Journal of Pathology Informatics 16 (2025) 100419

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Pathology Visions 2024 Overview

Pathology Visions 2024 (#PathVisions24) was a great success, and commemorates the 15th anniversary of the Digital Pathology Association (DPA). With a record-breaking 860 attendees at the Hyatt Regency in Orlando, Florida from November 3rd through the 5th, there was much to see, learn and enjoy. The pathologists, trainees, scientists, technologists, administrators, and industry partners took a break from the political backdrop of the US election and instead made the Decision Digital and worked on Advancing Digital Pathology together.

Sunday began with a selection of 11 companion meetings and industry-sponsored workshops that provided an extraordinary introduction to the many advances in this field. The official meeting began that evening with an opening reception that allowed attendees to meet friends and colleagues, to visit the sold-out exhibition floor featuring 65 innovative companies, and to see the first of more than 50 e-Posters showcasing cutting-edge research.

Monday morning, DPA President Michael Quick kicked things off with the DPA Executive Committee for an insightful panel discussion on "Looking Ahead to Vision 2030", followed by a keynote lecture by Anne Martel from Sunnybrook Research Institute in Toronto who told her story of "AI in Digital Pathology: Closing the Gap Between the Research Lab and the Clinic". This was followed by a global allied society panel discussion that included representation from the DPA, Asian Society of Digital Pathology (ASDP), College of American Pathologists (CAP), and European Society of Digital and Integrative Pathology (ESDIP).

Tuesday's events started with a plenary talk "Advances in data- and hypothesis-driven models in pathology" by Lee Cooper of Northwestern University. This was followed by two important discussions about regulatory and standards and reimbursement.

There were 30 platform presentations between two tracks, and lots of lively discussions following each one. These and the more than 50 e-Posters showcased state-of-the-art advancements, enhancing networking among experts and recognizing emerging talents in digital pathology investigation. Congratulations were extended to this year's poster award winners:

- **Best Faculty/Staff (Non-Industry):** Predicting TME from H&E images to characterize immunotherapy responses in lung cancer patients; *Presenting author: Tamara Jamaspishvili, SUNY Upstate Medical University*
- **Best Industry:** Building a FAIR digital pathology repository for pre-clinical safety assessment: tackling storage; *Presenting author: Benjamin Freiberg, Genentech*
- **Best Trainee:** Quantitative modeling of maternal inflammatory response in placental membranes; *Presenting author: Teresa Chou, Northwestern University*

Thank you to our 2024 judges: Adam Booth, Matthew Cecchini, Dibson Dibe Gondim, Joseph Gaut, Michael Isaacs, Hannah Krigman, David McClintock, Andrew Norgan, Jon Ritter, Rajendra Singh, Richard Torres and Mustafa Yousif.

Trainee and developing country award recipients were also recognized at Pathology Visions 2024. Please join us in congratulating them!

- Poombal, MBBS, Resident Pathologist, Baystate Medical Centre
- Mohammad Alexanderani, MD, Computational Pathology Fellow, Weill Cornell Medicine
- Teresa Chou, BS, MS, MD PhD Student, Northwestern University
- Emeka Enwere, MD, PhD, Resident (AP), University of Alberta *(DAPA Trainee)*
- Charles Herndon, BA, BS, Student Fellow, University of Pennsylvania
- Matan Kadosh, DO, Pathology Resident, Icahn School of Medicine at Mount Sinai
- Shuo "Sean" Niu, MD, PHD, MSQM, AP/CP Resident, Wake Forest University *(DAPA Trainee)*
- Phoenix Yu Wilkie, PhD Candidate, Sunnybrook Research Institute
- Shakti Kumar Yadav, DNB, Assistant Professor, All India Institute of Medical Sciences Bhopal *(Developing Country)*
- Haoyue Zhang, PhD, Postdoc Fellow, National Cancer Institute
- Mengxue Zhang, MD, PhD, AP/CP Resident, University of Chicago *(DAPA Trainee)*

The final highlight of the event was the closing session entitled "Elevating Patient Experiences and Shaping New Perspectives Through Digital Pathology" which emphasized why pathology is so important to patient care and showed how digital pathology can enhance the patient experience. Thanks to Pathologist John Groth, we heard directly from Michele Mitchell, BS, MS, PMP, and Patient Advocate, about her experiences with her Pathologist. This was followed by Bethany Williams from the UK describing initiatives to bring an understanding of Pathology to the public.

The 7.5 CME and CE credits offered through the College of American Pathologists (CAP) attest to the conference's commitment to education. The recorded presentations are accessible on the DPA website, and abstracts are published in the Journal of Pathology Informatics.

The success of #PathVisions24 owes much to the Program Committee, led by Co-Chairs Dr. Sylvia Asa and Dr. Matthew Hanna, and the active participation of committee members from around the globe. Special appreciation is extended to Ms. Abigail Norris, CAE, for her exceptional contributions.

The Committee is already working on the program for #PathVisions25 which will take place October 5–7, 2025 as we return to the Manchester Grand Hyatt in San Diego where the DPA began. The DPA remains dedicated to serving its members by providing educational initiatives that highlight progress in digital and computational pathology. The enclosed

http://dx.doi.org/10.1016/j.jpi.2025.100419

Available online 11 February 2025

abstracts will serve as a summary of the scientific advances showcased at the conference.

PV24 Oral Abstracts

Digital and computational cytology: Applications, guidelines and future directions

Zaibo Li
The Ohio State University Wexner Medical Center, Pathology, Dublin, OH, United States.
Email: Zaibo.Li@osumc.edu

The rapid evolution of digital technology and artificial intelligence (AI) has ushered in a new era in surgical pathology practice, significantly transforming traditional methodologies. While digital cytology and AI have been slower to integrate compared to other fields, they are steadily gaining traction within cytology laboratories. Recognizing the significance of this shift, the American Society of Cytopathology (ASC) convened a digital cytology task force to assess the current state of digital cytology and AI adoption in the field. Through comprehensive research and analysis, the ASC task force has not only identified the existing landscape of digital cytology and AI in cytology but has also issued a set of recommendations aimed at validating the implementation of digital cytology and AI in cytology practice. These recommendations serve as a roadmap for laboratories seeking to leverage digital technologies to enhance their diagnostic capabilities and improve patient outcomes.

This session aims to delve deeply into the critical aspects of the digital transformation of cytology practice. Attendees will gain a comprehensive understanding of how digital cytology and AI can be seamlessly integrated into their workflow, thereby optimizing efficiency and accuracy in diagnostic processes. By exploring the myriad benefits and challenges associated with digital cytology, participants will gain insights into how this technological revolution is reshaping the landscape of cytology practice. One of the key points of the session will be an exploration of the ways in which digital cytology and AI contribute to enhancing patient care. By streamlining workflows, facilitating collaboration among pathologists, and enabling more precise diagnoses, these technologies have the potential to significantly improve patient outcomes and overall healthcare delivery. Furthermore, the session will provide an overview of current AI models that are relevant to cytology. Attendees will gain valuable insights into the various development approaches employed in creating these models, ranging from traditional machine learning techniques to more advanced deep learning methodologies. By understanding the challenges and benefits associated with each approach, participants will be better equipped to evaluate and implement AI solutions in their own practice settings.

In summary, this session offers a comprehensive exploration of the intersection between digital technology, artificial intelligence, and cytology practice. By providing attendees with practical insights and actionable recommendations, it seeks to empower pathologists to embrace and harness the transformative potential of these technologies, ultimately driving improvements in patient care and advancing the field of cytology.

Digital pathology implementation in a multi-site academic hospital reflections and lessons learned

George Yousef[1], Charlotte Carment-Baker,[2] Blaise Clarke,[1] Ioannis Prassas[2]
[1]University Health Network, Laboratory Medicine Program, University of Toronto, Toronto, Canada
[2]University health network, Laboratory Medicine Program, Toronto, Canada
Email: George.Yousef@uhn.ca

University Health Network (UHN) is a multi-institutional academic hospital in Toronto, Canada with 4 primary hospital sites and 29 satellite locations served by the laboratory medicine program. We recently moved into a fully digital pathology practice.

We share this five-year journey of transformation into digital pathology for the benefit of other institutions on their way to digitalization. We share the challenges faced, key lessons learned, and the best practices that we developed during implementation.

We adopted a stepwise approach, starting with remote intraoperative consultation (frozen section) across different locations, then digital pathology sign-out for off-site locations and external consults. These were very useful learning experiences. The final step was to move into a fully-digital clinical practice across all sites. We decided to pursue three separate RFPs; the first one for scanners, the second for workflow solution (image management software), and the third for storage. This gave us more flexibility to get the best fit in each category. Open systems were emphasized. It is important to note that “compatibility” between the different systems needs to be confirmed by track record experience.

A key lesson learned is that pathology digitalization is an institutional, rather than departmental, project that needs support from executive leadership, department of pathology (especially pathologists and MLTs) and orchestration with other clinical departments.

A number of committees and working groups were established under a dedicated change management team. Monthly open discussion sessions were held for pathologists over a year period, including external invited speakers.

Careful assessment of budget is important. Estimation of the cost of smaller items is also critical. Budget goes beyond the sum of the big items.

We customized a validation / education learning process modified from both the UK Royal College protocol and the CAP protocol for validation.

A deep understanding of current laboratory processes must be captured as digital pathology necessitated workflow modifications, including the addition of new steps related to digital quality. Also, we paid special attention to the pathologists' office configuration.

It must be noted that digital pathology transformation is a gradual process, and it is risky to claim success too early. Key performance indicators must be determined early in order to assess beforeand after full digitization. We are now in the phase of evaluation and follow up fine tune the process.

AI-driven comparative study of high-grade cell features in urine cytology with biopsy correlation

Barbara Crothers[5], Jen-Fan Hang,[1] Pei-Yi Chu,[2] Chih-Jung Chen,[3] Tang-Yi Tsao,[4] Min-Che Tung,[4] Ming-Yu Lin,[5] Wei-Lei Yang,[5] Wen-Chi Yang,[2] Shin-Min Huang,[2] Shih-Wen Hsu,[5] Cheng-Hung Yeh,[5] Yi-Siou Liu,[5] Hsing-Ju Wu,[2] Tien-Jen Liu[5]
[1]Taipei Veterans General Hospital, Pathology, Taipei, Taiwan
[2]Show Chwan Memorial Hospital, Pathology, Changhua, Taiwan
[3]Taichung Veterans General Hospital, Pathology, Taichung, Taiwan
[4]Tung's Taichung MetroHarbor Hospital, Pathology, Taichung, Taiwan
[5]AIxMed, Inc., clinical research, Santa Clara, United States
Email: barbara.crothers@aixmed.com

Background

Urine cytology is a common and cost-effective screening and diagnostic tool for detecting high-grade urothelial carcinoma (HGUC). The Paris System (TPS) for reporting urine cytology defines the cytomorphologic characteristics for the reliable diagnosis of HGUC and emphasizes the nuclear-cytoplasmic (N/C) ratio as a key criterion. Additionally, nuclear size is also a relevant reference for reporting urine cytology. Currently, there is a lack of tools for generating quantitative and numerical data on N/C ratio

and nuclear size from large cell populations, hampering the correlation between TPS criteria, biopsy, and cytology findings. To address these subjective challenges, we developed AIxURO, an AI-powered software designed to generate cell-level representations from a whole slide image. Based on TPS morphological features, N/C ratio, and nuclear size, AIxURO performs quantitative analysis to assist in urine cytology reporting. This study also aimed to explore the cytologic-histopathologic correlations in biopsy-confirmed carcinoma in situ (CIS) and HGUC using AIxURO.

Methods

A multi-institutional retrospective study (4 hospitals) submitted 242 urine cytology slides (74 AUC, 56 SHGUC, and 112 HGUC) and their corresponding positive biopsy specimens collected within six months. Among these, 212 biopsies were diagnosed as high-grade urothelial carcinoma (Biopsy-HGUC) and 30 as carcinoma in situ (Biopsy-CIS). The 242 corresponding abnormal urine cytology slides were prepared using three preparation methods: Cytospin (162 slides), ThinPrep UroCyte (60 slides), and BD CytoRich (SurePath, 20 slides), followed by Papanicolaou staining. The AIxURO software categorized cells as Atypical Urothelial Cells (atypical cells) or Suspicious for High-Grade Urothelial Carcinoma (suspicious cells) and presented the top 24 "suspicious" cell images considered to have the highest risk of malignancy in a gallery for review. Statistical significance was evaluated using Kruskal-Wallis tests.

Results

Among 242 cytology slides, the software identified a total of 124,980 abnormal cells (16,662 suspicious cells and 108,318 atypical cells). The N/C ratio and nuclear size of each cell were analyzed. The average N/C ratio of the top 24 AI-selected suspicious cells (0.66, 95 % CI: 0.65–0.66) was larger than all suspicious cells detected (0.65, 95 % CI: 0.64–0.65, $p = 0.0035$) and significantly larger by 15.8 % than categorized atypical cells (0.57, 95 % CI: 0.57–0.57, $p < 0.0001$), respectively. The average nuclear size of the top 24 suspicious cells (110.5 μm^2, 95 % CI: 105.7–115.3) was significantly larger than total atypical cells (85.8 μm^2, 95 % CI: 83.1–88.5, $p < 0.0001$).

In the cytology HGUC group, the N/C ratio of the top 24 suspicious cells (0.66, 95 % CI: 0.65–0.67) was larger than the total number of suspicious cells (0.65, 95 % CI: 0.64–0.65, $p = 0.0056$) and significantly larger than total atypical cells (0.57, 95 % CI: 0.57–0.57, $p < 0.0001$). The N/C ratio of the top 24 suspicious cells was also larger than the total atypical cells, with a statistically significant difference ($p < 0.0001$) in cytology AUC+ and SHGUC+ groups. Additionally, in the HGUC group, the nuclear size of the top 24 suspicious cells (108.1 μm^2) was larger than those of atypical cells (84.5 μm^2, $p < 0.0001$). Similar results were also found in cytology AUC + and SHGUC+ groups. No significant difference in average nuclear size was found among the three sample preparation types (Cytospin, UroCyte, and CytoRich) for either suspicious or atypical cells.

Conclusions

The AIxURO software using Whole Slide Imaging demonstrates a significant advancement in quantitative cytology analysis, offering a rapid, consistent, and precise evaluation of the nuclear size and N/C ratio. In this study, significantly larger N/C ratios and nuclear sizes occurred in the top 24 suspicious cells displayed in the gallery than in the total suspicious cells, or total atypical cells analyzed from urine cytology slides correlated with 242 positive biopsy specimens.

These preliminary findings suggest that the cell characteristics and classification of abnormal urine cytology cells based on risk stratification with gallery presentation by AIxURO provide a valuable tool for quantitative analysis to assist in decision-making for urothelial carcinoma detection. The ability to rapidly and consistently quantify cytologic characteristics has the potential to mitigate interobserver discrepancies and enhance patient care.

Intra-patient co-registration of prostate whole-mount histopathology to magnetic resonance imaging

Fatemeh Zabihollahy[1,2,3,*], Holden H. Wu,[1] Anthony E. Sisk,[4] Robert E. Reiter,[5] Steven S. Raman,[1] Neil E. Fleshner,[3] George M Yousef,[2] and KyungHyun Sung[1]

[1]Department of Radiological Sciences, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

[2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada

[3]Divisions of Urology and Surgical Oncology, Department of Surgery, University Health Network, Toronto, Ontario, Canada

[4]Department of Pathology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

[5]Department of Urology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

*Corresponding author.

E-mail address: Fatemeh.Zabihollahy@uhn.ca (F. Zabihollahy)

Background

MRI is a useful non-invasive tool for evaluating the extent of prostate cancer (PCa). However, the inter-observer agreement of qualitative assessment of MRI is only modest for detecting and grading PCa lesions. Machine learning-based algorithms can improve PCa diagnosis. Accurate labeling of PCa lesions on MRI, however, is a prerequisite for training supervised learning-based models. Whole-mount histopathology (WMHP) provides an essential reference standard for the detection and grading of PCa tumors. Once the tumor is labeled on the WMHP image, the image is cognitively aligned to T2-weighted (T2W) MRI for mapping the PCa boundary from WMHP to MRI for further research studies.

Objectives

Our objective was to develop a novel artificial intelligence (AI)-based approach to register WMHP to pre-surgical T2W MRI for further research studies. High-resolution ex vivo MRI was used to discover and assess the structural relationship between in vivo MRI and WMHP.

Methods

This study was performed in compliance with the United States Health Insurance Portability and Accountability Act of 1996, where the local institutional review board waived the requirement for informed consent. The study cohort included prostate multi-parametric (mp)MRI of 315 patients before robotic-assisted laparoscopic radical prostatectomy performed between 2016 and 2020 on one of the 3 T MRI scanners (MAGNETOM Trio, Verio, Skyra, or Prisma, Siemens, Erlangen, Germany). A genitourinary radiologist used mpMRI to delineate a 3D contour of the prostate and PCa. The images were divided into 270 and 45 for train and test, respectively.

Since there are fundamental differences in image appearances between image pairs (i.e., in vivo MRI and WMHP), high-resolution ex vivo MRI was used as an interface to facilitate spatial alignment between image pairs. Each WMHP slide was compared to the 2D slices across 3D ex vivo MRI to find its corresponding MR image. In vivo and ex vivo MRIs were then resized and re-sampled to the same size and resolution. A 2D slice across in vivo MRI with a similar spatial position to the ex vivo MRI was selected to be registered to the WMHP slide.

Generally, the conventional registration methods, such as VoxelMorph, that solely focus on image intensity for learning the deformation field generate a noisy transformation domain when there are inherent dissimilar intensities between image pairs. The proposed registration method uses the anatomy of the prostate whole gland (P_{wg}) as a supervision signal for deformable field learning. An embedding layer was included in the VoxelMorph architecture that integrates the anatomy of P_{wg} into the network.

To acquire P_{wg} shape information on different imaging modalities, a target localization head was added to the pipeline to detect and segment the prostate on WMHP and MRI automatically. Regions with irrelevant information were then suppressed to accentuate the importance of the target for the learning transformation field required for registration. A simple encoder-decoder architecture was trained using a limited number of images for automated prostate contouring on WMHP slides. For prostate delineation from in vivo MRI, 3D U-Net-based was employed as an encoder-decoder architecture to effectively examine interslice information in a 3D image. A dense block was embedded into the 3D U-Net to mitigate learning redundant features and strengthen feature propagation inside the network.

The registration method contained affine and deformable transformations. For affine transformation, scaling and translation were calculated and applied to WMHP to align the image pair globally. To learn the deformable field, the optical flow was estimated, which is a related registration problem for 2D images, returns a dense displacement vector field between WMHP and MRI. Once the displacement field was learned during the training phase, it was applied to each pixel of the WMHP to warp it and register it to MRI.

The Dice similarity coefficient (DSC) was used to estimate prostate volume overlap between the registered MRI and WMHP. Hausdorff distance (HD) was also estimated between contours computed from the registered image pair. All metrics were measured for each patient, and the average across all patients in the dataset was reported.

Results

The registration module achieved DSC and HD of 0.95 ± 0.06 and 3.77 ± 0.77 pixels on the test dataset. The proposed registration method was compared to the state-of-the-art registration model; VoxelMorph which yielded DSC and HD of 0.84 ± 0.05 and 5.93 ± 0.78 on the same test cohort. The Wilcoxon rank sum test with $\alpha = 0.05$ was conducted to investigate the significance of the difference between DSC and HD computed by different registration methods. The p-values were < 0.0001 for both metrics, which demonstrates the effectiveness of the proposed algorithm for intra-patient multi-modal image registration.

Conclusions

We developed a novel multi-modality registration method between presurgical prostate MRI and histopathology images that allows accurate mapping of PCa from WMHP to MRI. A target localization module was added to the registration architecture that not only eliminated the need for manual segmentation of the prostate on both MRI and histopathology images but also significantly improved the registration accuracy compared to the VoxelMorph.

Advancing regulatory science in digital and computational pathology: A collaborative dialogue (updated)

Jochen Lennerz[1], Noor Falah,[2] Jithesh Veetil,[2] Kim Blenman,[3] Alexandra Kalof[4]; Keith Wharton,[5] Brandon Gallas[6]

[1]BostonGene, CSO, Waltham, MA, United States

[2]MDIC, MDIC, Arlington, VA, United States

[3]Yale University, Pathology, New Haven, CT, United States

[4]The University of Vermont Medical Center, Department of Pathology, Burlington, VT, United States

[5]Roche, Roche, Rockport, MA, United States

[6]FDA, OSEL, Silver Spring, DC, United States

Email: joe.lennerz@bostongene.com

Background

As digital and computational pathology become increasingly integrated into clinical practice, the Pathology Innovation Collaborative Community (PIcc), convened by the Medical Device Innovation Consortium (MDIC), recognizes the critical importance of regulatory science in this evolving landscape. We propose a focused session at the 2024 Visions meeting to foster collaborative dialogue on regulatory topics within digital pathology, specifically addressing its adoption in the USA.

