healthcare systems, where vast amounts of valuable data remain disorganized and underutilized, limiting their potential to support research and improve patient care.

Design: We developed **MAGIC** (Multimodal Analysis of Genomics, Imaging, and Clinical Data), an analytics platform enabling pathologists to leverage large-scale datasets for clinical and translational research. MAGIC applies natural language processing (NLP) to extract standardized indicators (e.g., tumor site, histologic grade, margin status) from scanned reports. Its interface supports cohort building based on clinical or molecular criteria, with seamless export for downstream analysis. Squamous cell carcinoma (SCC), a common histotype across multiple organ systems, served as a pilot use case.

Results: Using MAGIC, we assembled a cohort of 1,390 SCC cases and exported the dataset to cBioPortal for comparative genomic analysis. Our NLP pipeline accurately parsed and normalized the unstructured pathology reports in PDF formats into structured data elements, including demographics, tumor site, grade, stage, margin status, treatment effect, and histologic features such as differentiation, mitotic activity, lymphovascular invasion, and perineural invasion. These structured data were linked to sequencing datasets encompassing genomic alterations and gene expression profiles stored in the NCI Commons.

As a use case, we examined associations between tumor grade, primary site, and mutation profiles. Comparing 754 Grade 1–2 SCC cases with 532 Grade 3–4 cases revealed significant differences in anatomical site distribution (Figure 1). Notably, **ZFHX4** mutations were enriched in Grade 3–4 tumors ($\log_2$ ratio $= -1.04$, $p = 3.74\times10^{-7}$, $q = 6.61\times10^{-3}$), suggesting a potential biomarker of aggressiveness (Figure 2).



Figure 1. 100% stacked bar chart of "(Primary) Site" for Grade 1 & 2 Differentiation vs. Grade 3 & 4 Differentiation

Figure 1 - 1364.

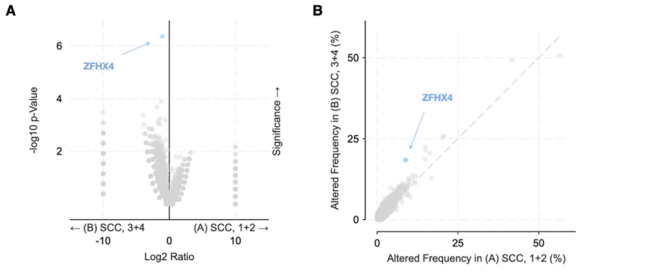


Figure 2. SCC cases: Grade 1 & 2 Differentiation vs. Grade 3 & 4 Differentiation. **A**: Volcano plot illustrating the $\log_2$ ratio of mutation frequencies (X-axis) versus statistical significance represented as $-\log_{10}$(p-value) (Y-axis). **B**: Scatter plot comparing alteration frequencies between SCC samples with grade 1–2 differentiation (X-axis) and grade 3–4 differentiation (Y-axis).

Figure 2 - 1364.

Conclusions: This proof-of-concept shows that analyses requiring months of manual curation can be completed in a single day using MAGIC with cBioPortal. By automating pathology report extraction and integrating multi-omic datasets, MAGIC offers an intuitive, scalable, and reproducible tool for translational research. Such platforms democratize access to underutilized data resources, accelerate hypothesis generation, and may ultimately advance biomarker discovery and precision oncology.

Lab Invest 106 (2026) 105658

https://doi.org/10.1016/j.labinv.2025.105658

1365 | Impact of Algorithm Assisted Scoring on Pathologist Concordance in PD-L1 Immunohistochemistry of Urothelial Carcinoma

Büşra Yaprak Bayrak[1], Kerem Karaca[2], Ece Ozturk[3], Nazif Alperen Yıldırım[4], Gizem Solmaz-Yilmaz[5], Irem Sonmez[5], Sercan Çayır[5], Burak Uzel[5], Mahmut Akgul[6], [1] *Kocaeli University School of Medicine, Izmit, Turkey;* [2] *Akdeniz University, Antalya, Turkey;* [3] *Acibadem University School of Medicine, Istanbul, Turkey;* [4] *Gulhane Military Medical Academy, Ankara, Turkey;* [5] *Virasoft Corporation, New York, NY;* [6] *Brigham and Women's Hospital, Harvard Medical School, Boston, MA*

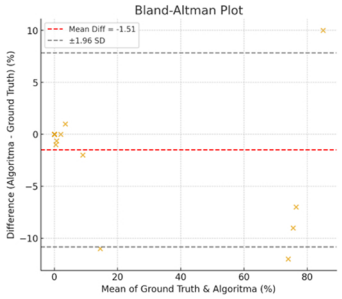
Disclosures: Büşra Yaprak Bayrak: None; Kerem Karaca: None; Ece Ozturk: None; Nazif Alperen Yıldırım: None; Gizem Solmaz-Yilmaz: None; Irem Sonmez: None; Sercan Çayır: None; Burak Uzel: None; Mahmut Akgul: None

Background: PD-L1 is a predictive biomarker in urothelial carcinoma. However, interobserver variability in PD-L1 scoring is a major challenge, particularly among less-experienced pathologists. Digital pathology algorithms have the potential to reduce variability and align junior assessments with expert standards.

Design: A total of 21 urothelial carcinoma cases with both PD-L1 IHC were included. Two senior pathologists identified hotspot regions and measured tumor proportion score (TPS%) for PD-L1 Three pathology residents (R1, R2, R3) first evaluated each slide independently without assistance, and with assistance of the algorithms' masks and scores. Accuracy was quantified using mean absolute error (MAE), root mean square error (RMSE), intraclass correlation coefficient (ICC), Cohen's kappa (using clinical cut-offs: <1%, 1–49%, ≥50%), and Bland–Altman analysis. Paired statistical tests assessed differences between algorithm-assisted and unassisted resident performance.

