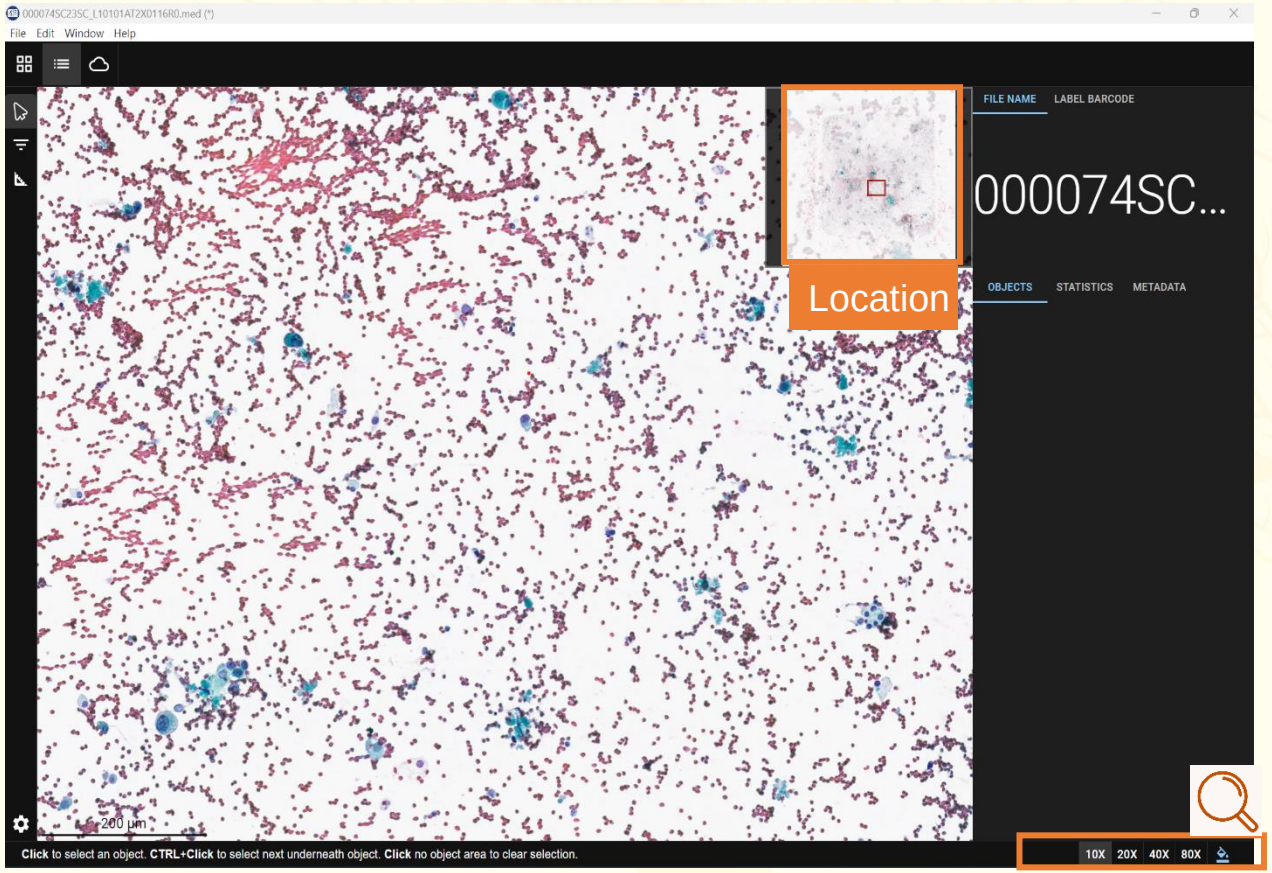


Evaluation metrics

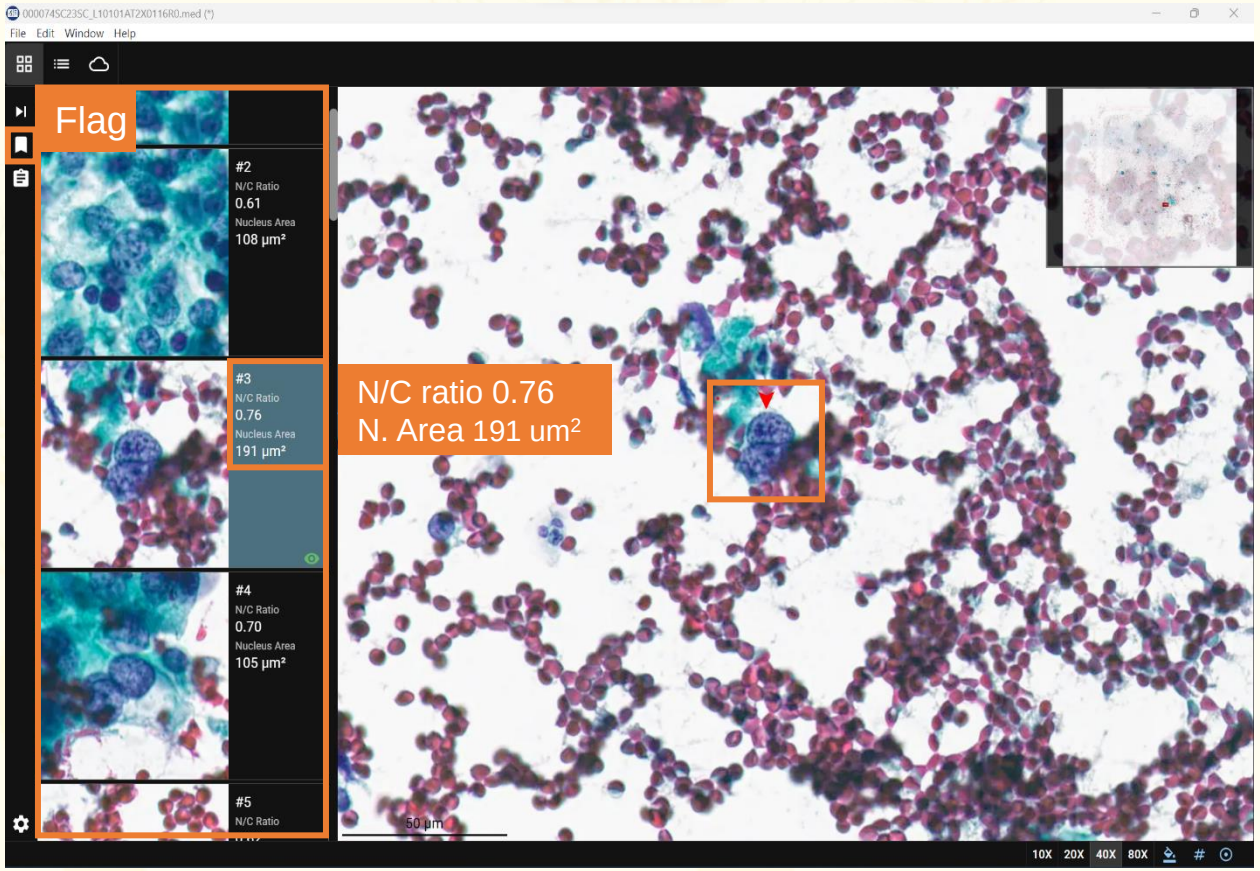
- Performances for each TPS category: Sensitivity/Specificity
- Time spent on diagnostics for each TPS category