Introduction

This session aims to facilitate dialogue among stakeholders from industry, academia, and regulatory agencies to address key regulatory challenges and opportunities in adopting digital pathology in the USA. Leveraging diverse expertise, we seek to advance regulatory science and accelerate the development and adoption of innovative medical devices in pathology.

Session objectives

1. Provide an overview of current regulatory advances and challenges in digital pathology adoption in the USA.
2. Discuss the results of the PIcc regulatory landscape survey and gather audience input through a digital survey.
3. Highlight the role of PIcc in advancing regulatory science and fostering industry-academia-regulatory partnerships
4. Share insights from regulatory agencies, including the FDA, on regulatory considerations for digital pathology technologies.

Session format

The session will include short impulse talks by representatives from PIcc, MDIC, regulatory agencies (e.g., FDA), and industry stakeholders, followed by a panel discussion. Talks will cover specific regulatory challenges, innovative approaches, and collaborative initiatives in digital pathology.

Interactive component

A digital survey will be used to gather real-time audience input on key regulatory issues and broader community needs, enhancing engagement and inclusivity.

Specific focus

The session will focus on "*Taking the First Steps to Digital Pathology Adoption in the USA*," including understanding rules and regulatory precedents. It will also address the results of the PIcc regulatory landscape survey.

Benefits of participation

- Gain insights into current regulatory trends and challenges in digital pathology adoption.
- Network with key stakeholders from industry, academia, and regulatory agencies.
- Contribute to the advancement of regulatory science and development of innovative medical devices.
- Enhance visibility and recognition within the pathology community through oral presentation and abstract publication in the Journal of Pathology Informatics.
- Participate in an interactive digital survey to provide real-time input on critical issues, ensuring inclusivity and capturing a wide range of perspectives.

Conclusion

This session represents a collaborative effort to address critical regulatory issues in digital pathology adoption, leveraging the results of the PIcc regulatory landscape survey. By engaging diverse stakeholders, we aim to

drive progress in regulatory science and improve patient care through innovative technologies.

Connectathons in digital pathology, driving interoperability and innovation for clinical workflow

John Groth[1], Kenneth Philbrick[2], Brian Napora[3], Norman Zerbe[4]
[1]Endeavor Health, Pathology, Endeavor Health, IL, United States
[2]Google, IT Research, Google, CA, United States
[3]Gestalt Diagnostics, WA, United States
[4]Charité - University Hospital, Berlin, Germany
Email: jgroth@northshore.org

As digital pathology systems gain momentum in clinical diagnostics, the importance of interoperability between systems from different vendors is paramount. This session will explore the evolution and significance of the Connectathon for digital pathology, highlighting key milestones, recent breakthroughs, and future developments. Participants will learn about the use of the Digital Imaging and Communications in Medicine (DICOM) standard for whole slide imaging and the importance of requesting DICOM features in pathology solutions. We will discuss lessons learned from recent Connectathon events, focusing on annotations, regulatory considerations, and the roadmap for expanded functionality and broader vendor participation. A joint initiative with the IHE Pathology and Laboratory Medicine (PaLM) group will also be introduced, aiming to resolve workflow issues and improve integration. By attending this session, pathologists will better understand how Connectathons are shaping the future of digital pathology.

This is not just a discussion—it's a call to action! Learn how you join the next Connectathon and be at the forefront of advancing digital pathology! As vendors, you can showcase your commitment to interoperability by actively participating in these events and demonstrating that your solutions meet the standards required for clinical workflows. Pathologists, now is the time to require standards-based interoperability from the very beginning of every procurement process. Ensure that DICOM and IHE PaLM compliance is a requirement in your contracts to guarantee seamless integration, enhanced flexibility, and future-proofing of your pathology practice.

Together, we can drive the widespread adoption of standards that will benefit the entire field of digital pathology, ensuring that both systems and workflows are not only interoperable but also scalable and ready for future innovations. Let's make standards-based interoperability the norm, not the exception.

Paving the way for regulatory and standards advances in digital pathology

John Groth[1], Esther Abels[2], Soumen Roy[3], Laura Chang[4]
[1]Endeavor Health, Pathology, Endeavor Health, IL, United States
[2]Solaris RTC, Solaris RTC, Boston, MA, United States
[3]FDA, Silver Spring, MD, United States
[4]Artera, Los Altos, CA, United States
Email: jgroth@northshore.org

The mission of the Digital Pathology Association (DPA) Regulatory and Standards Task Force is to advance digital pathology by clarifying the regulatory pathway, raising awareness of its evolution, and working towards the development and adoption of standards while promoting interoperability for clinical use. This session will provide a comprehensive update on the task force's four key initiatives: developing responses and educational materials for the FDA's Laboratory Developed Tests (LDT) Rule, creating guidelines for AI and ML-based technologies, advancing interoperability, and collaborating with the FDA.

Attendees will gain valuable insights into the task force's efforts, particularly the implications of the European Union Artificial Intelligence Act (EU AIA) for digital pathology. The session will also cover the ongoing collaboration with the FDA, focusing on addressing regulatory challenges for digital pathology devices, interoperability issues, and key updates from the regulatory landscape.

This session is more than just an information update—it's a call to action. Pathologists, industry leaders, and innovators are encouraged to actively participate in shaping the future of digital pathology by contributing to regulatory and standards development. Your involvement is crucial to ensure that the digital pathology ecosystem can thrive in a rapidly evolving regulatory environment. Together, we can drive innovation, enhance patient care, and shape the next generation of pathology practices.

Elevating patient experiences and shaping new perspectives through digital pathology

John Groth[1], Michele Mitchell,[2] Bethany Williams[3]
[1]Endeavor Health, Pathology, Endeavor Health, IL, United States
[2]Patient, Patient
[3]Leeds, Pathology, Leeds, United Kingdom
Email: jgroth@northshore.org

In this talk we will explore how the fusion of digital transformation and direct patient care in pathology is reshaping patient journeys and redefining our roles as pathologists. Through real-world cases, as told by a patient, a lead for training, educational and public/patient involvement and a pathologist, we will unveil how patient interactions with their pathology images through direct visualization, guide these journeys. Discover actionable insights to navigate the balance between innovation and compassionate care, all centered on enhancing the patient and pathologist experience.

AI in digital pathology: Closing the gap between the research lab and the clinic

Anne Martel
Sunnybrook Research Institute, Toronto, Physical Sciences, Toronto, Canada
Email: a.martel@utoronto.ca

One of the main advantages of adopting a digital workflow in pathology is the ability to use computers to analyze microscopy images, reducing the workoad of pathologists. Advances in Artificial Intelligence over the last 10 years have made it possible to count cells, detect tumors and classify disease with accuracy approaching, or even surpassing, that of pathologists. It is even possible for AI models to predict patient outcomes by assessing imaging features not discernable to the human eye. The first part of the talk will introduce the main AI methods and explore some of the exciting applications in digital pathology.

Despite this rapid progress, the number of applications where pathologists are able to make use of AI in their routine workflow is relatively small and there are still many barriers to overcome before AI reaches its full potential. In the second half of this talk I will delve into the challenges faced when translating algorithms developed in the research lab- into clinical settings and explore potential solutions.

Advances in data – And hypothesis – Driven models in pathology

Lee Cooper
Northwestern University, Pathology, Chicago, IL, United States
Email: lee.cooper@northwestern.edu

Artificial intelligence is advancing at a breathtaking pace. Consumer AI models trained on large and varied datasets have shown remarkable breadth of applicability. This success has translated into pathology, where recent examples include models have been trained with hundreds of thousands of whole slide images with diverse pathologies, or image-text pairs

harvested from internet databases and social media. These models present new opportunities to accelerate AI development in areas where data is lacking and are an exciting direction in a field defined by sub specialization of human experts. This talk will explore these developments and other pertinent topics, including the relevance of AI model interpretability and the explainability of predictions, and the challenges of studying AI in real-world settings.

Reimbursement task force update - Setting the stage for DP and AI reimbursement

Michael Rivers[1], Paul Gerrard[2], Matt Fickie[3]
[1]Roche, Digital Pathology, Roche Diagnostics, Santa Clara, CA, United States
[2]Artera, Government Affairs, Palo Alto, United States
[3]Highmark, Inc., Pittsburgh, PA, United States
Email: michael.rivers@roche.com

While digital pathology adoption is increasing and AI image analysis continues to show impressive potential to impact the practice of pathology, the overall returun on investment (ROI) for digital pathology implementation remains a significant challenge. With an increasing number of DP solutions being cleared by the FDA for IVD applications, the lack of reimbursement is quickly becoming the most signfiicant barrier to DP adoption.

Category III codes introduced in 2023 for tracking of whole slide imaging have been been difficult to implement and may not be generating the evidence required to move foward quickly with signficant change. Furthermore, to date these codes are focused solely on digitization in the lab and do not provide for tracking or reimbursement for emerging AI image analysis solutions developed for digital pathology.

The DPA Reimbursement Task Force has a mission to “To define and shape the pathway to reimbursement for digital pathology solutions enabling broad market access and delivery of value to patients and stakeholders.” To this end we have identified key strategic initiatives to enable the DPA to play a leading role in advocacy for a pathway to reimbursement for AI solutions. In addition, the task force is working through blog posts and webinars to educate the DPA members on the current state of reimbursement for DP and the opportunities ahead.

This session will provide an update on the key work of the task force and provide a detailed look into the framework for DP AI reimbursement that the task force is pursuing.

Generative AI in anatomical pathology: Today's innovations and tomorrow's possibilities

Ehsan Ullah[1], Victor Brodsky,[2] Andrey Bychkov,[3] Andrew Song,[4] Eric Walk,[5] Peter Louis,[6] Ghulam Rasool[7], Rajendra Singh[8], Faisal Mahmood,[4] Marilyn Bui[7], Anil Parwani[9]
[1]Health New Zealand, Auckland, Anatomical Pathology, Auckland City Hospital, Auckland, New Zealand
[2]Washington University in St. Louis, Department of Pathology and Immunology, St. Louis, United States
[3]Kameda Medical Center, Department of Pathology, Kamogawa, Japan
[4]Massachusetts General Hospital, Harvard Medical School, Department of Pathology, Boston, MA, United States
[5]PathAI, Pathology, Boston, AZ, United States
[6]Rutgers Biomedical and Health Sciences, Department of Pathology and Laboratory Medicine, New Brunswick, United States
[7]Moffitt Cancer Center, Department of Machine Learning, Tampa, FL, United States
[8]Summit Health, Dermatopathology, Woodland Park, NJ, United States
[9]Wexner Medical Center, The Ohio State University, Department of Pathology, Columbus, OH, United States
Email: ehsanu@adhb.govt.nz

Background

Generative Artificial Intelligence (AI) models such as Large Language Models, Vision Language Models, and Foundation Models have emerged as transformational tools across various disciplines. Anatomic pathologists are already interacting with generative AI chatbots during case sign out to get assistance on differential diagnosis work-up, recommended special stains as well as immunohistochemistry and molecular testing based on current guidelines, and to produce a draft of the pathology report. Generative AI is starting to reshape the technical and administrative processes within anatomic pathology too and holds immense potential for a host of new potential applications to completely transform the discipline of anatomic pathology and cancer diagnosis.

Aims

This presentation aims to delve into the generative AI applications, advantages, and challenges in anatomic pathology, emphasizing its influence on technical laboratory processes and pathologist sign out procedures, to highlight the opportunities to make workflow efficiency gains. This talk will also note the impact of generative AI applications on the educational paradigms and research advancements specifically within anatomic pathology.

Design

This presentation is based on a collaborative work conducted by a group of researchers and professionals encompassing pathology and AI fields who have conducted a thorough literature review into the recent developments in generative AI applications within anatomic pathology. These generative AI applications will be categorized into unimodal and multimodal applications and will be assessed for their current clinical utility, ethical implications, and future potential.

Results

Generative AI exhibits substantial promise across several domains in anatomic pathology. AI-driven image analysis, virtual staining, and synthetic data generation significantly enhance diagnostic precision. Automation of routine tasks, quality control, and reflex testing demonstrates potential for considerable workflow improvements leading to quicker turnaround times assisting faster treatment and better patient outcomes. AI-generated educational materials, synthetic histology images, and advanced data analysis methods foster enhanced educational and research opportunities. Initial findings suggest anatomic pathology workforce seems cautiously optimistic about the transformative potential of AI. Pathologists show interest in adopting AI tools for non-diagnostic tasks. There is a growing spectrum of various applications in academic settings. Dependable AI tools will need to go through rigorous testing and evaluation before and after each implementation to ensure quality.

Conclusions

Generative AI holds the potential to revolutionize anatomic pathology by enhancing diagnostic accuracy, improving workflow efficiency, and advancing education and research. However, its successful integration into clinical practice demands ongoing interdisciplinary collaboration, meticulous validation, and strict adherence to ethical standards to ensure that AI's benefits are fully realized while maintaining the highest levels of patient care. This talk will explore the transformative potential of generative AI in anatomic pathology, offering participants valuable insights into its current and future applications and addressing the necessary steps for its successful and ethical implementation in clinical practice.

Keywords

Generative AI, Anatomic pathology, Diagnostic accuracy, Workflow efficiency, Education, Research, Ethical considerations.

End-to-end AI-enabled automation of pathology accessioning

Eish Maheshwari[1], Alexander Kwon,[1] Yixing Jiang,[1] Jeya Maria Jose Valanarasu,[1] Michael Copenhaver,[2] Alex Dussaq,[2] Thomas Montine,[3] Andrew Ng[1], Eric Yang[3]

[1]Stanford University, Computer Science, Stanford, CA, United States

[2]Stanford Health Care, Laboratory Systems, Stanford, CA, United States

[3]Stanford University School of Medicine, Pathology, Stanford, CA, United States

Email: eishm@stanford.edu

Introduction and background

Timely and accurate pathology diagnoses are crucial for medical management decisions and clinical trial eligibility. Academic institutions and large healthcare enterprises offer expert consultations nationwide for the most complex cases requiring urgent pathology input. At Stanford Pathology alone, the volume of consult cases has surged by over 10 % in the past two years, now exceeding 14,000 cases annually. These cases represent 25 % of all surgical pathology cases, with further growth anticipated due to the aging baby boomer population. Despite this growth, the intake process for consult cases, known as *accessioning*, remains labor-intensive, requiring significant staff effort to manually enter numerous patient information fields. This time-consuming process is vulnerable to staff availability fluctuations, leading to substantial delays. Accessioning bottlenecks account for nearly 50 % of the total turnaround time per consult, delaying critical pathology results and undermining care for the sickest patients. In this study, we propose an AI-based solution for automating pathology accessioning. We present a novel end-to-end deep learning strategy for efficiently and accurately extracting patient information from document images, as well as an infrastructure strategy for clinical integration.

Model design

Understanding document images is challenging due to the need for both text recognition and holistic document comprehension. Traditional visual document extraction pipelines rely on off-the-shelf optical character recognition (OCR) for text reading and then focus on interpreting OCR outputs. While promising in certain clinical applications, these OCR-based methods are *computationally expensive* and *lack flexibility* across different languages and document types. Pathology accessioning, in particular, presents additional challenges as consult cases vary in structure, length, and writing quality, with many documents being handwritten or using non-text elements like checkboxes. This highlights the need for a more generalizable approach. We developed a transformer-based vision-language model (VLM) for processing document images end-to-end, eliminating the need for an intermediate text representation like OCR. Our model utilizes a vision encoder and a text decoder to directly extract patient information. This approach significantly improves generalizability, essential for effectively handling the diverse and complex nature of consult cases. By leveraging various pretraining and fine-tuning strategies, ensemble methods, and data augmentation techniques, our model achieves 88–95 % average normalized Levenshtein similarity (ANLS) in extracting patient information fields, with inference times of just 1–2 s. This performance significantly surpasses OCR-based approaches, with a 13 % improvement in accuracy and an 8-10× reduction in runtime. Additionally, we implemented a text localization strategy that integrates spatial information from OCR with model outputs to approximate bounding boxes for extracted text, facilitating manual reviews. Furthermore, we developed a quality-checking model that cross-references extracted patient information across document pages to automatically detect missing or incorrect data, ensuring the accurate reporting of confidential medical information.

Clinical integration

For clinical integration, we developed a human-in-the-loop pipeline to ensure zero tolerance for errors with patients' data. Documents are scanned into a PHI-safe cloud-based bucket, automatically triggering model inference to extract patient information and run data quality checks. This initial analysis happens locally and asynchronously without the need for human input. Accessioners then review cases using a custom user interface, allowing them to verify and, if necessary, correct the extracted information from the model, ensuring data accuracy is maintained. The verified patient data is then sent to Epic Beaker to create requisition orders. Our cloud-based serverless infrastructure allocates compute resources on-demand, saving costs and ensuring scalability. Thus, our semi-supervised pipeline maintains data accuracy and patient privacy while dramatically accelerating the accessioning workflow.

Concluding remarks

Our system accurately and efficiently extracts patient information from pathology documents in 1–2 s, compared to the several minutes typically spent per accessioning case, significantly reducing bottlenecks and treatment delays. This text extraction pipeline can be readily scaled to various medical document processing tasks, including the formatting and drafting of pathology reports, as well as automated metadata extraction for efficient management of whole slide images (WSI). Our solution has a profound impact on our ability to deliver timely and accurate pathology results.

Ergonomics for pathologists in the digital age: What can we learn from radiologists?

Les Folio, Marilyn Bui

Moffitt Cancer Center, Radiology, Moffitt Cancer Center/Tampa, FL, United States

Email: folio47@gmail.com

Though ergonomics is a well-established discipline in occupational and preventive medicine, there is a lack of awareness of basic principles and application in enterprise imaging. This is especially true for imaging centric medical specialties, namely pathology and radiology. In the digital age, like radiologists, pathologists are experiencing increasingly complex viewing workstations; many with increasingly common repetitive stress injury (RSI) compounded by lack of institutional oversight.

The presenters are a practicing radiologist and a pathologist from a leading academic medical center who will discuss the best practice examples and lessons learned from literature review, resources gathering, and a wellness initiative to shed light on the timely topic of ergonomics for pathologists in the digital age.

As part of our medical center's wellness program and quality improvement, we identified a need to increase ergonomics awareness in radiology, specifically viewing state, starting with a three-minute instructional video and short description centered around three basic ergonomic principles. Content is organized by the “three points of contact” well known in the imaging informatics community as follows:

1. Where the eyes meet the monitor(s)
2. Where the hands/fingers meet input devices (keyboard, mouse(s))
3. Where the feet meet the floor (when standing) or body meets the chair (when sitting)

The video is an informal discussion with a radiologist asking questions about another radiologist's workstation that includes an advanced standing/sitting workstation desk, four monitors, a dictation microphone stand

for hands free dictation, two gaming mouses and keyboard. A short description based on the three points outlined above is included with a short quiz on the topics covered in the video. The video was made available on an internal shared drive (OneDrive by Microsoft) and announced at a faculty meeting as optional viewing and optional participation in the quiz (initially optional to assess interest). Our efforts resulted in wellness funding allowing for a Certified Professional Ergonomist (CPE) to perform ergonomic evaluations during several site visits with a resultant presentation that was shocking for most attendees. For example, many with RSI suddenly realized why they seemed to be "falling apart" from work and not tennis, golf, or other activities.

This approach could be adopted by pathologists to assess the current state of ergonomic awareness and identify areas for improvement in their working environment, particularly within the context of a new digital sign-out workflow. By studying best practices and literature, and securing institutional support and funding, we can address ergonomic concerns in pathology in digital age. Doctor, heal thyself! Then the patients will benefit from our wellbeing!

Lessons from integrating digital pathology into clinical enterprise imaging at Michigan medicine

Mustafa Yousif
University of Michigan, Pathology, Ann Arbor, MI, United States
Email: mustafay@umich.edu

Our journey at Michigan Medicine involves integrating the pathology department into our Clinical Enterprise Imaging (CEI) program. We're using a centralized Picture Archiving and Communication System (PACS) to transform the department into a fully digital operation. Our goal is to achieve primary diagnosis by [summer 2024], a significant milestone in our digital transformation. This process will enhance interoperability and diagnostic precision through the Digital Imaging and Communications in Medicine (DICOM) standardization across various imaging modalities.