Results: The results of the algorithm (Figure) were similar to those of the senior pathologists (T-test = 0.16). For R1, MAE was 2.71 without the algorithm and 2.43 with it. RMSE decreased from 5.24 to 4.56; ICC increased from 0.986 to 0.990; and the Kappa decreased from 0.92 to 0.84. For R2, the MAE improved from 3.76 to 2.86, the RMSE from 7.42 to 6.18 and the ICC from 0.974 to 0.983; meanwhile, the Kappa was similar as 0.76. For R3, the MAE was 2.19 without the algorithm and 2.52 with it, the RMSE was 4.71 versus 4.90, the ICC was 0.989 versus 0.989, and the Kappa was 0.84. Paired t-tests comparing resident and senior pathologist scores yielded p-values >0.05 across all residents.

GT & Algorithm



T-test = 0.16

ICC = 0.988 [95% Confidence Interval: 0.971 – 0.995]

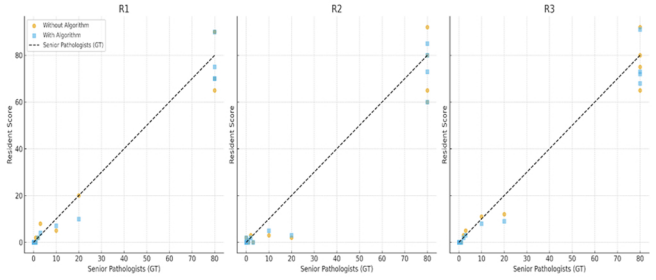


Figure 1. Scatter plots comparing residents' PD-L1 TPS assessment with and without algorithm assistance against senior pathologists (ground truth).
Each subplot shows one resident (R1, R2, R3). The dashed line represents perfect agreement with senior pathologists. Yellow circles indicate scores without the algorithm, while blue squares represent scores with the algorithm. Algorithm assistance led to closer alignment with ground truth for R1 and especially R2, while R3 already showed high concordance without substantial improvement.

Figure 1 - 1365.

Conclusions: Algorithm assistance reduced error and increased correlation for R1 and R2. However, R3 showed no additional benefit due to the high level of initial agreement. These results suggest that algorithm support could have the greatest impact on residents with lower baseline accuracy, improving reproducibility in PD-L1 scoring. However, larger datasets are needed to confirm statistical significance and clinical utility.

Lab Invest 106 (2026) 105659

https://doi.org/10.1016/j.labinv.2025.105659

1366 | Validating Live Whole Slide Imaging for Intraoperative Consultation

Sayeeda Yasmeen[1], Jingxin Qiu[1], Raima Memon[1], Jess LaVergne[1], Mihai Merzianu[1], Han Yu[1], Mokhtar Desouki[1], Deborah Sullivan[1], Thaer Khoury[1], [1] *Roswell Park Comprehensive Cancer Center, Buffalo, NY*

Disclosures: Sayeeda Yasmeen: None; Jingxin Qiu: None; Raima Memon: None; Jess LaVergne: None; Mihai Merzianu: None; Han Yu: None; Mokhtar Desouki: None; Deborah Sullivan: None; Thaer Khoury: None

Background: Whole Slide Imaging (WSI) is an increasingly emerging digital pathology tool for intraoperative consultations, particularly when remote access is needed. While WSI is approved for primary diagnosis, its application for intraoperative consultation (IOC) remains under investigation. The aim of the study is to assess accuracy and efficiency of WSI compared to glass slides (GS).

Design: Eighty previously submitted and verified frozen section slides across seven pathology subspecialties were randomly selected by one pathologist. The glass slides were provided with the anatomic site to three practicing pathologists and one fellow in training. The same slides were scanned using PreciPoint R8 digital light microscope Scanner resulting in high resolution WSIs. First, the GS were reviewed, and after 2-week washout period, the WSIs were reviewed. The pathologists had to choose one of these three diagnoses for each case: positive/malignant, negative/benign, or defer to permanent / request for 2nd opinion. Diagnostic accuracy, intraobserver concordance, sensitivity, specificity, and time to diagnosis were measured and compared between the two methods WSI vs. GS.

Results: There were 46 (57.5%) cases with benign diagnosis, 2 (2.5%) defer to permanent / request 2nd opinion and 32 (40%) malignant. For all raters combined, the concordance rate between WSI and GS was 91.9% with Cohen's Kappa 0.92 (0.88, 0.96). The sensitivity and specificity values are illustrated in table 1. The overall median and range of time of scanning was 28 (4, 38) seconds. The diagnosis time of WSI was 2.1-fold longer than that of GS (95% CI: 1.9, 2.33, $p < 0.0001$).

Table 1
Sensitivity and specificity with GS and WSI overall and per rater

Raters	Sensitivity		Specificity	
	GS, %	WSI, %	GS, %	WSI, %
Overall	78.9	80.5	89.1	90.2
Fellow	78.1	81.2	84.8	87
Pathologist 1	78.1	81.2	87	95.7
Pathologist 2	78.1	78.1	91.3	91.3
Pathologist 3	81.2	81.2	93.5	87

Conclusions: WSI demonstrates comparable diagnostic accuracy to GS, with longer diagnostic times. Given its safety, reliability, and remote accessibility, WSI is a viable alternative for GS and could enhance flexibility in the workflows of pathology, particularly during public health emergencies or in settings requiring subspecialty access.

Lab Invest 106 (2026) 105660

https://doi.org/10.1016/j.labinv.2025.105660