Materials and Methods

Arm 2: WSI review



Arm 3: AI-assisted (AIxURO) review

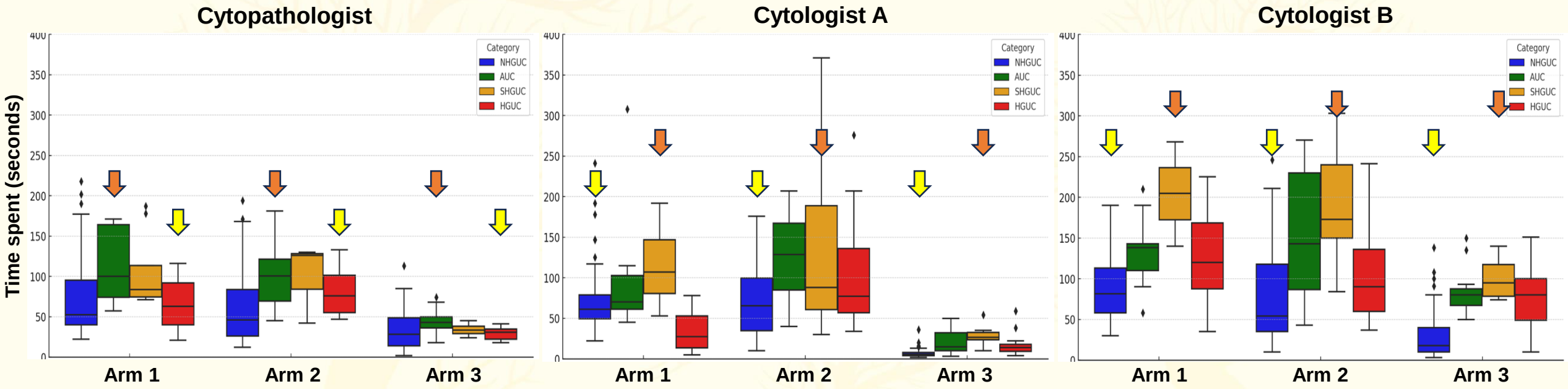


Results-Diagnosis

Ground Truth TPS Category		NHGUC (N= 86)			AUC (N=12)			SHGUC (N=11)			HGUC (N=7)		
Reader	Method	Arm 1	Arm 2	Arm 3	Arm 1	Arm 2	Arm 3	Arm 1	Arm 2	Arm 3	Arm 1	Arm 2	Arm 3
Cytopathologist	Sensitivity	96.5%	93.0%	93.0% ↓	25.0%	41.7%	75.0% ↑	54.5%	27.3%	36.4% ↓	100.0%	100.0%	100.0%
	Specificity	76.7%	76.7%	90.0% ↑	94.2%	91.3%	91.3% ↓	98.1%	100.0%	100.0% ↑	98.2%	95.4%	96.3% ↓
Cytologist A	Sensitivity	70.9%	80.2%	60.5% ↓	25.0%	25.0%	83.3% ↑	9.1%	9.1%	54.5% ↑	100.0%	85.7%	100.0%
	Specificity	93.3%	83.3%	100.0% ↑	88.5%	89.4%	74.0% ↓	93.3%	93.3%	96.2% ↑	78.9%	87.2%	90.8% ↑
Cytologist B	Sensitivity	94.2%	84.9%	86.0% ↓	25.0%	25.0%	33.3% ↑	18.2%	9.1%	9.1% ↓	100.0%	57.1%	85.7% ↓
	Specificity	76.7%	73.3%	80.0% ↑	90.4%	80.8%	85.6% ↓	99.0%	96.2%	97.1% ↓	95.4%	97.2%	93.6% ↓

- NHGUC Diagnostics: AIxURO (Arm 3) demonstrated lower sensitivity yet higher specificity compared to Microscopy (Arm 1).
- AUC Diagnostics: Sensitivity increased with AIxURO, albeit with a trade-off in specificity.
- SHGUC Diagnostics: Varied sensitivity among evaluators with AIxURO; Cytopathologist and Cytologist B noted a decrease, whereas Cytologist A showed an increase.
- HGUC Diagnostics: Cytopathologist and Cytologist A maintained sensitivity with AIxURO, but Cytologist B experienced a decrease.

Results- Time



- The Cytopathologist recorded the longest diagnostic time for AUC and the shortest for HGUC.
- Cytotechnologists required the most time for SHUC diagnosis, with the least time spent on NHGUC.
- Compared to Microscopy (Arm 1), AIxURO (Arm 3) demonstrated a reduction in diagnostic time across all categories—a time savings over microscopy of 32-45% for the Cytopathologist, 10-50% for Cytologist A, and 31-62% for Cytologist B.

Discussion & Summary

- AIxURO outperformed Microscopy in diagnostic accuracy for AUC, maintaining comparable accuracy across other TPS categories.
- AIxURO significantly reduced the diagnostic time for each TPS category.
- Observer variations in diagnostic accuracy and time were evident among the three methods tested.
- Reviewing WSI without AI assistance did not improve diagnostic accuracy or efficiency.
- The findings highlight the crucial role of AI in improving the clinical application of the Paris System in urine cytology reporting.

Thank You

Evaluating the Feasibility of a Two-Step Workflow in Artificial Intelligence-Assisted Digital Urine Cytology

Tien-Jen Liu[1], Wen-Chi Yang[2], Shin-Min Huang[2], Hui Wen Ho[3], Wei-Lei Yang[1], Hsing-Ju Wu[3], Cheng-Hung Yeh[1], Shih-Wen Hsu[1], Ming-Yu Lin[1] and Pei-Yi Chu[3]*

[1]AIxMed, Inc., Santa Clara, CA, USA; [2]Department of Pathology, Show Chwan Memorial Hospital, Changhua, Taiwan; [3]Research Assistant Center, Show Chwan Memorial Hospital, Changhua, Taiwan; *Correspondence to PY Chu ✉ chu.peiyi@msa.hinet.net

Disclosure: Tien-Jen Liu, Wei-Lei Yang, Cheng-Hung Yeh, Shih-Wen Hsu, and Ming-Yu Lin are employees of AIxMed, Inc.