Enterprise imaging is a comprehensive approach that involves capturing, analyzing, routing, and managing clinical images and multimedia content from various specialties in a seamless workflow. Our modern enterprise imaging systems integrate radiology, pathology, dermatology, and other fields using a vendor-neutral format for managing and storing images. This approach ensures comprehensive access to imaging data, enabling physicians to diagnose, share insights, and make informed patient-care decisions. By connecting previously siloed information and leveraging standards like DICOM and HL7, we're enhancing the diagnostic and treatment process across multiple specialties. The DICOM standardization plays a crucial role in this, ensuring compatibility with standards-compliant applications and avoiding future data migrations.

Our journey towards a cohesive digital pathology environment began with a collective effort. We standardized DICOM protocols across multiple scanners, addressing vendor hesitance and limitations. This collaborative approach has not only improved system integration and accessibility but also fostered a sense of shared accomplishment. Implementing robust quality control workflows using DICOM metadata has significantly enhanced the accuracy and reliability of our diagnostics, a testament to our collective dedication.

Integrating metadata from our Laboratory Information System (LIS) with our PACS has also streamlined diagnostics and boosted interdisciplinary collaboration. We utilize HL7 information to fill the DICOM headers before the data resides in the Vendor Neutral Archive (VNA). Our ongoing projects include developing a PACS-driven workflow that integrates LIS and Electronic Health Records (EHR) systems, further enhancing diagnostic efficiency and functionality.

In summary, this oral presentation will cover the lessons learned from integrating digital pathology into our Clinical Enterprise Imaging program at Michigan Medicine. It will focus on the five pillars of functionality in enterprise imaging at Michigan Medicine, including image capture, storage, viewing modality, image exchange workflow, and analytics. Additionally, it will highlight the importance of the CEI Cross-functional Technical Support Team and CEI Governance—steering, Operations, and Stewardship committees.

Deep zoom images to visualize unsupervised clustering of slide tiles

Dibson Dibe Gondim
University of Louisville, Pathology, Louisville, KY, United States
Email: dibson.gondim@louisville.edu

Unsupervised image classification techniques hold significant promise for pathology because they can group similar images without explicit annotations. These methods have been successfully applied in histopathology for tumor classification and identifying patches for training supervised approaches. Typically, the visualization process involves applying dimensionality reduction to embeddings and using unsupervised methods to create clusters, projecting images in two or three dimensions. Each tile is represented by a point, with cluster categories indicated by different colors. While methods exist to view corresponding images for each point or generate separate photomontages for each cluster, they do not allow for correlating the morphology of the tiles with their spatial distribution simultaneously. To address this, we created a deep zoom image with properties equivalent to a whole slide image, allowing visualization of individual tiles at full resolution, distributed in the embedding space. This work aims to demonstrate a proof-of-concept approach for visualizing tile images with spatial context in a deep zoom image.

We selected 25 random whole slide images from the TCGA bladder cancer dataset. Tiles were extracted at 512 $\times$ 512 pixels and downsampled by a factor of 4, yielding 156,000 tiles. A binary threshold excluded tiles with over 50 % white pixels. From this set, a balanced sample of 8000 tiles across all cases was randomly selected. Embeddings were extracted using a foundational model (Prov-GigaPath), trained on one billion 256 $\times$ 256 pathology image tiles from over 170,000 whole slides. Dimensionality was reduced using Principal Component Analysis (PCA), and clustering was performed using the K-means algorithm with multiple cluster numbers (5, 6, 7, 8, and 9). To create the deep zoom image, we calculated the new location of each tile based on its coordinates in the K-means embedding space, generating a large canvas with tiles distributed proportionally to their cluster locations. Once the canvas was ready, each tile was placed in its appropriate position, enabling visualization of individual tiles at full resolution within their spatial context.

The resulting deep zoom image spanned 260,000 pixels in each dimension. This spatial representation provided more information than a photomontage without spatial context. For bladder cancer whole slide images, the clusters included invasive cancer with desmoplasia, carcinoma without desmoplasia, urothelium, papillary urothelial carcinoma, stroma, blood, and artifacts such as blurring and cropped edges. Despite some contaminants between classes, the clusters were predominantly morphologically consistent. Impressively, even tiles with color pen marks (green, blue, red) were correctly clustered with tiles of the same class without ink, demonstrating the quality of embeddings created by the foundational model.

Evaluating the deep zoom image provided greater context and understanding of how tiles were clustering. This method allows for a more robust evaluation of visual semantic clustering performance by visualizing the spatial distribution of tiles. It can be applied to multiple slides or cases, as well as single slides. The primary computational costs are associated with extracting embeddings from multiple patches and storing large images. Deep zoom images offer significant benefits by providing detailed spatial context, helping to understand the relationships between different clusters and the morphological characteristics of each tile. However, further exploration in various settings is needed to fully determine the method's value and potential applications.

PathAssist: Accelerating dermatopathology diagnosis with integrated knowledge graphs and LLMs

Rajendra Singh[1], Aneesh Sathe,[2] Anil Parwani[2]
[1]Summit Health, Pathology, Woodland Park, NJ, United States
[2]Osumc, Bioinformatics, San Diego, CA, United States
Email: skinpathology@gmail.com

Clinicians usually refer to a diverse array of published articles, textbooks, personal notes, and online resources for diagnostic decision-making, particularly in complex cases. However, the vast volume of available data often exceeds human memory capacity, leading to inadvertent loss of critical information over time. While recent advancements in Large Language Models (LLMs) and Retrieval Augmented Generation (RAG) based systems enable a first pass search but fail to capture the depth and interconnected nature of medical knowledge required for accurate diagnostics. Dermatopathology is the perfect example as dermatopathologists often refer to not only pathology references but also to a lot of dermatology resources. In our study we use the integration of LLMs and Knowledge Graphs (KGs) for dermatopathology resources to highlight the enhancement of diagnostic capabilities through the structured understanding and dynamic processing of the vast resources of data as compared to a RAG based system.

This presentation highlights an advanced AI framework that combines the precision of KGs with the contextual comprehension of LLMs to create superior query tools for dermatopathology knowledge. We opted for KGs over traditional RAG systems due to their advanced ability to structure and interconnect vast arrays of data. This structuring not only links directly queried information but also intelligently surfaces related concepts and context not explicitly mentioned in the search terms, significantly enhancing the depth and relevance of insights for improved diagnostic accuracy. We compare query tools using the RAG model with LLM with those using KGs with LLMs. Our approach harnesses four key advantages of KG-based systems:

1. Structured Contextual Understanding: By structuring dermatology and pathology data into a detailed KG that encapsulates diseases, symptoms, histopathological features, and treatments, our AI tool utilizes the LLM's capabilities to access a rich, interconnected knowledge base for accurate diagnostics, substantially enhancing both specificity and sensitivity.
2. Provenance and Trust: Each diagnostic suggestion by our system is accompanied by provenance, linking back to the data points within the KG that support the diagnosis. This transparency allows clinicians to verify the AI's recommendations, fostering trust and facilitating wider acceptance in clinical practice.
3. Handling Complex Queries: Leveraging the organized structure of the KG, our system excels at interpreting complex, multi-symptom patient cases that traditional AI systems struggle with. This capability is critical for accurately diagnosing multifaceted dermatological conditions.
4. Efficiency and Resource Management: The use of KGs reduces the need for extensive data traversal during each query, enabling faster response times and reducing computational demands. This efficiency is particularly beneficial in high-volume clinical settings where timely decision-making is crucial.

We used the author's personal knowledge repository and automatically created a knowledge graph using LLMs as an intermediate tool. Further with LLMs as the interaction layer made extracting information from the KGs simple and distributable. We used the same knowledge repository to create a RAG based system as a baseline comparison.

For testing we worked with a cohort of 10 dermatopathologists of which 5 worked with a standard RAG + LLM system and 5 with the KG + LLM system, testing for complex questions that required multi-hop associations as in real life drawn from an identical question bank. In the preliminary cohort of 100 questions, the KG + LLM system gave 100 % accuracy compared to the RAG + LLM system, which completely failed on multi-hop associations. We will expand the testing with a larger cohort of users as well as an expanded question bank. Future work will focus on continuously expanding the KG to include new research and adapting the LLM integration to incorporate these updates seamlessly.

The implications of our research extend beyond improved diagnostic accuracy, suggesting significant enhancements in clinical workflows, patient outcomes, and the economic aspects of healthcare delivery. This system also facilitates the creation of personalized knowledge graphs for individual practitioners, enhancing their diagnostic capabilities and operational efficiency.

Optimizing and evaluating pathology foundation models for both low and high resolution tasks

Saloni Agarwal, David Steiner, Kenneth Philbrick, Theo Guidroz, Supriya Vijay, Faruk Ahmed, Ellery Wulczyn, Dale Webster, Yun Liu, Jeremy Lai
Google, Google Health, Mountain View, United States
Email: saloniag@google.com

Pathology specialized foundation models have emerged as a powerful tool for representing histopathology images and facilitating development of downstream applications for a wide variety of use cases, using less task-specific data and computational resources as compared to traditional machine learning methods. However, one challenge that can potentially limit the utilization of these models stems from the fact that nearly all available foundation models rely on creating thousands to hundreds of thousands of cropped image "patches" from each whole slide image (WSI). Processing these many patches per slide (and case) can be computationally expensive. In addition, most foundation model development and evaluation has focused on relatively high magnification image interpretation tasks with relatively less attention on lower magnification tasks, such as quality control or tissue type, even though these can also be a critical part of interpretation workflows. In this work we develop and evaluate foundation models for both high and low magnification patches and tasks, for which entire WSIs can be represented with orders of magnitude fewer patch embeddings. Specifically, we train models using patches across a range of magnifications, ranging from $\sim 4mm^2$ per patch (256×256 pixels at ~ 16 μm per pixel or $0.625\times$) up to $\sim 0.1mm^2$ per patch (~ 0.5 μm per pixel). For evaluation, we augment high resolution benchmark tasks with "low resolution" tasks such as stain quality, specimen type, tissue type, and tumor grading. We use this set to evaluate models via linear probing on held out data. Area under the receiver operating characteristic curve (AUC) was calculated for individual tasks as well as averaged across tasks to help summarize findings. We find that training using either low resolution and high resolution patches results in models that generally perform better on tasks corresponding to the matched resolution (ie. training with low resolution patches results in better performance on low resolution tasks). Training with a combination of low resolution and high resolution patches resulted in performance on par with a low resolution model for low resolution tasks (average AUC across low resolution tasks of 0.897 for combined model and 0.900 for low res model) and on par with a high resolution model for high resolution tasks (average AUC across high resolution tasks of 0.928 vs. 0.935). We propose this type of "pan-magnification model", that is flexible to the input image size and resolution, offers an important option to consider when choosing a WSI embedding strategy for optimizing performance across tasks of varying magnifications and enabling computational efficiency for lower resolution tasks.

Foundation models: A deep dive into their applications in pathology

Kingsley Ebare
Mayo Clinic, Pathology and Laboratory Medicine, Scottsdale, AZ, United States
Email: ebarekings@gmail.com

Foundation models trained on massive and diverse datasets, hold immense potential to revolutionize pathology by automating image analysis and assisting pathologists in diagnosis. This presentation will delve into introducing foundation models to the audience, exploring their key characteristics, applications in digital pathology, and a specific use case for classifying tumors. We will discuss the benefits of adopting these models, including increased diagnostic accuracy and reduced analysis time. Furthermore, we will also address the challenges associated with data privacy, ethical considerations, and the need for explainable AI. By exploring the the application of foundation models in pathology, this presentation aims to spark further discussion and collaboration between AI researchers and pathologists to optimize their integration into digital pathology research and practice.

FROG: An unsupervised deep learning framework for scalable and robust end-to-end IHC quantification

Carl Molnar, Mark Zarella, David McClintock, Thomas Tavolara, Christopher Garcia, Wenchao Han

Computational Pathology and Artificial Intelligence, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, United States

Email: han.wenchao@mayo.edu

Quantitative immunohistochemistry (IHC) readout is of importance for multiple disease types for accurate diagnosis, prognosis, and treatment guidance. Currently, most reporting is based on manual assessment which is time-consuming, labor intensive, and subject to inter and intra-observer variability.

Artificial intelligence (AI) technology, especially deep neural networks, helps the quantification by 1) instance cell segmentation followed by counting the segmented single cells for final scoring (a powerful approach providing both ROI/case-level IHC score and instance cell segmentation, which facilitates results verification, interpretation, and the derivation of other IHC quantification); or 2) directly performing IHC score estimation at region-of-interest (ROI) or case-level. To achieve either, existing methods rely on supervised learning approach (including weakly-supervised method), necessitating extensive annotated data for model training. In addition, well-known challenges have been persisting in domain adaptation and generalization, where models often experience substantially reduced performance or complete failure when applied to out-of-distribution data (e.g., data drift, different tissue/stain type)—image data that differs from the samples used during training. Data drift, a known issue in practice, in IHC images can occur at any stage of tissue processing, staining, or image acquisition, often unnoticed. This poses a significant risk, as models may provide inferior results without user awareness, potentially leading to incorrect diagnoses. The primary solution is training (including retraining/fine-tuning) the model with out-of-distribution data. However, for supervised methods, especially for the single cell model, the need for handers of thousands cell-level annotation makes it labor intensive with substantially high operational cost, especially if frequently performed. Crucially, such approach does not resolve the issue of unnoticed data drift.

To overcome these challenges, we designed and developed a novel deep network framework (called FROG), the first unsupervised deep learning framework for end-to-end IHC quantification without the need for any annotations, where the model can self-train on out-of- distribution data in an unsupervised manner. FROG learns to generate colored instance cell segmentation masks while simultaneously predicting cell center point and biomarker expressions for positively vs. negatively stained cells, made possible through our novel dual-branch generative network structure.

We rigorously and comprehensively validated our model on multiple IHC quantification tasks for both case and cell-level quantification using internal, external, and public datasets. These datasets encompass multiple tissue types of breasts, lung, bladder, and prostate, and include multiple nuclear IHC markers: Ki67, Estrogen Receptor (ER), Progesterone Receptor (PR). We benchmarked FROG with the widely used model (Unet) and the state-of-the-art model (DeepLIIF).

For case-level classification of Ki67-stained breast cancer, our model identified positive and negative cells to compute scores, using a 20 % cutoff based on ASCO guidelines. We used 2136 clinically signed-out cases (pathologists' diagnosis are used as for validation), constituting 678, 134 image patches, spanning 11 years from Mayo clinic, which training and testing data are split randomly for overall performance, and chronologically for the data drift situation. FROG achieved an F1 score (the harmonic mean of precision and recall) of 0.95, accuracy of 0.97, and AUC of 0.99 with random splits, substantially outperforming Unet and DeepLIIF. Under data drift conditions, FROG maintained an F1 score of 0.92, while Unet and DeepLIIF dropped to 0.63 and 0.66, respectively, highlighting FROG's robustness to data drift.

For cell-level validation, we used two public datasets: BCDatasets (181,074 annotated Ki67 cells in 1338 breast cancer image patches, 800 patches for training) and bladder and non-small cell lung cancer datasets (600 image patches). FROG matched DeepLIIF's performance (averaged about 5 counts/patch error), with Unet showing higher errors of additional 5 counts/patch for BCDatasets and 7 counts/patch for the bladder and lung datasets. The public datasets have limited sample sizes, we anticipate even better performance for FROG with a sample size exceeding 2000 image patches, as evidenced by our sample size experiments on other datasets with larger sample sizes. Note that FROG did not use any annotation during training, while the benchmark methods (Unet and DeepLIIF) used hundreds of thousands cell-level annotations for similar or inferior performance.

The third task was performed qualitatively by manually inspecting the model identified positively and negatively expressed cells. We performed this task on ER and PR-stained breast cancer tissue from a cohort of 563 cases; Ki67- stained bladder, lung and prostate cancer tissue from a cohort of 300 cases, which includes internal and external consulting cases. The results showed consistent and reliable performance across different tissue and IHC stain types.

In conclusion, we proposed a groundbreaking paradigm providing a way for unsupervised learning for cell-level quantification for IHC image by self-training without any annotations, thereby overcoming significant challenges in domain adaptation and generalization. Through rigorous validation across diverse datasets and tasks, FROG consistently outperformed benchmark models, demonstrating superior to state-of-the-art performance, robustness, and generalizability, which supports accurate, efficient and reliable clinical implementation at scale.

Moving beyond data bias: Integrating AI ethics in computational pathology for improved patient care

Chhavi Chauhan[1], Chris Garcia,[2] Mark Zarella[2]

[1]ASIP, Publicaitons, Rockville, MD, United States

[2]Mayo Clinic, Computational Pathology and Artificial Intelligence, Rochester, MN, United States

Email: chhavich@gmail.com

Abstract

As artificial intelligence (AI) advances and integrates into digital and computational pathology, it is crucial to address the potential risks associated with equity and bias in these systems. Bias in AI can arise from various sources, including unrepresentative training data, flawed algorithm designs, and human biases embedded in the development process. These biases can lead to disparities in diagnostic accuracy, treatment recommendations, and ultimately, patient care. To ensure that AI models perform effectively in real-world scenarios, it is essential to design studies that accurately reflect the intended clinical use case, the characteristics of the target patient population, and the data processing methods used in

clinical practice. Failure to align study design with these real-world factors can result in models that do not generalize well to their intended applications, leading to suboptimal performance and potential risks to patient care.

This roundtable discussion aims to raise awareness of equity and bias issues in AI for digital and computational pathology, demonstrate their relevance in current practice, and share rapidly evolving best practices for mitigating risks while promoting transparency. We will explore real-world examples of how equity and bias manifest in digital and computational pathology, examining the implications of these issues on patient care, research, and the overall trustworthiness of AI systems. By delving into case studies and current research, we aim to provide attendees with a comprehensive understanding of the challenges at hand.

Furthermore, we will discuss the rapidly evolving best practices for developing and deploying AI solutions responsibly, including strategies for ensuring diverse and representative training data, implementing robust validation processes, and promoting transparency in AI development. We will also highlight the importance of multidisciplinary collaboration, involving pathologists, computer scientists, ethicists, and other stakeholders in the development and evaluation of AI tools.

Attendees will gain valuable insights into the current landscape of equity and bias in AI for digital and computational pathology, learning practical approaches for mitigating risks, promoting transparency, and fostering trust in AI-assisted pathology workflows. By the end of the roundtable, attendees will be equipped with the knowledge and tools necessary to advocate for responsible AI practices in their own institutions and research endeavors.

As the field of digital and computational pathology continues to evolve, it is imperative that we proactively address the challenges of equity and bias in AI. This roundtable discussion serves as a call to action, encouraging the pathology community to engage in ongoing dialogue, collaboration, and education to ensure the responsible development and use of AI solutions. Together, we can harness the potential of AI to transform pathology while upholding the highest standards of equity, fairness, and patient care.

Predicting responses to NAC(ypT0) from initial TURBT specimen using Artificial Intelligence

Haoyue Zhang[1], Meera Chappidi[2], Zhijun Chen[1], Lucas Liu[3], Erolcan Sayar[3], Helen Richards[3], Jonathan Wright[4], Stephanie Harmon[1], Michael Haffner[3]

[1]National Cancer Institute, Molecular Imaging Branch, Bethesda, United States

[2]University of Washington School of Medicine, Urology, Seattle, WA, United States

[3]Fred Hutchinson Cancer Center, Data Science Integrated Research Center, Seattle, WA, United States

[4]University of Washington, Urology, Seattle, WA, United States

Email: harry.zhang@nih.gov

Background

Neoadjuvant Chemotherapy (NAC) followed by radical cystectomy is the first-line treatment for patients with muscle-invasive bladder cancer. Histomorphologic features from transurethral resection of bladder tumor (TURBT) are currently underutilized in clinical practice. We hypothesize that accurately predicting response to NAC using specimens from TURBT will provide crucial information for better treatment planning and patient stratification. Multiple instance learning (MIL) is an efficient A.I. imaging approach for patient-level or slide-level Whole Slide Image (WSI). Effective feature extraction is a major challenge for WSI, particularly in smaller cohorts. To this end, we propose a novel framework, NAC-AI, that leverages bladder WSIs from TCGA for patch-level tumor detection, a large-scale digital pathology foundational model for feature extraction; and a transformer-based cross-attention MIL model for NAC response prediction.

Methods

NAC-AI was developed with a publicly available TCGA dataset, pretrained foundational model UNI, and in-house TURBT data from patients who received NAC and RC at the University of Washington and were classified as responders (pT0 or pTis) or non-responders (pT2 and/or pN+). The framework included three components: 1) A patch-level tumor detection model using vision transformer (ViT-L16) architecture that pre-trained on TCGA bladder data and fine-tuned on TURBT slides to learn to discriminate tumors against artifacts that were unique to TURBT slides; 2) Slides were patchified, and patch-level features embedding were extracted using UNI, a ViT based digital pathology foundational model, along with tumor probability information from the tumor detection model; 3) The extracted features forming a bag were fed into a cross-attention MIL model built based on TransMIL for the final patient-level binary classification of NAC response. Patch-level probability information served as an extra hard-attention mechanism to guide the model in focusing on tumor regions. State-of-the-art approaches such as Attention-based MIL (ABMIL) and Clustering-constrained-attention MIL (CLAM) with both ResNet50 features and UNI features were compared.

Results

A total of 309 slides (1.04 million 40× patches) from TCGA were used for the path-level tumor detection model. A model with a test ROC-AUC of 0.920 was used for step 2. A balanced 138-slide (70 responders vs 68 non-responders) dataset from the University of Washington was used for tumor detection fine-tuning and NAC-AI MIL model development. Five-fold cross-validation metrics were reported. Compared to ABMIL-ResNet50, CLAM-ResNet50, and CLAM-UNI, our approach achieved a ROC-AUC of 0.70 ± 0.07 vs 0.52 ± 0.07 vs 0.55 ± 0.10 vs 0.65 ± 0.12. The corresponding accuracy, sensitivity, specificity, and precision scores are 0.72 ± 0.05, 0.81 ± 0.9,0.62 ± 0.10, and 0.69 ± 0.06.

Conclusions

Our study addresses the critical clinical need of predicting NAC responses from TURBT specimens. NAC-AI has shown promising correlations between the two, potentially improving treatment planning. Our findings underscore the value of the TURBT specimen after artifact removal using a fully automated algorithm and call for larger cohort and multi-center studies to enhance prediction accuracy.

Advancing precision pathology: Deep CNN model for forecasting of liver cancer recurrence on WSI

Mohammad Alexanderani, Robert Chen, Mohamed Omar, Luigi Marchionni

Weill Cornell Medicine, Pathology & Laboratory Medicine, New York, NY, United States

Email: mom4011@med.cornell.edu

Background

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related deaths worldwide. Worryingly, the rate of recurrence among patients who have undergone curative treatment can be as high as 88 %. Despite recent technological advancements in the field, tumor recurrence remains a significant challenge, necessitating careful reappraisal of patient and disease status. To address this, we aimed to develop a precise deep learning algorithm to predict liver cancer recurrence utilizing whole slide images (WSIs).

Methods

We developed an attention-based deep learning model to predict liver cancer recurrence from digital slides of hematoxylin and eosin-stained liver tissue. The dataset, sourced from the TCGA database, underwent preprocessing that included tissue segmentation, tiling, and extraction of histopathological features. The labeled dataset was randomly split into training (70 %), validation (15 %), and testing (15 %) sets. The attention-based model was trained on the training set and optimized using the validation set, while the testing set was kept completely unseen to be used only for the final evaluation of the model's performance. Performance was evaluated using standard metrics like AUC, sensitivity, specificity, and accuracy. Attention heatmaps were generated to provide interpretability and insights into the model's decision-making process. Additionally, we trained a HoverNet model to analyze the distribution and organization of cells in the liver cancer microenvironment, comparing recurrent and non-recurrent cases. The study is supported by the NIH-NCI for Next Generation Onco-Pathologists Program at our institution.

Results

The dataset comprised 450 whole slide images (WSIs), which were classified as either post-therapeutic liver cancer recurrence or no recurrence. The model demonstrated robust performance on the validation dataset, achieving an Area Under the Curve (AUC) of ~0.73.

Conclusion

Our study presents an innovative deep learning approach that accurately predicts liver cancer recurrence utilizing H&E whole slide images (WSIs). By leveraging attention-guided deep neural networks, we were able to develop a powerful prognostic tool for forecasting the risk of liver cancer recurrence. These findings have critical implications for optimizing personalized therapeutic interventions and surveillance strategies in liver cancer management, thus advancing the field of precision pathology. Additional, experiments are being conducted to further expand our findings.

A comprehensive AI education framework based on experiential learning

Mark Zarella, Thomas Flotte, Christopher Garcia, Katelyn Reed
Mayo Clinic, Laboratory Medicine and Pathology, Rochester, United States
Email: Zarella.Mark@mayo.edu

The application of AI in pathology benefits from the users of these tools possessing a fundamental understanding of what it takes to develop and put an AI model into practice. This includes not only existing pathologists and technical staff but also future practitioners. Knowledge and experience in AI also encourages engagement and interest, which is vitally needed to stave off workforce shortages and promote recruitment into the specialty. Extending AI fluency to our patients is also necessary to build trust in these burgeoning technologies and is essential for their ability to make informed decisions about their own care. To address each of these needs, we developed an educational approach that seeks to expose practitioners and patients to AI through experiential learning. The framework includes: 1) implementing infrastructure to directly support AI studies, including democratizing AI through a no-code AI platform, 2) establishing an internship program to reach learners within and outside the institution, 3) leveraging formal engagements with internal programs and mentoring AI related projects, 4) formalizing relationships with external academic institutions focused on engineering and computer science, 5) reaching the local community by upskilling area high school teachers using a train-the-trainer approach.

First we implemented a no-code AI solution designed to provide our pathologists, technical staff, and trainees the ability to pursue their own AI projects without programming expertise. We created an RFA process to support over 50 projects based primarily on whole-slide imaging. Several publications and conference abstracts were generated as a direct result of this effort. In parallel, we sought to complement institutional cloud computing infrastructure with on-premises computing resources targeted to junior faculty and trainees through the Major Research Instrumentation program at the NSF.

Second, we established an internship program to provide learners outside the institution with opportunities to engage in ongoing AI research projects with translational potential, mentored by established AI investigators and with access to large data sets and pathology expertise. Internships varied from 2 to 6 months and interns were at the early undergraduate, graduate, and postgraduate levels, representing medical and computing trainees alike. Our initial cohort of 14 interns were involved in projects across multiple pathology divisions and generated several first-author publications.

Third, we collaborated with formal programs within our institution to provide trainees with opportunities to fulfill their educational requirements through involvement in AI projects. These trainees, often fellows, actively seek projects to gain practical experience in clinical AI. Initially, we engaged with the Clinical Informatics fellowship program, where second-year fellows acquire first-hand knowledge by participating in operational and research activities across the enterprise. Throughout the year-long engagement, our first fellow made significant contributions to the formalization of our AI lifecycle processes.

Fourth, existing formal agreements between the Mayo Clinic and other academic institutions were leveraged to involve trainees in AI and digital pathology research, complementing Mayo Clinic's healthcare focus with the engineering and computing academic expertise found elsewhere. We participated in new programs as well, including co-op programs, capstone projects, and collaborative research opportunities.

Fifth, we established a collaboration with the Mayo Office of Education to pursue extramural funding to support the Research Experiences for Teachers program focused on AI in Healthcare. The intent of the program is to introduce AI/healthcare curricular modules to area high schools with a focus on those in rural districts less likely to have access to these resources. This program adopts a train-the-trainer approach by providing AI research experiences integrated with AI curriculum support to area high school teachers, who will then implement curricula in their schools incorporating lessons learned from their summer research experience. We received letters of collaboration from 11 partner school districts, primarily in rural districts throughout Olmstead County. More than 10 research faculty at Mayo Clinic pledged to participate as project mentors within and beyond pathology.

Together, these initiatives focused on extending AI education to all levels of learner from high school to graduate level to professional, and ensuring that learners outside the institution got access to the unique opportunities available at an institution well along in its digital journey. We suggest that an experiential-focused approach, especially when complemented by lectures, online content, and traditional didactics, can serve as a blueprint for delivering AI education to existing and future workforce in pathology.

Digital pathology implementation: Addressing impact on culture, teamwork, and communication

Orly Ardon[1], Suzanne Dintzis[2]
[1]Memorial Sloan Kettering Cancer Center, Pathology and Laboratory Medicine, New York City, NY, United States
[2]University of Washington Medical Center, Department of Laboratory Medicine and Pathology, Seattle, WA, United States
Email: ardono@mskcc.org

Background

Digital pathology is slowly being adopted in laboratories around the world, yet discussions of barriers to digital implementation in pathology tend to focus on hardware and software choices, technical interfaces and compatibility, data management, and the adoption of new AI based decision support tools. The cost and quality of the digital transformation are cited as

recognized as the major barriers in wide adoption. Another challenge in successful digital pathology implementation is the impact on work culture, prioritizing digital workflows in a high paced environment, the alignment of pathologists and laboratory staff on digital image availability turnaround time, and the need to define the requirements and expectations around communication and performance. This adoption of disruptive technologies preceding transformation of processes and employee interactions is not unique to digital pathology, and can result in costly failures, lack of workflow change, and flawed organizational practices.

Methods

An assessment and identification of workflows and teams affected by the introduction of digital workflow was conducted at Memorial Sloan Kettering's (MSK) Department of Pathology and Laboratory Medicine, an early adopter of digital pathology. The study aimed to define the training needs of the different teams and the change management activities required for the digital pathology transformation through workflow observations and interviews of team members.

Results

The digital transformation at MSK started years before the clinical implementation of digital pathology in 2020, and it is still ongoing. These activities included team identification, new departmental workflows development, training materials and townhall meetings, as well as small group training, digital adoption surveys and leadership/champion identification and development in the different teams. The activities are continuing well past the introduction of the new digital workflows into the histology laboratories and are aimed to allow seamless migration from analog to digital workflows while maintaining quality operations metrics and all stakeholder approval.

As the current laboratory technologists training program in the US does not include digital pathology workflows, most histotechnologists are not familiar with the Digital Pathology Certificate NSH/DPA. In addition, most laboratory staff is not involved in discussions of digital pathology workflow adoption and will only get trained on the technologies once they get brought into the laboratory. Pathologists were given training sessions in a group setting as well as access to training materials. There were departmental wide town hall meetings that discussed the change in the workflows and gave updates on the timelines and the teams involved.

Conclusions

The practice of laboratory medicine and pathology requires multiple teams to be physically present in the workplace to conduct tests and provide coordinated care with clinical teams. Transformations occurring without attention to the changes' impact on worker experience, and how our clinical colleagues experience us, may affect teamwork and morale. To prevent transactional relationships from replacing deeper relationships, full engagement of the pathology and laboratory community in the coming sea change will be necessary. Technology adoption across pathology departments should include clear communications, all stakeholders' identification and involvement, clear goals and policies, review of all existing analog and future digital workflows, and the setting of new roles and responsibilities. Other change management activities include administration support and leadership guidance before, during and after the digital pathology workflows implementation.

Histology hide-and-seek: Visually navigating latent space clustering for pathology exploration

Phoenix Wilkie[1], Anne Martel[2], Kelvin Jok[2]
[1]University of Toronto, Medical Biophysics, Toronto, Canada
[2]Sunnybrook Research Institute, Physical Sciences, Toronto, Canada
Email: phoenix.wilkie@mail.utoronto.ca

Background

Clustering is used in weakly and self-supervised learning to group similar images of tissue samples together. Unsupervised clustering allows for exploration of the latent space which is beneficial in digital pathology classification tasks. There exist new and improved methods for clustering pathology images. Yet, it remains difficult to assess which clustering method would be most effective for certain datasets when designing experimental research plans.

It is currently difficult to assess cluster quality quantitatively for histology patches. Standard clustering evaluation typically quantifies the inter- and intra-cluster distances. However, in histology-based analysis, the variations in morphological features are subtle, necessitating labeled downstream tasks for cluster validation.

Dunn's Index (DI) is useful for comparing clustering algorithms or parameter settings by assessing their ability to generate compact, well-separated clusters. However, in pathology images, DI often yields values of 0 due to sensitivity to noise and outliers. Significant overlap and poor cluster separation are common when assessing clustering in uncurated patch datasets. Furthermore, metrics such as the average Silhouette Coefficient (aSC) frequently approach 0, suggesting that many mixed tissue patches tend to reside near or on the decision boundary between neighboring clusters.

Therefore, we propose a visual dimensionality reduction software pipeline. This will allow rapid, visual assessment of clusters.

Methods

A graphical user interface (GUI) application was developed to address histopathology clustering assessment challenges. This software pipeline facilitates rapid visual evaluation of clustering methods by enabling users to upload feature-extracted data and corresponding patches. Users can upload results from any clustering algorithm or utilize the built-in unsupervised clustering method provided by the program.

The included unsupervised clustering method was pretrained on histology patches using the SIMCLR framework for contrastive self-supervised learning. This approach groups similar images while separating dissimilar ones. Hierarchical agglomerative k-means is then applied for further refinement.

The software allows users to conduct the Elbow Test for optimal cluster number estimation and determine the minimum number of parent clusters in hierarchical clustering. It offers automated evaluation using established metrics and generates centroid patches as output. Users with ground truth labels can select from evaluation metrics such as Adjusted Rand Index, Normalized Mutual Information, Homogeneity, Completeness, and V-Measure. Internal validation metrics like aSC, DI, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index (CHI) are available for users without ground truth labels.

The GUI allows users to reduce dimensionality, orient plots, and access patches. Users can visualize patches in latent space using TSNE, PCA, Sammon, and UMAP in 2D or 3D. Interactive navigation enables users to right-click on data points to view associated images, facilitating manual inspection of clusters. This interactive visualization approach supports a comprehensive examination of clustered data.

Results

Two downstream tasks were conducted using public and private datasets. Firstly, patches containing cross-sectional views of complete tubules were identified with 98 % (± 1.7). This was done using the public BreCaHAD dataset.

Secondly, tissue classification, particularly adipose and blood patches, was performed on an unlabelled private dataset. Initial evaluation metrics included aSC $= -0.02$, DI $= 0.0$, DBI $= 3.47$, and CHI $= 41.71$. By retaining only the 60 % of patches closest to centroids, metrics approached optimal values for aSC, DI, and DBI, indicating tighter clusters and

increased separation between neighbors. Consequently, the remaining patches contained a single tissue type per patch.

Conclusion

Assessing latent space proximity involved examining patches from diverse tissue labels. Visual inspection revealed evident morphological similarities among patches within clusters. The GUI streamlines the selection of optimal clustering methods for experimental pipelines, enabling swift quantification using standard equations for assessing cluster quality. Moreover, qualitative assessment is facilitated through interactive exploration, i.e. clicking on points within graphical clusters displays corresponding histology patches. Additional evaluation metrics are currently being implemented and integrated into the GUI software.

Extraction of discrete information from pathology reports using local and private LLMs

Ghulam Rasool[1], Asim Waqas[1], Aakash Tripathi[1], Ehsan Ullah[2], Marilyn Bui[1]

[1]Moffitt Cancer Center, Machine Learning, Tampa, FL, United States

[2]Auckland City Hospital, Auckland City Hospital, Auckland City Hospital, Auckland, New Zealand

Email: ghulam.rasool@moffitt.org

Background

Surgical pathology reports provide detailed descriptions of tumor samples and are the primary communication tool between pathologists and other clinical specialists involved in a patient's treatment journey. These reports may include critical information such as cancer site, laterality, tumor stage and grade, histology, behavior, and disease codes. Extracting this information as discrete variables has numerous downstream applications, including the maintenance of cancer registries. Cancer registries are databases that capture essential information about cancer patients, and pathology reports are a vital source for creating and maintaining accurate tumor records. Currently, the extraction of relevant information from pathology reports for cancer registries is performed manually, with human experts reviewing the documents and populating the records.

Current solution and their limitation

Various natural language processing (NLP) methods have been proposed for extracting information from pathology reports. However, these methods often fail to achieve the desired accuracy due to the complex nature of pathology reports, including cancer typing, sub-typing, and specialized medical terminology. Additionally, pathology reports may be stored in formats such as PDF or RTF, necessitating pre-processing steps like optical character recognition (OCR), which introduces additional artifacts and noise into the data. Recently, techniques based on large language models (LLMs) have been proposed. However, using LLMs presents several challenges, including (1) **privacy**: most LLMs are accessed via APIs, requiring the transmission of user data over the internet to the LLM server for processing; (2) **cost**: LLMs are billed per token (approximately three-quarters of a word in English), and the cost of processing pathology reports from both historical data and current clinical records can accumulate to substantial amounts; (3) **computational requirements**: running LLMs on-premise necessitates significant investment in computational infrastructure.

Methods

We addressed these challenges by running compressed and quantized LLMs locally within Moffitt's firewall to process over 7000 pathology reports from the TCGA project. **Data:** The PDF pathology reports from twelve solid cancers (bladder, brain, cervix, colorectal, head and neck, kidney, lung, liver, ovarian, pancreas, prostate, and uterus) were downloaded and stored locally. We developed a software pipeline to load the PDF files, perform OCR, and then use an LLM to extract six discrete variables, along with an explanation for each output selection. The variables included cancer site, laterality, stage, grade, histology, and behavior. Our prompt strategy involved two calls to the LLMs using the LangChain library: (1) we prompted the LLM to extract the six variables and provide an explanation for each extraction, and (2) we used Pydantic to force the LLM to output the variables in a JSON dictionary format using the results from the first call as the input for the LLM in the second call. The pipeline's output was stored in JSON and CSV formats. We experimented with different LLMs, including Mistral, Llama-2, Llama-3, and Mixtal, and found that the Mixtral 8x7b model (quantized at Q4_0 with 46.7B parameters) provided the best balance between processing time and accuracy. Our experiments were conducted on a desktop computer with an NVIDIA RTX A4500 (30GB VRAM) GPU and a data center compute node with an NVIDIA A30 (24GB VRAM) GPU. A pathology expert on our team analyzed the extracted variables. Given the large number of reports (6944 in total across all cancers), the experts randomly selected between 10 and 30 reports from each cancer and manually verified the correctness of the LLM-extracted variables. Reports where OCR failed to produce correct text were excluded from the experiment. Any variables not present in the original report or not applicable (as determined by the subject matter expert) were excluded from the analysis.

Preliminary results

The LLM was able to extract all six variables with an average accuracy of 99.2 % across all variables and cancer sites in 145 reports. Specifically, the extraction accuracy for each variable was as follows: cancer site 100 %, laterality 99.3 %, stage 99.3 %, grade 98.6 %, histological entity 100 %, and behavior 97.9 %. When analyzing accuracy by cancer type, we observed the following results: bladder 100 % (n = 22), brain 83.3 % (n = 12), cervix 90.9 % (n = 11), colorectal 92.8 % (n = 14), head and neck 95 % (n = 18), kidney 100 % (N = 9), lung 100 % (n = 22), liver 100 % (n = 11), prostate 100 % (n = 15). The PDF files, source code, LLM prompt, and configurations have been publicly shared via GitHub (https://github.com/grasool/tcga-path-report).

Going digital - Roadmap for implementation

Bryan Dangott, Michael Gellner, Khela Pursell, Marie Johnson, Joann Kirkland, Heather Confiado, Zeynettin Akkus, Kevin Wu, Aziza Nassar

Mayo Clinic Florida, Department of Laboratory Medicine and Pathology, Jacksonville, United States

Email: Dangott.Bryan@mayo.edu

Transitioning from analog to digital workflows with digital pathology involves many considerations which may seem daunting without developing the right plan, expectations, and support. Our department began converting to digital workflows in early 2021 and now has all pathologists reading digitally for greater than 90 % of the case volume (excluding hematology and cytology). We will discuss the process and lessons learned for scanner selection, slide preparation, PACS, LIS integration, infrastructure, deployment, and workflow adoption.

Optimal scanner selection is probably the most critical component for going digital. While there are various scanners on the market, careful consideration should be given to choosing the right scanner for the intended application. An assessment of image quality and magnification requirements should rank highly in the factors to consider. Scanner redundancy and robustness may also be high priorities for scanners deployed in time sensitive environments such as frozen section or other intra-procedural environments. Ultimately, multiple scanner platforms may be needed to accommodate various modalities such as paraffin sections, frozen sections, immunoflourescence, and smears from hematology or cytology samples. The daily slide volume, scanner throughput, and expected turn around time will impact the total number of scanners needed for deployment.

As processes move from manual to automated workflows, slide preparation may be subject to tighter tolerances and requirements. While slide bar codes are common in routine histology, hand labeling methods used by many frozen section laboratories may not be compatible with automated case assembly. In addition, outside consult cases may need to be relabeled in a manner that is compatible with internal accessioning systems. Slide preparation steps such as cover slipping, staining, and defect detection may also impact scanning operations. For example, cover slips or label stickers which extend beyond the edge of the glass slide can impact automated slide handling equipment. Wet slides or adhesive residue can impact slide gripper mechanisms and contribute to scanner downtime. Air bubbles, fingerprints, and debris can cause out of focus regions or obscure histology and thus require re-scan. Tissue layout on the slide can impact scanner throughput or even lead to scanner errors or re-scans if the tissue is too close to the edge of the slide. Stain intensity can impact automated tissue detection algorithms and result in tissue being missed by the scanner.