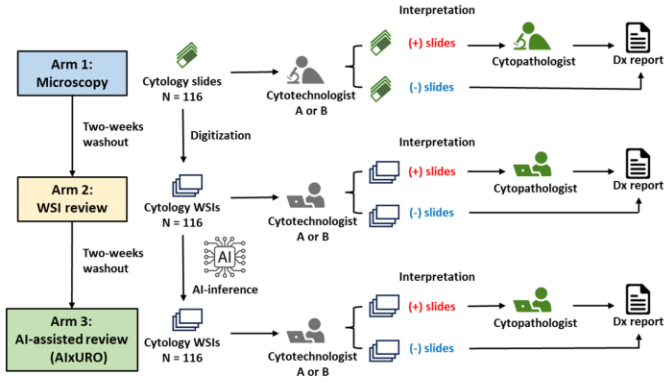
USCAP 113TH ANNUAL MEETING BRINGING EDUCATION TO LIFE

Poster #14

Introduction

- Routine urine cytology screenings for bladder cancer contribute to a significant clinical workload, primarily due to the low frequency of abnormalities, with the majority of cases being negative.
- Cervical cancer screening employs a two-step Pap test workflow to enhance efficiency: cytotechnologists initially assess slides for interpretations, and slides deemed "abnormal" are referred to cytopathologists for further evaluation, thereby optimizing the workflow and reducing the burden on cytopathologists.
- The efficacy of artificial intelligence (AI)-based digital urine cytology tools in improving the performance and efficiency of bladder cancer screenings has been established.
- This study investigates the potential of incorporating a two-step urine cytology workflow, augmented by AI-assisted tools, into bladder cancer screening protocols to optimize case management and diagnostic accuracy (Figure 1).

Figure 1. Two-Step Workflow for Urine Cytology in Cancer Screening



Materials and Methods

- An expert panel diagnosed 116 urine cytology slides, identifying 30 as positive and 86 as negative for bladder cancer. These slides, forming the "ground truth," were digitized into whole-slide images (WSIs).
- The slides/images were evaluated by three readers (two cytotechnologists and one cytopathologist) using three distinct methods: microscopy (Arm 1), WSI review (Arm 2), and AI-assisted review using AIxURO (Arm 3).
- The evaluation initiated with Arm 1, adopting the two-step workflow for bladder cancer screening. Following a two-week washout period, the readers progressed to the subsequent methods.
- For each method, performance metrics such as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and total diagnostic time were calculated for each reader.
- The outcomes were analyzed to assess the feasibility of adopting the workflow in cancer screening (Table 1).

Results

Table 1. Performance and Interpretation Time Across Three Methods in the Two-Step Workflow

The Two-step Workflow						Step 1: Cytotechnologist			Step 2: Cytopathologist			Workflow Spent Time (min)
Method	Order of Readers	Performance of Binary Diagnosis (Positive/Negative)				Intepretation of Total Cases		Total Spent Time (min)	Intepretation of Positive Cases		Total Spent Time (min)	
		Sensitivity	Specificity	PPV	NPV	Positive	Negative		Positive	Negative		
Arm 1 Microscopy	Cytotech A + Cytopath	76.7%	96.5%	88.5%	92.2%	53	63	130.2	26	27	80.0	210.2
	Cytotech B + Cytopath	76.7%	97.7%	92.0%	92.3%	28	88	197.3	25	3	47.4	244.7
Arm 2 WSI review	Cytotech A + Cytopath	70.0%	92.4%	80.8%	90.0%	42	74	170.3	26	16	56.8	227.1
	Cytotech B + Cytopath	63.3%	95.4%	82.6%	88.2%	35	81	187.9	23	12	55.9	243.8
Arm 3 AI-assisted review (AIxURO)	Cytotech A + Cytopath	90.0%	93.0%	81.8%	96.4%	64	52	29.0	33	31	43.2	72.2
	Cytotech B + Cytopath	76.7%	93.0%	79.3%	92.0%	36	80	87.0	29	7	23.4	110.4

Cytotech: Cytotechnologist; Cytopath: Cytopathologist; PPV: Positive Predictive Value; NPV: Negative Predictive Value

- In Arm 3, Cytotechnologists A and B identified more positive cases (64 and 36, respectively) than in Arm 1 (53 and 28) and Arm 2 (42 and 35).
- Arm 3 enhanced the sensitivity (90.0% vs. 76.7% in Arm 1) and NPV (96.4% vs. 92.2% in Arm 1), highlighting the efficiency of the AI-assisted two-step workflow. The performance of Arm 2 paralleled that of Arm 1.
- The adoption of the two-step workflow in Arm 1 reduced the total diagnostic time to 210.2 (74.4%) and 244.7 (70.0%) minutes, a reduction from the 282.5 and 349.6 minutes incurred by combining Cytotechnologists A or B's evaluations with the Cytopathologist's.
- Notably, the adoption of the two-step workflow in Arm 3 further condensed the diagnostic time to 72.2 and 110.4 minutes, demonstrating marked time savings compared to both Arm 1 and Arm 2.

Conclusion

- AI-assisted review (Arm 3) matched the diagnostic accuracy of both microscopy (Arm 1) and WSI review (Arm 2) while significantly reducing the time required for cancer screening.
- Arm 3 improved cytotechnologists' efficiency by identifying more positive cases, highlighting its potential to enhance the preliminary screening process.
- The lack of performance or time improvements with Arm 2 emphasizes the crucial role of AI in optimizing the screening workflow.
- These findings underline AI's pivotal contribution to digital urine cytology and its potential for integrating a two-step workflow in bladder cancer screening.