Pathologists using glass slides may have a variety of methods for sorting and working through flats of glass. The case management system, PACS, and slide viewer should provide a comfortable and efficient environment to transition some of those analog workflows to digital workflows. In many ways, the digital environment has an advantage because the capacity to sort, annotate, and navigate slides can all happen on one screen. However, if core functions are missing or poorly implemented, pathologists may consider these as defects and detractors from going digital. Familiarizing and training pathologists how to optimally use the software is essential. In addition, at the elbow support is helpful in preventing frustration. Finally, pathologist workstations may need to be upgraded to handle slide viewing and case management functions.

Any digital pathology operation requires a plan for hosting and storing whole slide images. Some groups prefer to purge digital images a few months after the case is signed out while others prefer to archive digital files for much longer periods. In either case, the storage plan, budget, and architecture needs to be compatible with the clinical and administrative expectations. There may also be considerations for optimizing the location of the slide preparation and scanning facilities. Topics such as physical plant, environmental systems, network capacity, and redundant storage systems are all relevant to implementation.

Understanding all of the upstream and downstream impacts of digital pathology is critical for successful implementation. Digitizing glass slides is the obvious change. However, more subtle changes include addressing upstream workflow changes in histology and downstream changes in transcription and in the LIS. Centralization of glass slides for scanning allows cases to be distributed electronically. However, the glass may still need to be readily available so that pathologists may request re-scans or physical glass slides if necessary. The LIS and PACS systems can be leveraged to provide insights into the daily workflows to anticipate busy periods and slide volumes. The LIS may need to accommodate and track cases signed out digitally for compliance and regulatory purposes.

While the hardware, software, and infrastructure needs are critical, the people and change management issues are equally important, if not more so. Successful implementation of digital pathology requires a team approach and the support of many different stakeholders. A multidisciplinary team encompassing representatives from groups impacted by digital pathology would be well positioned to manage concerns and set expectations.

Infrastructure for real-time integration of digital tools into the clinical ecosystem

John Philip[1], Luke Geneslaw[1], Julie Garcia[1], Jonathan Alarcon[1], Porselvi Vaithianathan[1], Noreen Wu[1], Susan Fu[1], Jack Cortez[1], Dylan Conroy[1], Mushfiq Rahman[1], Evangelos Stamelos[1], Orly Ardon[1], Ron Pearson[1], Saad Nadeem[1], Chad Vanderbilt[1], Gregory Goldgof[1], Aijazuddin Syed[1], Matthew Hanna[2], Victor Reuter[1], Meera Hameed[1], Ahmet Dogan[1]

[1]Memorial Sloan Kettering Cancer Center, Digtial Informatics and Technology Solutions, Memorial Sloan Kettering Cancer Center, United States

[2]University of Pittsburg Medical Center, Pathology, Pittsburgh, PA, United States

Email: philipj@mskcc.org

Introduction

While the core operations of clinical digital pathology are primarily concerned with the scanning and viewing of slides such that microscopes can be replaced with digital viewers, there are numerous additional software tools available to augment clinical operations within the digital pathology space. For example, algorithms can automate reporting or provide clinical decision support (CDS) to the pathologist or to the scan team for quality control purposes. The integration of these tools into the clinical workflow is complex and in many cases requires collaboration between multiple parties to build custom infrastructure in order to bring the tool online.

Methods

At Memorial Sloan Kettering Cancer Center, we have built custom middleware which provides a platform for the integration of viewing and AI tools to provide decision support, workflow enhancements, or other forms of automation. Specifically, we employ (1) a digital pathology database which gathers image metadata alongside orders data from the laboratory information system (LIS), (2) an API which provides an access point for downstream systems to request digital pathology metadata from the database, and (3) event-driven cloud-based architecture for subscribing to image scans and pushing images and clinical metadata downstream accordingly. This event-driven architecture allows for the filtering of image scans such that models or platforms of a specific scope may receive access to images and associated metadata in accordance with that scope, i.e. an immunohistochemistry model is able to receive immunohistochemistry scans while passing over hematoxylin and eosin scans.

Results

Through this integration layer, we have achieved the integration of 4 distinct viewing and/or AI platforms which subscribe to and receive image files and clinical metadata in near-real-time. These platforms house 5 distinct AI models and generate inferences on the order of 1000 images per day.

Conclusion

By building custom middleware infrastructure, a platform can be created for near-real-time integrations of viewing and AI tools into the clinical ecosystem such that pathologists may interact with these tools for evaluation and use in clinical practice. This integration layer implements automated real-time decision-making around whether images are eligible for one or more available viewers or AI models, and for the automated routing of those images to the associated platform(s) to ensure the data is available on demand for clinical workflows.

Mutation detection in lung cancer using heterogeneity derived from quantitative pathology features

Prasanna Porwal,[1] Fariba Dambandkhameneh,[1] Xiao Li[2], Paolo Ocampo,[2] Yao Nie[1]

[1]Roche Diagnostic Solution, Computational Science & Informatics, Santa Clara, United States

[2]Genentech, Inc., Personalized Healthcare, Data, Analytics and Imaging group, South San Francisco, CA, United States

Email: yao.nie@roche.com

Background and objective

The identification of actionable variants is fundamental for guiding critical therapeutic decisions in the management of an increasing number of cancer types, including non-small cell lung cancer. However, conventional

DNA-based genetic testing methods are expensive and time-consuming. In this study, we test the hypothesis that morphological heterogeneity in tumor cell populations reflects the level of genetic variation, which in turn provides information about the presence of mutations. This study aims to harness tumor heterogeneity derived from Quantitative Pathology Features (QPFs) for the classification of known (mutation or fusion) and unknown drivers in lung cancer patients.

Methods

The analysis employed a comprehensive dataset from Foundation Medicine Inc., consisting of 2422 whole slide image (WSI) slides. This dataset was divided into a training set (60 %) and two distinct testing sets (20 % each). Random Field-Of-View (FOV) images were extracted from tumor regions in the WSI slides. These FOV images were further refined using Reinhard color normalization technique prior to feature extraction. QPFs characterizing individual tumor cell nuclei were then extracted and aggregated at the WSI level, providing a comprehensive tumor morphology information for each slide. The WSI level features were used to compute four different heterogeneity metrics for each QPF, namely, Standard Deviation, Kolmogorov–Smirnov statistics, Quadratic Entropy and Outlier Percentage. These morphological heterogeneity metrics for all QPFs were concatenated and used to train and test the XGBoost model for mutation detection. Model optimization was achieved by performing a five-fold cross-validation technique on the training set, while the model's stability was ascertained using two separate testing sets. The model results were cross-referenced with ground-truth labels detected through an orthogonal mutation detection method, next-generation sequencing (NGS). Model performance was assessed using the area under the precision-recall curve (PR-AUC) due to class imbalance in the data distribution.

Results

Findings revealed heterogeneous patterns indicative of different genetic drivers, with driver oncogenes (KRAS, EGFR, BRAF mutations, and ALK, ROS1, RET fusions) showing distinctive heterogeneity profiles compared to tumor suppressors. Notably, remarkable predictive performance was achieved, with a PR-AUC of 0.88 and 0.85 respectively on two separate test sets. This signifies the high precision-recall capacity of our model in distinguishing classes of genetic drivers based on tumor heterogeneity.

Conclusions

Our findings demonstrate the potential of quantitative pathology-based tumor heterogeneity to classify genomic drivers. It reveals that the heterogeneity inherent in tumors provides a valuable source of information for characterizing the molecular landscape of cancer. Importantly, this method can potentially enable more accurate and efficient diagnosis of cancer, as well as guide treatment strategies based on the detected mutations. This work sheds light on the underexplored connection between tumor heterogeneity and genomic alterations, underscoring the potential of morphological heterogeneity to serve as a novel predictive biomarker.

Quality processes in clinical digital pathology operations at memorial sloan kettering cancer center

Orly Ardon, Allyne Manzo, Marc Labasin, Peter Ntiamoah, Luke Geneslaw, John Philip, Matthew G. Hanna, Victor Reuter, Meera Hameed
Memorial Sloan Kettering Cancer Center, Pathology and Laboratory Medicine, New York City, NY, United States
Email: ardono@mskcc.org

Background

Pathologists are becoming increasingly familiar with digital pathology for primary diagnostic use, and the field is witnessing a growth in pathologist buy-in and comfort using the technology. However, in order to enable successful clinical implementation, a suitable, multi-component digital pathology infrastructure must be in place. Each laboratory and organization will require foundational components, including hardware, software, and network components to support the digital pathology system. The exact infrastructure blueprint chosen for institutional deployment will depend on the laboratory use cases, resources, and overall strategy. An often-overlooked component of the clinical digital pathology operation is the need to ensure quality whole slide imaging workflows that will control image quality for the downstream pathology use cases. These quality considerations need to be applied in all phases of the digital pathology process: preanalytical, analytical and post analytical phases and get updated with new technologies and workflows that are being adopted and as scanning volumes increase. The objective of this study is to identify quality processes required for clinical grade digital pathology operations.

Methods

Our team at Memorial Sloan Kettering Cancer Center (MSK) started developing a quality management system (QMS) shortly after the integration of digital pathology into our clinical systems in 2020. With the expansion of our operation in recent years, a departmental wide effort to identify additional potential quality needs for the different digitization phases was identified. The effort included multiple teams and a systemic review of all digital pathology related workflows from specimen accessioning to slide storage that led to changes in some of these workflows. The financial implication of these new workflows was calculated to ensure long term sustainability.

Results

Identified new workflows in different phases of the pathology operation resulted in changes in glass slide delivery, the addition of automated tracking capabilities, added quality dashboards, turnaround time monitoring and overall lean operations that were integrated in our daily workflows. New centralized scanning laboratory workflows resulted in a decrease in the downtime of the digital scanners. A major effort was given to finding an automated solution for the labor-intensive image quality review process, which was previously found to be essential for our large-scale clinical pathology effort.

Conclusions

The transition to fully digital pathology operations requires adjustments to current workflows. A systemic approach to identify quality improvements should be ongoing as technologies and scanning operations expand. The talk will describe these integration efforts and the financial implications of these improvements on our overall pathology operations.

Informatics and pathology innovation highlights of the college of American pathologists

Marilyn Bui[1], Monica de Baca,[2] Joe Saad[3]
[1]Moffitt Cancer Center, Pathology, Tampa, United States
[2]Synmex America Inc., Medical Affairs, LincoInshire, United States
[3]Methodist Health System/UT Southwestern Medical Center, Pathology, Dallas, TX, United States
Email: marilyn.bui@moffitt.org

This presentation is designed for those interested in understanding the strategic initiatives of the College of American Pathologists (CAP) regarding Digital Pathology (DP) and AI reimbursement. It will also address the impact of recent FDA Laboratory Developed Test (LDT) rules on pathology, DP, and AI. Attendees will gain insights into the resources CAP provides in these areas and learn about CAP's efforts to advance the practice and science of DP and AI. Additionally, the presentation will explore how the

pathology community can collaborate to advance various areas, including implementation, quality improvement, standardization, cancer protocols, electronic reporting, and the regulatory science of DP and AI.

The CAP established the Council on Informatics and Pathology Innovation (CIPI) in 2022. It is charged to identify and recommend strategic direction on current and emerging medical information science, date science, and computational technologists that could impact the practice of pathology; provide informatics domain information and expertise to the CAP in furtherance of its programs and mission; and support appropriate engagement with external stakeholders. CIPI has embarked on numerous initiatives in education, quality, proficiency testing, and regulatory science of DP, AI, and beyond. The leadership and resources provided by CIPI and its various committee are valuable to pathologists and laboratories undergoing a digital transformation. In addition, the CAP's Council on Government and Professional Affairs (CGPA) is working diligently on behalf of pathologists to identify current and emerging issues in the legislative, regulatory, and private sector arenas that may impact the practice of pathology; to develop policies and strategies to positively influence these issues to benefit patients and pathologists; and to implement these policies and strategies by educating pathologists, conducting advocacy programs and maintaining liaison with health-related organization.

The presenters and panelists are pathologists and the chairs of the CIPI, CGPA, and Digital and Computational Pathology Committee. A Q&A session will follow the presentation. This one-hour session aims to provide valuable insights and foster collaboration among pathologists and laboratories experiencing a digital transformation. The leadership and resources offered by CAP, through CIPI and CGPA, are instrumental in navigating the evolving landscape of digital pathology and AI.

DICOM and pathology: It's a hit, not a myth

Nabile Safdar, Thomas Crawford
Society for Imaging Informatics in Medicine, Learning, Leesburg, VA, United States
Email: nabile.m.safdar@emory.edu

Traditionally rooted in radiology, Digital Imaging and Communications in Medicine (DICOM) standards offer significant benefits for digital pathology, including improved interoperability, streamlined workflows, and enhanced data management. This presentation will explore the crucial role DICOM plays in the digital pathology landscape, providing insights into its practical applications and future potential. Attendees will gain a comprehensive understanding of how DICOM can bridge the gap between radiology and pathology, fostering a more collaborative and efficient diagnostic environment.

During this session, we will discuss the technical and practical aspects of implementing DICOM in pathology, address common challenges, and highlight success stories from institutions that have successfully integrated DICOM into their pathology practices. Whether you are a pathologist, radiologist, IT professional, or healthcare administrator, this session will equip you with the knowledge and tools necessary to leverage DICOM for improved diagnostic accuracy and patient outcomes.

Multi-expert workflow for image patch labelling: An example in megakaryocyte detection

Ethan Okoshi[1], Wataru Uegami,[2] Tom Bisson,[3] Sosuke Ishijima,[1] Kris Lami,[1] Andrey Bychkov,[2] Junya Fukuoka[1]
[1]Nagasaki University Graduate School of Biomedical Sciences, Department of Pathology Informatics, Nagaski University School of Biomedical Sciences, Japan
[2]Kameda Medical Center, Department of Pathology, Kamogawa, Japan
[3]Charité - Universitätsmedizin Berlin, Department of Pathology, Berlin, Germany
Email: bb55324008@ms.nagasaki-u.ac.jp

Background

Digital pathology has, in the past decade, seen a boom in machine learning (ML) models trained on histopathological images for diagnostic and assistive purposes [1]. ML models can assist pathologists on tasks of identifying morphological structures and provide a hedge against the issue of interobserver variability [2,3]. However, researchers must often ask pathologists to complete meticulous hand-drawn annotations for tens or hundreds of whole slide images (WSI) to generate the necessary data to train these models. The alternative is to label WSI at the slide level, which is less effort for pathologists, but is far less efficient when training ML models [5]. We propose a new method for generating patch-level labels which can be used as training data for ML vision models. We then applied this method to create two datasets, one for training a hematopoietic bone marrow (BM) detection model, and another for training a megakaryocyte detection model. Megakaryocyte identification suffers from interobserver variability [4], yet is crucial in accurately diagnosing diseases such as immune thrombocytopenic purpora (ITP), myeloproliferative disorders, and myelodysplastic syndromes (MDS) [6]. In a workflow featuring the models working in tandem, the first model will detect patches of hematopoietic BM tissue, which will then be fed into the megakaryocyte model to finally output an annotated WSI with predicted BM and megakaryocytes highlighted.

Methods

We developed an image sorting app, SortImg, which allows pathologists to easily sort histology image patches into pre-determined categories. WSI are tiled or patched; these patches then are automatically clustered by a clustering algorithm. SortImg is then used to purify these clusters to ensure that only the desired morphologies or cell types are included. These image patch sets can be sorted by multiple pathologists, which can be leveraged to reduce interobserver variability, i.e. by taking a majority vote, clustering pathologists to obtain varied features. For our case, we first extracted 256px x 256px patches from bone tissue WSI and sorted these patches into 50 % hematopoietic BM (hematopoietic cells and adipocytes) vs. “other” (including 50 % bone, fibrosis, blood, or background). We then extracted 100px x 100px patches from the same WSIs and classified these as containing megakaryocytes vs. not, both using the SortImg app. We collected 2000 total patches for each set, which was then split into 80 % train and 20 % validation. We finetuned two ResNet101 models pretrained on ImageNet-1 k for each dataset.

Results

The BM detection model was able to differentiate between patches with 50 % hematopoietic BM and “other” patches with 98.25 % accuracy, 0.9996 ROC AUC, and 0.9822 F1 score. On the validation set ($n = 400$), the model was able to correctly classify all 200 negative patches, but had 7 false negatives out of the 200 true positive patches. For the megakaryocyte detection model, the model achieved 99.00 % accuracy, 0.9978 ROC AUC, and 0.9899 F1 score. On the validation set (n = 400), the model achieved 0 false positive predictions of 200 true negatives and 4 false negative predictions of 200 true positives.

Conclusions

We were able to train high-performance models to recognize hematopoietic BM tissue and megakaryocytes using a new pathologist-friendly workflow tool, SortImg. Labeling at the patch level is finer than at the slide level, allowing for computer vision models to learn morphology with greater specificity, while taking less time than meticulous hand-drawn annotations. This new method of labeling image patch sets could greatly improve dataset compilation workflows for researchers seeking to build assistive computer vision models for a wide range of tissues and diseases. In the current age of human oriented medicine, we may have surpassed the need for single ground truth labels, opting for a more

nuanced multi-label approach. SortImg, with its multi-pathologist functionality, is a tool that is suited for such a task.

References

1. McGenity C, Clarke EL, Jennings C, et al. Artificial intelligence in digital pathology: a systematic review and meta-analysis of diagnostic test accuracy. Npj Digit Med. 2024;7(1):1-19. doi:10.1038/s41746-024-01106-8

2. Lami K, Bychkov A, Matsumoto K, et al. Overcoming the Interobserver Variability in Lung Adenocarcinoma Subtyping: A Clustering Approach to Establish a Ground Truth for Downstream Applications. Arch Pathol Lab Med. 2023;147(8):885-895. doi:10.5858/arpa.2022-0051-OA

3. Homeyer A, Geißler C, Schwen LO, et al. Recommendations on compiling test datasets for evaluating artificial intelligence solutions in pathology. Mod Pathol. 2022;35(12):1759-1769. doi:10.1038/s41379-022-01147-y

4. Wilkins BS, Erber WN, Bareford D, et al. Bone marrow pathology in essential thrombocythemia: interobserver reliability and utility for identifying disease subtypes. Blood. 2008;111(1):60-70. doi:10.1182/blood-2007-05-091850

5. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat Biomed Eng. 2021;5(6):555-570. doi:10.1038/s41551-020-00682-w

6. Kumar V. Robbins and Cotran Pathologic Basis of Disease. Saunders/Elsevier; 2010.

Opportunities and challenges in refining grading of neuroendocrine tumors using digital pathology

Monika Vyas[1], Buthania Hakamy[1], Ruben Oganesyan[1], Nandan Padmanabha[1], Raul Gonzalez[2]

[1]Beth Israel Deaconess Medical center, Pathology, Boston, United States

[2]Emory University Medical center, Pathology, Atlanta, United States

Email: mvyas1@bidmc.harvard.edu

Grading of gastroenteropancreatic (GEP) well-differentiated neuroendocrine tumors (WD-NETs), based on the current WHO classification, depends on the Ki67 proliferation index (PI) and mitotic index (MI) per 2 mm^2. Accurate classification and grading are important for determining treatment strategies and prognostication of these tumors. The WHO recommends manual estimation as the gold standard. The criteria for performing manual PI involve counting the number of positive cells in at least 500 tumor cells in the area of highest apparent proliferative activity (the so-called hotspot region). For MI estimation, the WHO recommends counting the number of mitotic figs. (MF) in a 2 mm^2 area in the most mitotically active region. Both measures are considered when assigning the final grade. The cutoffs for PI and MI are as follows:

- **Grade 1**: PI <3 %, MI < 2 per 2 mm^2
- **Grade 2**: PI = 3–20 %, MI = 2–20 per 2 mm^2
- **Grade 3**: PI 20 %, MI 20 per 2 mm^2

Estimation of PI and MI is cumbersome and subject to interobserver variation. The challenges in Ki67 PI estimation start with choosing a block for staining, defining a positively staining cell, and determining a hotspot area, among other difficulties. Pathologists often select a block showing more mitoses, atypia, or high-risk features; however, none of these may be present in a given tumor, making block selection arbitrary. Evaluation of Ki67 on multiple blocks is cumbersome and not cost-effective. Choosing the right focus for MI estimation is limited by the human eye and compromised by tissue processing/preservation artifacts and inflammatory cells.

Our group has been working on using computational pathology to refine and simplify practices in grading GEP WD-NETs. Image analysis makes it possible to study multiple tissue blocks to perform PI and MI for accurate grading. This is feasible with both camera-captured images and whole-slide images (WSI). Automated hotspot detection minimizes subjectivity in PI estimation. We have also compared multiple open-source platforms to perform PI on WSI and camera-captured images, finding concordant grade assignments in most cases.

As digital pathology adoption gains traction, one must adapt to using WSI for MI, but no guidelines are currently available for this. We conducted a pilot study comparing MI estimation on WSI to the gold standard method on glass slides. Our methodology for MI estimation on glass slides includes low-power panning to find hotspots. Once such an area is found, the first MF is counted, and MFs in the next 10 consecutive 40× fields (2 mm^2) are recorded. On WSI, we tested two strategies:

1. Placing a grid overlay on the WSI, with each square measuring 2 mm^2. The slide is panned at 10× to find any MF. Each MF is annotated, and the maximum number of MFs in any one square is recorded as the MI.
2. Panning the slide at 10×, annotating the first MF, and then drawing a 2 mm^2 square annotation around it. Additional MFs within this square are counted, and the total number is recorded as the MI.

We selected 10 cases (4 small intestine and 6 pancreas) for assessment using these methods. We found excellent agreement between the glass and digital grid methods (interclass correlation coefficient, ICC = 0.97), the digital square annotation method (ICC = 0.98), and between the two digital methods (ICC = 0.99). There was no difference in the final grade assigned by any method. However, MI estimation on WSI took longer (average 8.3 min per slide for the grid method, 7.3 min per slide for the square annotation method, compared to 3.05 min per slide for glass slides). The reasons for this are multiple. Pathologists are not used to assessing MI on WSI. The inability to perform fine adjustments can limit the definitive identification of MF. Scanning a WSI at 10× is cumbersome and time-consuming.

Despite this, we observed some advantages of using WSI. Counting in a standard area of 2 mm^2 (either grid or annotation) is easier than counting MF in consecutive high-power fields. Furthermore, annotating the MF on WSI allows for the comparison of multiple hotspots and future review. Machine learning algorithms can reduce the time spent finding MF, aid in the accurate detection of hotspots, and minimize subjectivity. Deploying these algorithms on sections from multiple tumor blocks can increase accuracy of grade assignment. However, developing these tools is fraught with challenges, such as tumor cell detection, staining intensity, debris, fixation artifacts, chromatin heterogeneity, and inflammatory cells particularly neutrophils. Balancing a sensitive and specific algorithm is often difficult to achieve.

The aim of this presentation is to share our experiences and the lessons we have learned. Much work remains in developing a standardized, consistent, and accurate approach to grading GEP WD-NETs. Nevertheless, we believe that image analysis will be critical in refining the grading of these tumors, and using open-source technology will be very helpful in ensuring wider applicability and uniformity in grading.

Driving the advancement of digital pathology and AI through collaboration: An allied society panel discussion

Michael Quick[1], Junya Fukuoka,[2] Norman Zerba,[3] Donald Karcher,[4] Marilyn Bui[1]

[1]Digital Pathology Association, Carmel, Indiana, USA

[2]Asian Society of Digital Pathology, Singapore

[3]European Society of Digital and Integrative Pathology, Lisboa, Portugal

[4]College of American Pathologists, Chicago, Illinois, USA

Email: michael.quick@hologic.com

Since its inception in 2009, the Digital Pathology Association (DPA) has been at the forefront of driving innovation in digital pathology (DP) and augmented intelligence (AI) within healthcare and life sciences. Recognizing the power of collaboration, DPA has forged strong partnerships with like-minded professional societies and organizations. Join us for this exclusive panel discussion featuring the Presidents of four distinguished allied

societies. Together, they will provide a high-level overview of the current landscape of DP and AI, highlighting both the advancements and challenges facing the field. Panelists will delve into the strategies and initiatives implemented by their respective societies to propel DP and AI forward. They will also explore opportunities to strengthen collaborations and address critical issues such as education, regulation, and reimbursement. This interactive session is designed to foster open dialogue and encourage audience participation in this vital conversation about the future of pathology.

PV24 Poster Abstracts

Decoding colon cancer recurrence: Unveiling accurate predictions with attention-guided deep CNN

Mohammad Alexanderani, Mohamed Omar, Matthew Greenblatt, Luigi Marchionni

Weill Cornell Medicine, Pathology & Laboratory Medicine, New York, NY, United States

Email: mom4011@med.cornell.edu

Background

Up to 60 % of colon cancer patients are at high risk of cancer recurrence, yet accurate and timely prediction tools are lacking. Leveraging whole slide images (WSIs) and deep learning models, we aimed to develop precise algorithm for prediction of colon cancer recurrence. Thus, enables risk stratification for optimized therapeutic interventions and improved health outcomes.

Design

We developed an attention-based deep learning model to predict colon cancer recurrence using digital slides of hematoxylin and eosin-stained colon tissue biopsies. The dataset, sourced from the TCGA database, underwent preprocessing steps including tissue segmentation, tiling, and histopathological feature extraction. The labeled dataset was split into training (85 %), validation (7.5 %), and unseen testing sets (7.5 %). Model performance was evaluated using standard metrics including AUC (Area Under the Curve), sensitivity, specificity, and accuracy. Interpretability was achieved through attention heatmaps, providing insights into relevant histological features.

Results

235 WSIs satisfied the criteria and were classified as either post-therapeutic colon cancer recurrence or no recurrence. The model consistently showcased strong performance across both the validation and testing datasets, achieving robust performance with AUC of 0.86. On the unseen testing set, the model attained a sensitivity of 1, specificity of 0.80, and an accuracy of 0.83.

Conclusion

Our study introduces an innovative deep learning approach that accurately predicts colon cancer recurrence using solely H&E Whole Slide Images (WSIs). It highlights the substantial contribution of the stroma in recurrence prediction. These findings carry critical implications for optimizing therapeutic interventions and implementing effective surveillance strategies in personalized colon cancer management.

Institutional development and prospective validation of a clinical grade model for prostate biopsies

Ramin Nateghi,[1] Madeline Saft,[2] Ruoji Zhou,[3] Jeffery A. Goldstein,[1] Ximing Yang,[1] Ashley E. Ross,[2] Lee A.D. Cooper[1, 4]

[1]Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[2]Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[3]Department of Pathology, University of California Los Angeles David Geffen School of Medicine, Los Angeles, CA

[4]Chicago Chan Zuckerberg Biohub, Chicago, IL, USA

Email: ramin.nateghi@northwestern.edu

Introduction

Large collections of annotated whole slide images are a common ingredient in the development of many successful AI tools. We sought develop a model with clinical grade performance in prostate cancer detection and Gleason scoring using real world data generated at Northwestern Memorial Hospital. Our institution is in the early phases of our digital pathology implementation, so we established a research scanning and informatics pipeline to scan, link, and annotate prostate biopsy slides generated at our hospital, as well as efficient software tools for training AI models on enterprise scale datasets.

Methods

We established an IRB approved protocol for prospective scanning of prostate biopsy slides and scanned all slides from positive biopsies collected at Northwestern Memorial Hospital from August 2021 and March 2023 using a Leica GT450 scanner (total 21,695 slides from 1042 subjects). Subsequently, for the month of April we scanned all slides from all biopsies to generate a validation dataset that is representative of our patient population (total 1922 slides from 95 subjects). Our Enterprise Data Warehouse team developed and validated a pipeline for extracting structured data for prostate biopsy slides from Epic Beaker including subject, specimen, block, primary and secondary Gleason pattern, level, stain, and other pathology variables. Scanned images were linked to the pathology database by reading barcodes in the scanned images and using manual review of optical character recognition where barcode reading fails. We developed a carcinoma detection and Gleason pattern prediction model in two stages. First, we used the PANDAS dataset to train a simple EfficientNetB0 convolutional network to classify high power fields for the presence or absence of carcinoma (all patterns). Second, we used latent features from this network as an embedding to train a weakly supervised model to detect carcinoma at the block-level and to predict primary and secondary Gleason pattern. This model uses a novel transformer approach to simultaneously model interactions between fields at both near and far distances.

Results

The carcinoma detection AUC of our model is 0.991, with a sensitivity of 97.3 % and a specificity of 97.4 %. The primary and secondary Gleason pattern AUCs for our model were 98.5 % and 98.0 % respectively. We also found that among equivocal tissue blocks where immunohistochemistry was ordered, our model predictions were perfect on the 50 % of blocks with the most confident model scores.

Discussion

We envision this model being used in a second read capacity where discordant pathologist and model reads that could result in changes to clinical management are used to identify cases for additional review. We are currently working with our healthcare system to develop a framework to deploy this and other models to collect additional data on their use and impact on quality. Future work includes performing a consensus review of our validation dataset and comparing our task-specific embedding approach to newly published foundation models.

Histotype Px® Colorectal as an AI biomarker to guide adjuvant therapy in colon cancer patients

Vidya Arole,[1] Sepp DeRaedt,[2] Jessica Maupin,[1] Sandra Boutros,[1] Demond Handley,[1] Anne Noonan,[1] Eric Miller,[1] Giovanni Lujan[1]

[1]The Ohio state university, Pathology, Columbus, United States

[2]DoMore diagnostics, DoMore diagnostics, DoMore diagnostics, Norway

Email: Vidya.arole@osumc.edu

Background

The decision to offer adjuvant chemotherapy for patients with resected high-risk stage II and stage III colon cancer remains challenging for clinicians. Per current guidelines, most patients in these groups should receive adjuvant chemotherapy, but up to 80 % of patients do not benefit from this treatment and only experience side effects. On the other hand, some patients are not advised to receive adjuvant chemotherapy because of small perceived clinical benefit. Thus, it is critical to equip clinicians with additional tools to aid in decision-making for this patient group. One of the known shortcomings has been that the current histopathologic grading of tumors into three tiers as well, moderately and poorly differentiated has shown inconsistent correlation with clinical outcome. Histotype Px Colorectal is a novel state-of-the-art artificial intelligence (AI)-based biomarker validated for R0 resected stage II and III colorectal adenocarcinoma that analyzes digitized routine whole slide images of hematoxylin and eosin-stained formalin-fixed, paraffin-embedded tumor resections. Combined with clinical parameters, patients are stratified into distinct low, intermediate, and high-risk groups. This pilot study seeks to determine the association between Histotype Px Colorectal and patient outcomes in an independent cohort from the United States.

Methods

This was a retrospective analysis of patients diagnosed with surgically resected pathological stage II-III colon adenocarcinoma at Ohio State University from 2016 to 2020. Clinical parameters (pT status, pN status, number of lymph nodes) and clinical outcome data were extracted from the electronic medical record. Anonymized slides were prepared for each patient and digitized using the Aperio AT2 scanner. Histotype Px Colorectal biomarker was blindly applied to each scan and subsequently linked to clinical outcomes. The primary outcome of interest was cancer-specific mortality and secondary outcomes were all-cause mortality and recurrence. Univariate Cox proportional-hazards regression models were used to assess the association between the Histotype Px Colorectal biomarker and disease outcomes. A threshold of 0.05 was used to determine statistical significance.

Results

Baseline characteristics of the 159 eligible patients included a median age of 63 years (range 22–91), 52 % were female, 52 % had stage III disease, 64 % had right-sided tumors, and 31 % had microsatellite instable tumors. 19 % of stage II and 71 % of stage III patients received ACT, respectively, with FOLFOX in 61 %. Median follow-up for all patients was 54.2 months. Of those who received adjuvant chemotherapy ($N = 72$), 19.4 %, 26.4 %, and 54.2 % were classified as high, intermediate, and low-risk by the Histotype Px Colorectal biomarker, respectively. Of those did not receive adjuvant chemotherapy ($N = 87$), 9.2 %, 17.2 %, and 73.6 % were classified as high, intermediate, and low-risk, respectively. When evaluating those patients who did not receive adjuvant chemotherapy, an association between Histotype Px and all-cause mortality (23 events) was observed for high-risk (HR = 4.41, 95 % CI 1.37–14.24) and intermediate-risk (HR = 2.67, 95 % CI 1.09–6.54) patients which was statistically significant ($p = 0.017$). Significant associations were also observed for the high-risk (HR = 15.66, 95 % CI 3.74–65.52) and intermediate-risk (HR = 5.61, 95 % CI 1.31–23.94) patients with regards to cancer-specific mortality (12 events, $p < 0.001$) and recurrence (18 events; high-risk with HR = 28.67, 95 % CI 8.32–98.77; intermediate-risk with HR = 5.89, 95 % CI 1.84–18.84, $p < 0.001$). For patients who did receive adjuvant chemotherapy, no statistically significant associations were observed between Histotype Px risk group and all-cause mortality (16 events, $p = 0.344$) or cancer-specific mortality (12 events, $p = 0.103$), but with a trend observed for recurrence (20 events, $p = 0.090$) given the limitations of a small sample size. When separated by stage, Histotype Px risk group was significantly associated with recurrence in intermediate-risk patients (9 events; HR = 4.73, 95 % CI 1.29–17.41, $p = 0.019$) but not with cancer-specific mortality (6 events, $p = 0.132$) or all-cause mortality (13 events, $p = 0.903$) in those with stage II disease. For patients with stage III disease, Histotype Px risk group was significantly associated with all-cause mortality (26 events; high-risk with HR = 3.14, 95 % CI 1.22–8.04; intermediate risk with HR = 2.38, 95 % CI 0.94–6.04, $p = 0.049$), cancer-specific mortality (18 events; high-risk with HR = 5.96, 95 % CI 1.92–18.56; intermediate risk with HR = 2.37, 95 % CI 0.61–9.23, $p = 0.007$), and recurrence (29 events; high-risk with HR = 4.40, 95 % CI 1.92–10.06; intermediate risk with HR = 1.67, 95 % CI 0.61–4.60, $p = 0.002$).

Conclusions

The findings from this small pilot study highlight the potential utility of this AI biomarker in assisting clinicians with decisions about offering adjuvant chemotherapy to patients with resected stage II and III colon cancer. Identifying patients who would benefit from an escalation of therapy as well as potential therapy de-escalation serves as a critical next step in advancing personalized medicine. Further research including a larger and more diverse patient cohort and subsequent clinical studies are planned to strengthen these initial findings and to integrate this promising biomarker in the clinic.

Transformative pathomics: Predicting gene expression in pancreatic cancer using whole slide images

Shrey Sukhadia, Bing Ren

Dartmouth-Hitchcock Medical Center, Pathology and Laboratory Medicine, Lebanon, NH, United States

Email: shrey.s.sukhadia@hitchcock.org

ABSTRACT

Pancreatic ductal adenocarcinoma (PDAC) presents significant challenges for early diagnosis and effective treatment due to its aggressive nature and late-stage detection. Traditional diagnostic methods, which rely on pathologists microscopically analyzing H&*E*-stained FFPE tissue slides, provide essential insights but can benefit from enhanced precision for early and accurate diagnosis. Pathomics, the study of quantitative imaging from such samples, aims to complement traditional methods by uncovering intricate tissue and cellular details. Our research focuses on leveraging pathomic features extracted from whole slide images (WSIs) to predict gene expression profiles, integrating histopathological and molecular data to improve molecular profiling and targeted therapy for pancreatic adenocarcinoma. By correlating pathomic features with gene expression, we aim to identify novel biomarkers that can facilitate early detection, better prognostic indicators, and personalized treatment strategies, ultimately advancing the clinical management of pancreatic cancer.

Background

Pancreatic cancer is a highly lethal malignancy, ranking as the fourth leading cause of cancer-related deaths worldwide, with a five-year survival rate of less than 10 % due to its aggressive nature, late-stage diagnosis, and limited efficacy of current treatments. The most common type, PDAC, is known for rapid progression and resistance to conventional therapies.

Early detection and accurate prognosis are critical but challenging due to the lack of specific biomarkers. Current diagnostic methods are often inadequate, highlighting the urgent need for novel biomarkers. One promising approach involves extracting pathomic features from whole slide images (WSIs) using advanced computational techniques to analyze tumor histopathology, providing insights into tumor behavior and clinical outcomes. Integrating pathomic and genomic data through advancements in AI and machine learning can facilitate the discovery of new biomarkers, leading to better diagnostic tools, prognostic indicators, and personalized treatment strategies for pancreatic cancer patients.

Method

We annotated 195 regions of interest (ROIs) on whole slide images (WSIs) of pancreatic adenocarcinoma specimens from The Cancer Genome Atlas (TCGA) Pancreatic Adenocarcinoma (TCGA-PAAD) registry. The corresponding gene expression data, measured as FPKM using RNASeq assays, were also downloaded from the TCGA-PAAD registry. Additionally, 83 pathomic features, including cellular morphometry and the intensity and gradient of Hematoxylin and Eosin (H&E) staining of the respective tissues, were extracted from the ROIs using a custom histomics analysis pipeline. A Pearson-based correlation analysis followed by hierarchical clustering of the 83 pathomic features and the gene expression (FPKM measures) of 20,000 genes was performed using the correlation module of a multiomic software named “ImaGene”. Pathomic feature clusters that exhibited significant correlations with clusters of gene FPKMs were selected for further multi-task machine learning analysis.

The dataset of 195 ROIs was randomly split into training (80 %) and testing (20 %) sets. A Multi-task Elastic Net (MTEN) model was trained using the pathomic and gene FPKM features from the training set, employing the Machine Learning module in ImaGene. Five-fold cross-validation of the training set was conducted to minimize data overfitting. The trained model was subsequently tested on the testing set, and the Area Under the Receiver Operating Curve (AUROC) and the coefficient of determination (RSirintrapun et al. (2017)[2]) were measured to evaluate the prediction of gene FPKMs from pathomic features in the testing set. These predictions were assessed for significance using a permutation-of-label approach, and the resulting *p*-values were recorded. Finally, the set of predicted genes was queried against existing literature to identify their biological significance as reported by other orthogonal studies in pancreatic cancer and other human malignancies.

Results

The MTEN model predicted high expression of a set of nine genes at high AUC (AUC 0.8) and R^2 (R^2 0.3) in the testing set. These predicted genes are: YBX1, SUOX, SNX16, ECE2, IGKV1OR-2, KCNJ2, ODF3, RXFP3, and AC013489. The key pathomic features that contributed to the prediction of these nine genes include: the equivalent diameter of cells, the number of pixels in the convex hull image (which is the smallest convex polygon that encloses the region of interest), and the length of the major and minor axes of the ellipse that shares the same normalized second central moments as the annotated tumor region of interest.

Biological relevance of the identified gene expression

A literature review of the gene set from the pathomics-based MTEN model revealed key findings: YBX1 is crucial for PDAC development, invasion, and cisplatin resistance. SUOX is linked to mitochondrial metabolism and poorer prognosis in PDAC. Elevated SNX16 enhances pancreatic cell migration and invasion. KCNJ2 is associated with tumor-associated macrophage infiltration in various cancers. IGKV1OR-2 and ODF3 are overexpressed in several cancer types. RFXP3 is influenced by DNA damage and oxidative stress, common in cancer, while AC013489, a long non-coding RNA, is documented in liver and breast cancers.

Conclusion

This study highlights cellular size and shape-based markers as indicators of biologically relevant gene expression, linking cellular phenotypes to biological processes in PDAC.

Can artificial intelligence be leveraged for automated evaluation of donor kidney frozen sections?

Keefer Wu,[1] John Liang[2], Patricia Tsang[2]
[1]Northeastern University, Computer Science, Boston, MA, United States
[2]MedStar Washington Hospital Center, Pathology & Laboratory Medicine, Silver Spring, MD, United States
Email: keeferwuu@gmail.com

Background

Intraoperative frozen section evaluation of transplant donor kidneys is challenging for many general pathologists. A pilot study was conducted to assess if artificial intelligence (AI) has the potential to be leveraged as a tool to aid the interpretation of whole-slide scanned images of kidney frozen sections.

Method

A small-scale pilot utilizing 132 kidney still images was conducted to train a publicly available AI base program to identify certain parameters recommended by the Organ Procurement & Transplant Network: 1) glomeruli, 2) global sclerosed glomeruli, 3) nodular mesangial glomerulosclerosis, and 4) arteriolar hyalinosis. The accuracy and level of confidence of the analyses were recorded on 37 unknown kidney digital images scanned at 20×.

Results

After training, the AI program was able to correctly identify 96 of the 98 glomeruli (98 %) on the unknown slides, 15 of the 16 globally sclerosed glomeruli (94 %), 12 of the 12 nodular mesangial glomerulosclerosis (100 %), and 16 of 17 arteriolar hyalinosis (94 %). False positive identification was seen in two instances, including an atherosclerotic artery being identified as global glomerulosclerosis (0.55 level of confidence) and a small artery with medial thickening as arteriolar hyalinosis (0.78 level of confidence). Of note, three glomeruli correctly identified by the AI program were originally missed by a practicing general pathologist as these structures were only partially present at the edges of the biopsies. A rather subtle hyalinized arteriole identified by the AI program was also overlooked by a general pathologist. The software took no more than one second to yield the result per still frame.

Conclusion

This pilot study has demonstrated that AI has the potential to be developed into a fast and useful tool for identifying parameters that are important for the evaluation of transplant donor kidney frozen sections. The current study suggests that a properly trained AI program can enhance the pathologist's ability to identify subtle findings. In the intraoperative setting, particularly during off-hours, an AI-based automated tool can potentially help general pathologists improve the consistency and efficiency of donor kidney evaluation. Building on the success of this pilot, a much larger training set of digital images has been planned to improve the accuracy of the image analysis before the program can be implemented for pathology practice. Furthermore, a plan is in place to expand the study to include additional pathologic parameters, such as vascular disease grading and interstitial fibrosis/tubular atrophy.

Implementation of digitally native pathology workflows with multi-system integrations

Dibson Dibe Gondim, Mercia Gondim
University of Louisville, Pathology, Louisville, KY, United States
Email: dibson.gondim@louisville.edu

Digital pathology and artificial intelligence are rapidly becoming top priorities for laboratories, underscoring their strategic importance. The convergence of technologies like cloud computing, artificial intelligence, high-throughput scanners, and system integration is driving large-scale digital pathology deployments. Seamless integration between Digital Pathology Systems and Laboratory Information Systems (LIS) is crucial for optimal usability. Moreover, further integration across multiple systems not only enhances traditional workflows but also paves the way for entirely new and innovative processes that were previously unattainable due to the physical limitations of glass slides. This work aims to showcase the implementation of digitally native pathology workflows that support clinical, educational, research, and operational activities within an academic pathology department.

The basic framework used for building digitally native pathology workflows involves understanding the use case and the value it will create, identifying the necessary data sources across one or multiple systems, processing and integrating the data, and developing a data presentation layer, which includes webpages, mobile applications, emails, or integration within other health system applications.

We implemented various solutions to present data on web pages and through email reports. The primary data sources included the LIS, digital pathology system — the Paige platform utilizing the tag and AI results application programming interface (API), a vector database for pathology reports, and the service scheduling system. Using data from the Paige tag API, which provides information on slide-level tags created in Paige, we developed a dashboard to support multiple operations. These operations include organizing cases for consensus conferences and tumor boards. The slide-level tags are also used for curating educational cases, creating prospective datasets for research, and documenting the most relevant slides of cases, such as those that define the pathologic stage or show lymphovascular invasion. The resident feedback system leverages the report database, which contains both the final version created by the resident and the final report from the pathologists. A compiled report integrates links to whole slide images in the Paige viewer, allowing residents to efficiently review many cases from any location at their leisure.

These solutions have been used daily by faculty, residents, and staff, becoming well-integrated and part of the fabric of our operations. The resident feedback system was particularly praised as the most useful addition to the program for supporting education during the annual educational retreat. While digitally native pathology workflows provide significant efficiencies, they remain underexplored. It is essential for vendors to provide APIs that facilitate the creation of new digital workflows and for institutions to share their experiences to further develop this field.

Assessment of a prostate AI algorithm in a busy academic urological pathology diagnostic service

Simran Chandawarkar,[1] Anil Parwani[2], Swati Satturwar[2]
[1]Northeast Ohio Medical University, College of Medicine, Rootstown, OH, United States
[2]The Ohio State University Wexner Medical Center, Department of Pathology, Columbus, OH, United States
Email: schandawarkar@neomed.edu

Purpose

The purpose of this study is to evaluate the effectiveness of a commercial artificial intelligence (AI) program in diagnosing prostate cancer and varying presentations of it. We aim to assess the utility of commercial AI algorithms in daily pathological practice by comparing their utility to that of pathologists using a dataset collected over a substantial timeframe.

Study design and methods

This study will use data from the Department of Pathology archives at The Ohio State University Wexner Medical Center from January 2010 to May 2023 for patients with different subsets of prostate cancers. Digital images and other relevant information, including the final diagnosis made by a clinical pathologist, will be extracted and made anonymous with distinctive participant codes to ensure privacy and data security.

Implementation

A secure platform shall be used to upload Whole Slide Images (WSIs) and annotate them which will be vital in validating the AI algorithms. The AI diagnoses will then be compared to the clinical pathologist's diagnoses. The study's collected data will include the commercial AI's diagnostic accuracy compared to traditional pathological diagnoses made by clinicians. Secondary data will assess the time and cost savings of using AI processes.

Future applications and conclusion

Upon completion of the retrospective analysis, if the commercial AI is found to be as accurate or more accurate than a traditional pathologist's diagnosis, AI algorithms can be implemented in both academic and community medical centers to revolutionize the diagnosis of prostate cancer and potentially enhance accuracy, efficiency, and cost-effectiveness. This study aims to bridge the gap between AI innovation and clinical applications in pathology by thoroughly confirming commercial AI protocols and accuracy in prostate cancer diagnosis. This research will provide invaluable comprehension of the potential benefits and limitations of AI in Pathology. The anticipated outcome will not only demonstrate the level of feasibility of using commercial AI algorithms in routine pathology but also build a foundation for future innovations in medical diagnostics.

Digital pathology implementation in the Kaiser Southern California Permanente Medical Group

Alexander Nobori[1], Patricia Neder-Eckman[2], Brian Roehmholdt[1], Sophia Liang[3], Sony Wirio[1]
[1]Southern California Permanente Medical Group, Pathology, Los Angeles, CA, United States
[2]Kaiser Permanente Southern California Regional Reference Lab, Informatics, Chino Hills, CA, United States
[3]Kaiser Permanente Bernard J. Tyson School of Medicine, N/A, Seattle, WA, United States
Email: anobori@gmail.com

The Kaiser Southern California Permanente Medical Group (SCPMG) is a large physician partnership which includes over 8000 physicians (with over 100 pathologists across 13 pathology departments), 16 hospitals, and 4.8 million patients. In 2018, SCPMG Pathology initiated its digital pathology program with the initial goal of central immunohistochemistry interpretation with digital image analysis, which later expanded to primary case evaluation using digital slides.

To date, our institution has scanned 1,032,703 slides available to 111 pathologists within the medical system. Cases are SCPMG scans slides using 4 Leica Aperio AT2 scanners and 6 Leica Aperio GT450 scanners across three scanning sites. SCPMG also utilizes 2 Mikroscan SL5 telemicroscopy systems for remote evaluation of intraoperative frozen section procedures. In addition to primary diagnosis and frozen section diagnosis via telemicroscopy, digital pathology is also used for local, regional, and outside consultation; local and regional conferences; image analysis

for Ki-67; identification of tissue for next generation sequencing; and incorporation of de-identified whole slide images for teaching at the Kaiser Permanente Bernard J. Tyson School of Medicine.

A 2023 survey of Kaiser pathologists found that 82 % of the 112 respondents used digital pathology either weekly (50 %) or daily (32 %). 82 % of respondents reported they are comfortable performing primary sign-out with whole slide images without the use of glass slides in most cases. 72 % of respondents stated they are "very" or "somewhat" dependent on digital pathology to accomplish their daily work including remote sign-out with digital slides.

Herein we describe the history of the development of digital pathology within the Kaiser SCPMG system, the roadblocks faced and how they were overcome, and future directions for digital pathology within the partnership including computational pathology and AI solutions.

Measuring acceptance, readiness and influence of AI tools for Ki67 scoring among 90 pathologists

Amanda Dy[1], Ngoc-Nhu Jennifer Nguyen[2], Julien Meyer[1], Melanie Dawe[3], Anthony Fyles[3], Fei-Fei Liu[3], Dimitrios Androutsos[1], Susan Done[2], April Khademi[1]

[1]Toronto Metropolitan University, Electrical and Computer Engineering, Toronto, Canada

[2]University of Toronto, Department of Laboratory Medicine and Pathobiology, Toronto, Canada

[3]University Health Network, Princess Margaret Cancer Centre, Toronto, Canada

Email: amanda.dy@torontomu.ca

Background

The Ki67 proliferation index (PI) helps determine eligibility for adjuvant chemotherapy in early-stage, high-risk, HR + /HER2- breast cancer in Canada. Human assessment of PI, however, is time-consuming and error-prone. As pathology labs begin to integrate artificial intelligence (AI) tools, understanding pathologists' readiness and trust in AI becomes essential. This study investigates the degree of AI influence on pathologists' Ki67 scoring in invasive breast cancer.

Hypothesis

Pathologists are positively influenced by the AI tool and accept the AI recommendation.

Methods

We conducted a study involving the Ki67 PI scoring of 10 invasive breast carcinoma tissue microarrays (TMAs) by 90 blinded pathologists, with and without AI assistance. The AI tool for Ki67 assessment, which uses UV-Net architecture, was previously validated (PMID: 38218973). Data for training the AI differed from this study's data sources. No pathologists in this study contributed to the AI training or validation.

The AI tool demonstrated high accuracy when compared to gold standard counts, with a mean PI error rate of 0.6 % across the 10 TMAs. The tool provided a PI and a Ki67 +/− tumor nuclei overlay. The survey was prepared using Qualtrics (Provo, Utah, USA), with weblinks to the digitized TMAs and **an** integrated AI tool using PathcoreFlow (Toronto, Ontario, Canada).

To assess the impact of the AI tool on pathologists' Ki67 scoring, we define a novel AI influence metric (as a proxy for AI reliance) computed as follows:

$$\text{AI Influence} = 100^{*}(1 - |\text{Assisted Score} - \text{AI Score}|/ |\text{Non} - \text{Assisted Score} - \text{AI Score}|)$$

This measures how close the assisted score is to the AI recommendation compared to without assistance. The interpretations are 100 %: full influence (assisted score = AI score); 0 % to 100 %: partial influence (assisted score moved closer to the AI score); 0 %: no influence (assisted score = non-assisted score)**;** 0 % to −100 %**:** negative influence (assisted score diverged from the AI score more than the non-assisted score). Respondents were stratified by groups specified by years of experience (10-year intervals), career stage (trainee, practicing, retired), self-reported Ki67 scoring proficiency, utilization of the AI tool in the study (agree, disagree, neutral), and time to personal implementation of AI tools (<5 years, 5–10 years, not presently, never, not sure).

The Kruskal-Wallis test was used to evaluate statistical significance between groups due to the non-normal distribution of the data as shown by the Shapiro-Wilk test. All statistical analyses were set at $p < 0.05$ for significance. Metrics are shown as median [25th–75th percentile].

Results

In total, 900 responses (90 pathologists × 10 TMAs) were collected and analyzed. Pathologists demonstrated a strong reliance on the AI tool, with an AI influence measured at 89 % [72–95 %].

Years of experience and career stage significantly impacted AI influence ($p < 0.05$). Participants with 10–19, 30–39 and 40–49 years of experience were influenced the most by AI (92 [82–96]%, 92 [83–95]%, and 92 [84–95]%, respectively) whereas those with 20–29 years of experience had moderately less reliance (88 [75–95]%), and 0–9 years exhibited the least reliance (74 [50–91]%). In terms of career stages, retired pathologists demonstrated the highest AI influence of 91 [81–95]%, followed by practicing pathologists at 89 [76–95]%. Trainees were the least influenced by AI at 85 [53–91]%. Pathologists with greater experience were more likely to incorporate AI recommendations as the degree of AI influence increased proportionally to their years of experience and career progression.

Self-reported Ki67 scoring proficiency and views on the suitability of AI for Ki67 scoring did not significantly affect AI influence, demonstrating the impact of AI is independent of the pathologists' perceived personal proficiency or their views on AI suitability, suggesting that AI influence transcends subjective views.

Self-reported views on their implementation of AI tools significantly affected AI influence ($p < 0.05$). Interestingly, those who responded with "not presently" had the highest influence at 92 [83–95]%, followed by those who said 5–10 years with 89 [67–95]%. This paradox suggests that while some pathologists may have reservations about the current state of AI, their decisions are still influenced by the AI tool.

Self-reported use of the AI tool significantly affected AI influence ($p < 0.05$). The highest level of AI influence of 93 [86–96]% was demonstrated by the neutral group followed by those who agreed with an influence of 88 [69–95]%, and lastly, those who disagreed 61 [0–88]%. This demonstrates the AI tool substantially impacted scoring decisions, regardless of the participants' acknowledgement or declared usage of the tool. Even respondents **who** were indifferent or skeptical about using the AI recommendation were influenced by the tool. Most participants reported using AI when it was provided, only 7.7 % of trainees and 3.3 % of retired pathologists stated they did not use the AI tool.

Conclusion

In this study, we have defined a novel way of measuring AI reliance. We found that AI's influence on Ki67 scoring among pathologists varies with their experience and views on AI usability and implementation, with more experienced pathologists showing greater reliance. Pathologists showed a significant reliance on and trust in AI recommendations which demonstrates the importance of developing accurate and robust AI tools. The tool used in this study showed high accuracy when validated against ground truth manual counts. However, if the AI tool produces errors and pathologists rely on its recommendations, it could compromise quality of care. Understanding the trust and influence of AI on digital pathology is complex and needs to be studied further in other cohorts and problem areas.

Gram stain interpretation using artificial intelligence

Angel Strasburg[1], Jordan Starkey[1], Alex Donadio[1], Miguel Dimas[2], Michael Dowdle[2], Kami Kies[1], Audrey Schuetz[1]
[1]Mayo Clinic, Department of Laboratory Medicine and Pathology, Rochester, MN, United States
[2]Techcyte, Techcyte, Orem, Mexico
Email: strasburg.angela@mayo.edu

Introduction

Automation of clinical microbiology laboratory processes via artificial intelligence (AI) is a rapidly growing field. Gram stains are critical for patient care, providing rapid results that can guide initial clinical management. Here, performance of AI compared to traditional manual microscopy was evaluated for Gram stains.

Methods

The co-developed Mayo Clinic and Techcyte Bacteriology Gram Stain AI algorithm was paired with the Hamamatsu NanoZoomer S360 scanner to assess 530 smears from various specimen types (including biopsies/tissues, upper and lower respiratory, genitourinary, synovial fluid, swabs, and body fluids) submitted to the Mayo Clinic clinical microbiology laboratory from October 2023 to March 2024. Results were evaluated based on Gram stain reaction classification (Gram-positive or -negative), organism morphology (rod vs. coccus), and semi-quantitative values of organisms and human cellular elements (white blood cells, epithelial cells). Results from AI-assisted review were compared to manual reads; discrepancies were resolved using culture correlation and re-review of digital and manual smears.

Results

Morphologies missed by traditional manual microscopy were correctly identified in 70/530 (13 %) specimens using AI, compared to 2 % of specimens for which AI missed morphologies due to low prevalence of organisms. AI was superior to manual microscopy in subclassifying Gram-positive cocci into those resembling *Staphylococcus* and *Streptococcus* for 15 % of specimens. Small Gram-positive bacilli were occasionally difficult to report with AI due difficulties in distinguishing from streptococci. Human cellular elements were more often reported with AI for 18 % of specimens. Equivalent performance was noted for AI for Gram stain reaction and semi-quantitation of organisms.

Conclusions

Compared to manual reads, AI-assisted reading demonstrated a higher rate of detection for organism morphologies and human cellular elements. AI presents technologists with easily accessible images of organisms, detects organisms in rare quantities, and enhances reporting of cells otherwise missed by manual review. After appropriate training in digital images, technologists were pleased with images and results provided by AI, and efficiencies in workflow processes and ergonomics were achieved. Continued model develop is expected to demonstrate additional reporting improvements.

Automated detection of acid-fast bacilli in Kinyoun-stained digital slide images

Brendan O'Fallon[1], Paul English[2], Muir Morrison[2], David Ng[2], Elizabeth Enrico[2], Salika Shakir[2], Leo Lin[2], Katie Knight[2]
[1]ARUP Laboratories, Applied AI and Bioinformatics, Salt Lake City, UT, United States
[2]ARUP Labs, Research and Innovation, Salt Lake City, UT, United States
Email: brendan.ofallon@aruplab.com

Despite decades of research, *Mycobacterium tuberculosis* infection and disease (TB) remains a persistent global health challenge. Rapid and accurate identification of new TB infections is essential from both a public health and clinical management perspective. While nucleic-acid amplification tests can offer higher sensitivity, CDC guidelines recommend acid-fast bacilli (AFB) smear microscopy in all individuals suspected of having pulmonary TB (Lin, n.d.[1)], and manual review of auramine-O (AO) – stained smears remains a broadly employed and cost effective method for AFB detection, especially in resource poor communities. However, the sensitivity of smear analysis remains low compared to culture results, and in higher volume settings the manual review of slide smears may impede rapid reporting of results.

Here we investigate the use of modern deep-learning methods for detection of AFB-positive samples in digitized slide images. While digitization offers some advantages over traditional microscopy techniques, one challenge to AFB detection is that few modern scanners support fluorescent imaging. Hence, we examine the ability of deep-learning object detection tools to identify AFB in Kinyoun-stained smears, instead of the more common AO stain. Our goal is to develop a model that can accurately detect AFB-negative slides, reducing the manual review in the majority of cases.

Kinyoun-stained slides were prepared from decontaminated and concentrated samples of 202 specimens before culture. 1002 Whole Slide Image (WSI) scans were created from 3 different scanners to make up the dataset used in this research. Scanners include a Pramana HT2 unit, and Hamamatsu S20 and S360 models under various scanning profiles, including both 40× and 80× magnification. We extracted 450,000 tiles from each WSI and split tiles into training and validation sets at the specimen level, resulting in 362,805 training tiles and 89,012 validation tiles. The dataset includes approximately 11,000 human generated labels for AFB created by trained microbiology lab technologists. We train RCNN and FCOS object detection networks on bounding boxes labeling the extracted tiles to compare performance. Inference results are computed for positive labeled tiles and negative tiles to estimate sensitivity and specificity of the object detector at the bounding box and tile level. Inference results across full WSI for the entire dataset is run to estimate the density of organisms and AFB positive status at the WSI level where we have case results but may not have bounding box labels.

While the object detection model is trained to detect individual AFB organisms, our ultimate goal is to accurately predict which slides are negative for AFB. In our preliminary experiments, individual AFB detection models achieve precision, recall, and F1 scores in the range of 0.3–0.6. Despite this, by aggregating object-level predictions over an entire WSI and classifying the WSI based on these aggregated results, our model can make useful predictions at the WSI level. Our validation set consisted of 42 WSIs of which 22 were positive by culture, and 18 of these 22 were also positive by manual review of AO stain. Of these, our model correctly detected 17/18 of the AO positive cases while identifying 15/20 true negatives. These preliminary results suggest that we can accurately detect a substantial fraction of negative slides with little risk of mis-classifying a positive slide. With more data, and as we continue to refine our annotation process with our domain experts, we are optimistic our WSI-level classification can be improved further.

Deep learning quantification of cells in volumetric pseudo-H&E multiphoton images of renal biopsies

Charles Robbins[1], Gabriel Lerner[1], Kelsey Moore[2], Sudhir Perincheri[1], Richard Torres[2]
[1]Yale School of Medicine, Department of Pathology, New Haven, United States
[2]Applikate Technologies, Inc., Director of Market Development, Fairfield, CT, United States
Email: jack.robbins@yale.edu

Background

Nephropathologist assessment of renal biopsies clarify the underlying diagnosis or change patient management in up to 40 % of adult patients with acute or chronic renal injury. Quantification of mononuclear cell infiltrates and mesangial cell cellularity are tedious and normally imprecise tasks, but nonetheless are an important component of medical kidney biopsy assessment, including renal allograft pathology. Clearing Histology with Multiphoton Microscopy (CHiMP) (Applikate Technologies, Fairfield, CT) produces z-stacks of analogous pseudo-H&E whole slide images (WSIs) with benefits of markedly reduced labor, short turn-around time, compatibility with downstream processing with tissue preservation, more complete tissue assessment, and a digital output amenable to computerized analysis, potentially augmenting accuracy and precision of kidney diagnostics. However, the application of standard WSI deep learning approaches to multiphoton pseudo-colored data is largely unexplored. Given this, we developed deep learning algorithms for renal cell segmentation to create quantifiable features for the extent of lymphocyte infiltration and mesangial cell proliferation in renal biopsies to assist in renal pathology assessment on multiphoton microscopy images.

Methods

Portions of formalin-fixed core needle clinical renal biopsy samples were dehydrated, stained, and cleared over a period of 2 h using specialized reagents analogous to standard tissue processing (Applikate Technologies, Fairfield, CT). Samples were imaged on a CHiMP microscope to generate 10–15 digital, optical slices at 50-μm steps via a mathematical conversion of eosin and DAPI staining. 23,449 annotations of lymphocyte, tubule, endothelial, fibroblast, mesangial, parietal, and podocyte nuclei were labeled by renal pathologists from fully annotated fields of view of kidney multiphoton H&E WSIs. A DeepLabV3 + with ResNet-101 backbone was trained on patches of 512 × 512 pixels at 0.25 um/px (1940 patches with 70/15/15 train/validation/test split). Additional image postprocessing produced classified nuclei instance maps that can be overlayed onto kidney biopsies for visualization of predictions. QuPath image analysis workflows were used to create human interpretable features of lymphocyte infiltrates and mesangial cell counts per glomeruli.

Results

Prediction of kidney cell classification on the test dataset by our algorithm demonstrated F1-scores above 0.8 for all cell types, except parietal cells (F1-score = 0.78). Instance segmentation performance yielded a Dice score of 0.81. The proportion of impacted tissue area and cell density of lymphocytic infiltrates within pseudo-H&E multiphoton kidney biopsies are readily quantifiable within sections and across specimen volumes. Mesangial cell count per glomeruli was also extracted across the sample cohort and illustrates the enhanced precision achieved with z-stack datasets. Comparative data of these metrics throughout the 3D volume of the specimens demonstrate variable degree of section-to-section variance.

Conclusions

Digital multiphoton microscopy combined with deep learning-based classification of kidney cells enables robust quantification of relevant diagnostic criteria of renal allograft pathology – including extent of lymphocytic infiltrates, mesangial cell proliferation, and key identification of mononuclear cells within renal tissue compartments.

Automated whole-slide imaging capture quality assessment using convolutional neural network

Mathew Francis[1], Matthew Jones[2], Stefan Wagner[2], David Breen[2], Fernando Garcia[3], Mark Zarella[1]

[1]Mayo Clinic, Laboratory Medicine and Pathology, Rochester, MN, United States

[2]Drexel University, Computer Science, Philadelphia, PA, United States

[3]Reading Hospital, Department of Pathology and Laboratory Medicine, West Reading, PA, United States

Email: francis.mathew@mayo.edu

Whole-slide images (WSI) have revolutionized the field of pathology enabling multiple use cases including primary diagnosis. This transformation relies on two very important assumptions: the WSI is a true (free of artifacts) and complete representation (no missed or extra area) of the tissue present on the glass slide. The latter assumption is more important as the pathologist generally won't be able to verify the fact, as they are typically presented only with the scanned region. Even if the scanners capture a macro snapshot (MS) of the entire glass slide, it is not routinely presented to the pathologist. Modern scanners struggle with lightly stained images and adipose tissue, and oversensitive settings that could overcome this might result in extra area scanning. The situation is currently remedied by adding a manual quality assessment (QC) stage either before or after scanning to ensure that all the tissue is captured in the scan. However, this step is not practical in case of large capture volumes and is against the spirit of digitization and automation. Several open-source and commercial software packages exist that could be adopted for solving this problem, but they are mostly intended for data cleaning for artificial intelligence training rather than in-line QC. We propose a lightweight convolution neural network (CNN) trained on MS images to eliminate the need for manual QC. Importantly, the method does not directly require user annotations.

We analyzed a total of 954 WSI of hematoxylin and eosin (H&E) and estrogen receptor (ER) stained normal and breast cancer biopsies and resections, acquired in support of patient care at Drexel University College of Medicine. All the WSI were in the Aperio SVS format (Leica Biosystems Imaging Inc., Illinois, United States of America). Each WSI was from one subject and was captured, immediately digitized and then passed to manual QC over the time period from 2009 to 2018. The wide period was taken to incorporate real world variance introduced due to changes in laboratory protocols and tissue handling differences. The ground truth, pixel mask, for the CNN training was created using a publicly available tissue detector, PathProfiler, without retraining the model. The model was originally trained on prostate and colon tissue, appears to generalize well to breast tissue, but other tissue responses are uncertain. Ground truth mask labels for ER images were from the H&E images, used after co-registration using the extracted common hematoxylin channel. The mask was then down sampled (20 μm per pixel) and registered with the MS images. MS image patches of 64 × 64 pixels size was randomly sampled from the tissue region and non-tissue region guided by the mask such that minimum of 30 % and less than 5 % region was tissue in the patches, respectively. Total of 1024 patches were drawn for tissue and non-tissue classes each. CNN was composed of input layer, three set of convolutional layers, fully connected layer and an output layer. The CNN weights were optimized from random numbers using stochastic gradient descent at a learning rate of 10–3.

Among all the WSIs that were part of this study 122 (12.8 %) images were discarded due to unidentifiable scan region on macro images, as a result of scant tissue, or unusually small scan region. The dataset was selected as it contained long-term images of a single stain type from a single site. Therefore, tailoring an algorithm to this set would enable us to later test the data dependence of generalizing to a more diverse data set sourced from multiple laboratories, comprised of different tissues, and a variety of stains. Among all the H&E stained WSIs analyzed we found a range of 0 to 120 mm2 of missed tissue, with a median of 8.52 mm2. Likewise applying to the algorithm to ER stained images with an adjacent H&E stained image after removing any pair with co-registration issue, due to poor macro image quality, we found a comparable result, demonstrating that the approach can be applied to different stains without retraining.

Virtual immunohistochemistry for fast and accurate breast cancer biomarker analysis

Christopher Jackson[1], Kenneth To[1], Rafay Azhar[1], Digvijay Yadav[1], Louis Vaickus[2]
[1]ViewsML, Research and Development, Hershey, PA, United States
[2]Dartmouth Hitchcock Medical Center, Pathology and Laboratory Medicine, Lebanon, NH, United States
Email: cjackson@viewsml.com

Introduction

Estrogen receptor (ER), progesterone receptor (PR), and HER2 status are crucial biomarkers guiding breast cancer diagnosis, treatment decisions, and clinical trial eligibility. Traditional immunohistochemistry (IHC) for these markers is time-consuming, expensive, and prone to inter-laboratory variability due to subjective interpretation. This study presents a novel virtual immunohistochemistry (vIHC) software platform with significant potential to revolutionize breast cancer diagnostics by offering faster, more cost-effective, and potentially more standardized analysis, all from a single hematoxylin and eosin (H&E) slide.

Innovation

vIHC utilizes machine learning to analyze H&E stained breast cancer tissue sections. This approach predicts the presence or absence of ER, PR, and HER2 expression within individual cells from a single slide, eliminating the need for physical IHC staining. This has the potential to lead to faster turnaround times, reduced costs, and improved diagnostic consistency across laboratories and pathologists. By providing results as either positive or negative for each cell, vIHC minimizes the subjectivity inherent in visual assessment by pathologists, potentially leading to enhanced reproducibility. Notably, our area-under-the-curve (AUC) was evaluated by comparing its predictions on a per-cell basis to physical IHC results.

Methods

A large dataset comprising 8.1 million annotated cells for estrogen receptor (ER), 10.6 million annotated cells for progesterone receptor (PR), and 4.8 million annotated cells for HER2 was utilized. This dataset included tissue microarrays ($n = 15$) and tissue blocks ($n = 3$) from 564 patients, covering various breast cancer subtypes and normal tissue. The slides were initially stained with hematoxylin and eosin (H&E), scanned, then destained and restained with immunohistochemistry (IHC) markers before being re-scanned. A nuclear detection algorithm localized cells on the H&E images, and the corresponding IHC expression was used for categorization. Experienced pathologists set thresholds to classify IHC staining intensity. ER and PR expression was classified as positive or negative while HER2 was classified as either strong, moderate, weak, or negative. Machine learning algorithms were trained on these H&E images to predict per-cell biomarker expression. Model performance was evaluated using hold-out validation and testing sets, each comprising 10 % of the data.

Results

The vIHC models achieved AUC values of 0.885 and 0.902 for ER and PR, respectively, while HER2 model AUCs ranged from 0.738 to 0.948 for different staining categories.

Conclusion

Our vIHC software demonstrates promising potential for accurate and automated breast cancer biomarker analysis. This approach has the potential to significantly impact diagnostic workflows by offering faster, more cost-effective, and potentially more standardized analysis compared to traditional IHC. Further studies with data from diverse pathology labs are warranted to validate and address generalizability for broader clinical adoption. Our platform is well-suited to facilitate this data collection, and we plan to expand the number of biomarkers analyzed using this approach.

Keywords

Virtual immunohistochemistry, Breast cancer biomarkers, Machine learning in pathology, Digital pathology, Computational pathology, Diagnostic automation, Per-cell analysis, Biomarker standardization, Artificial intelligence.

Modeling the impact of recall bias on AI performance evaluation and validation

Sajda Adam[2], Mark Zarella[1]
[1]Mayo Clinic, Laboratory Medicine and Pathology, Rochester, United States
[2]Drexel University, College of Arts and Sciences, Philadelphia, PA, United States
Email: sa3754@drexel.edu

Intra-pathologist concordance is a fundamental measure to examine the potential impact of an intervention on the diagnostic outcome. For example, for a typical digital pathology validation, a pathologist may be asked to review the same slide digitally and again by microscope, and noninferiority will be based on whether they render the same diagnosis under both modalities. Similarly, when AI is introduced into the workflow, the pathologist may be asked to review the same case with and without AI assistance, sometimes measured against a gold standard, and the results may demonstrate potential improvement of AI in terms of accuracy, speed, or reproducibility. However, the influence of the study design must also be considered, as viewing the same slide twice, even when separated by some interval of time, can introduce bias due to recall. For example, when the subject consistently performs better or faster the second time on the same slide, it may be due to the intervention or it may simply be due to the subject having previously remembered the slide (recall bias).

Common approaches to reduce the impact of recall bias are:

- **Introduce a washout period.** By adding time between presentations, the likelihood of recall diminishes. However, as previously shown, even after doubling washout period from 2 weeks to 4 weeks nearly 30 % of slides may still be recalled by the subject. While this is likely to further diminish with longer washout periods, it is unlikely to reduce to 0 % and adding time potentially introduces extended project timelines.
- **Randomize the order of presentation.** An effective solution can be to interleave the intervention (e.g. AI trials) with the control (e.g. trials without AI). This can destroy the correlation between presentation order and the intervention thus reducing the impact of recall bias on the results, but can also introduce complications to the study. For example, if a pathologist is asked to switch back and forth between digital and glass in the same session or needs to remember which tools or procedures to use for a given case, it can introduce artifice (and maybe aggravation). Furthermore, in concordance studies where matching diagnoses under both conditions is the primary metric, the order of presentation is of little consequence because recall may be producing higher-than-typical concordance regardless of presentation order.
- **Introduce a 3rd trial.** Another option can be to present the slide to the subject a third time to explicitly measure the impact of recall. For example, this could mean asking the subject to view without AI assistance, then with, then without again to provide a baseline for recall bias and confirm that improvement declined on the third presentation of the same slide. This helps disentangle the effects of recall from the true objective of the study. One of the drawbacks to this approach is that more sessions are needed from the subject.
- **Rely on inter-pathologist review.** The study can be designed so that only inter-pathologist review of slides is needed. For example, two or

more pathologists will each review the same slide, but with different conditions or interventions. With sufficient number of combinations, each pathologist's performance can be measured against their population results as a whole to surmise the effect of the intervention. However, concordance even without the intervention can be relatively low between subjects, so this must be considered before embarking on this design.

We developed a computational model capable of simulating the influence of recall bias in a validation study. We found that as recall becomes more prevalent concordance naturally becomes inflated, which can effectively overestimate concordance and provide significant risk to the interpretation of validation results. We found that the magnitude of this effect depended on a number of additional factors including inter- and intra-pathologist concordance (without the intervention) for a given task, strategic and cognitive biases, and imbalanced data sets. We then tested how each of the study design modifications described above potentially mitigates recall bias, enabling us to provide example tasks where each design may be most effective. Together, these results emphasize the importance of study design in any validation and provide a reference frame from which to better interpret studies in digital pathology and AI.

Improving PR-ER-HER2 detection in triplex IHC using a multi-label UNet and adversarial training

Xingwei Wang[1], Kaining Mao,[2] Matthew Olson,[1] Mike Flores,[1] Mehrdad Gangeh,[1] Yao Nie,[1] Jim Martin,[1] Patricia (Trish) Thorne-Nuzzo,[1] Terry Landowski,[1] Erika Walker,[1] Danielle Brands,[1] Paco Delgado,[1] Gianni Ferreri[1]

[1]Roche, Computational Science Pathology Lab, santa clara, CA, United States

[2]University of Alberta, Electrical Computer Engineering, Edmonton, AB, Canada

Email: xingwei.wang@roche.com

Introduction

Estrogen receptor (ER), progesterone receptor (PR) proteins, and human epidermal growth factor receptor 2 (HER2) are important biomarkers for breast cancer treatment and diagnosis. Triplex immunohistochemistry (IHC) staining can simultaneously detect three biomarkers and their co-expression, but it is difficult for humans to score reliably. In this study, we developed a Multi-label UNet Pipeline to detect ER, PR, and HER2 in triplex images, which allows pathologists to score triplex images and aid breast cancer diagnosis and treatment. Cell classification models can be sensitive to image noise, color, and intensity changes from different staining protocols, tissue types, and scanners. To solve this problem, we proposed a novel approach based on customized augmentation and adversarial training. Through adversarial training and customized augmentation, the model is less sensitive to noise and image intensity and color variations caused by different scanners and staining differences.

Methods

Using chromogenic multiplexing technology, breast carcinoma tissue was stained on a BenchMark ULTRA staining system. PR (clone 1E2) was detected with Tyr-TAMRA (Magenta) chromogen, while ER (clone SP1) was detected using QM-Dabsyl (yellow) chromogen. Her2(Cal27) was detected using DISCOVERY Green HRP Kit. 50 images were selected from six triplex slides with a range of stain intensity. A color deconvolution algorithm was developed to unmix 50 triplex images to generate synthetic Dabsyl ER, TAMRA PR, and green HER2 images. Three pre-trained models were used to detect nuclei and evaluate ER/PR as positive or negative. HER2 was scored 0, 1+, 2+, and 3+. Two pathologists reviewed and corrected wrongly labeled cells. To avoid a single multi-classification model identifying 17 different classes/phenotypes in triplex images with complex training on an unbalanced dataset, or training three multi-classification models that need to merge phenotypes, a multi-label UNet algorithm with customized augmentation was developed to detect positive/negative ER and PR markers, HER2 with 1+, 2+, 3+, and negative, and their corresponding co-localization in triplex images. An adversarial training process was implemented to generate perturbed pathology images to improve the robustness of the multi-label model.

Results

In this study, the ground truth agreed by two pathologists contained 69,571 cells; out of these, 22,129 cells were ER+, 20,714 were ER-, 29,062 were PR+, 13,781 were PR-, and 2841 were HER2 3+, 5618 cells were HER2 2+, 19,837 as HER2 2+, and 17,618 as HER2-. The remaining were other cells. We split the original dataset for training and testing: 80 % for training and 20 % for testing. The learning rate was set to 1e-4 and the Adam optimizer was used. The accuracy for detecting positive and negative ER/PR markers was 0.89 and 0.93 respectively. For HER2, the accuracy was 0.80. The overall accuracy for all phenotypes was 0.86. Through customized augmentation and adversarial training, the multi-label UNet model becomes less sensitive to noise and image intensity variations, and its validation and training losses are smaller and converge faster than without this process.

Conclusion

An innovative Multi-label UNet algorithm was developed, which can detect PR-HER2-ER in a triplex image and help the pathologists score and eliminate the need for merging different types of biomarker locations predicted by a multi-classification model or reduce the complexity of training multiple models. The multi-label UNet model is more robust to diverse triplex images with color and intensity variations through customized augmentation and adversarial training.

Utility of artificial intelligence algorithms for urologic pathology

Swati Satturwar[1], Anil Parwani[2]

[1]The Ohio State University, Wexner Medical Center, Pathology, Columbus, United States

[2]The Ohio State University, Pathology, Columbus, OH, United States

Email: swati.satturwar@osumc.edu

In this era of precision medicine, artificial intelligence (AI) based diagnostics in urologic pathology is making its way for deployment to routine clinical use. In recent years there have been tremendous development and advancement in the research studies analyzing use of AI tools for urologic pathology particularly for prostate cancer. AI applications have been used not only in histologic diagnosis of urologic cancers but also used at every step of urologic cancer detection and management that includes patient education, identification of cancer (radiologic and pathologic), classification and grading of cancer, devising treatment plans and predicting treatment response and prognosis. With an ever-increasing body of literature about AI in urologic pathology, it is evident that AI has the potential to outperform pathologists in many aspects and has the potential to upgrade the conventional methods of urologic cancer detection, subtyping and grading of urologic cancers using whole slide images (WSI). AI has the potential to reduce overall workload for pathologists and assist general pathologists in the community hospital setting or developing countries to achieve subspecialized diagnostic capabilities and reduce healthcare disparities. A handful of studies are exploring potential use of AI tools beyond diagnosis, for predicting outcomes based on single most representative WSI. AI is set to improve the accuracy, the speed, the reliability, cost-effectiveness and decrease diagnostic disparities for urologic pathology and eventually improve patient care. There are limited studies on implementation of these AI tools in routine clinical use. However, several questions regarding the actual role and utility of AI for routine diagnostic use in urologic pathology

remains to be answered. There is limited knowledge about the performance of such AI tools for benign mimickers of urologic carcinomas and carcinomas with variant morphologies.

Despite of these challenges there are several prospects for AI use for urologic malignancies as outlined below.

- As pre-sign-out screening tool
- As a quality control tool for post-sign-out cases
- Real-time pathologist digital diagnostic assist tool
- For second review of challenging cases
- Routine primary diagnosis with pathologist supervision
- Routine primary diagnosis of negative prostate biopsy cases without pathologist supervision
- Deploying automated templates for reporting of prostate cancer CNBs
- Research aspect of AI as a tool to improve the understanding of biology of prostate cancers and beyond
- Prostate cancer risk stratification using WSIs
- Diagnosis, subtyping, and grading of renal cancers
- Quantification of tumor necrosis of clear cell renal cell carcinoma resection cases
- Risk stratification based on morphologic features
- Diagnosis and subtyping of bladder cancers
- Automated urine cytology tools for diagnosis of bladder cancer for routine clinical use
- Risk stratification of bladder cancer patients using morphologic features
- Diagnosis, subtyping, and quantification of different components of mixed germ cell tumors
- Improved Ki-67 quantification and risk stratification
- Discovery of new knowledge by analyzing large scale histopathologic data
- Automated quantification of prognostic biomarkers e.g., PD-L1 IHC
- Combining pathology with radiomics and genomics to provide highest level of integrated diagnosis for urologic cancers.

This presentation will provide our institutional experience as well as provide an update on the current usage and barriers for AI algorithm deployment for urologic pathology diagnostics.

Integrating AI model for tumor microenvironment analysis & spatial distribution in colorectal cancer

Nada Shaker[1], Noor Shaker,[2] Mohammed Shaker,[2] Anil Parwani,[1] Nuha Shaker[3]

[1]University of California, San Francisco, CA, United States

[2]Spatial X, NA, London, United Kingdom

[3]University of Pittsburgh Medical Center, Department of Pathology, Pittsburgh, United States

Email: Nada.shaker@ucsf.edu

We employ advanced AI models to identify the tumor bed and we use AI segmentation algorithms for eosinophils, plasma cells, and lymphocyte predictions, aiming to gain a deeper understanding of the tumor microenvironment (TME). Our methodology was tested on eight colorectal cancer slides sourced from The Cancer Genome Atlas (TCGA). The selected cases involved patients diagnosed between one and two years prior; four of these patients survived beyond one year, while the rest succumbed to the disease within a year.

We conducted a comparative analysis of biomarker counts within the tumor bed, generating heatmaps for visual inspection of individual biomarkers at the slide level. Using AI, the system was able to segment thousands of cells (ranging from 886 for one patient to 27,880 for the patient with most segmented cells). Cell distribution also varied among patients and cell type. Eosinophils distribution, for example, ranges from 0.06 Eosinophils/mm2 for one patient to 0.07 cells/mm2 for another, while plasma cell distribution was between 0.03 cells/mm2 to 0.15 plasma cells/mm2. The results revealed how the use of AI allow cell segmentation and counting at scale, revealing clear differences in both the quantity and spatial distribution of these biomarkers among patients and among different biomarkers. This study suggests that analyzing the tumor microenvironment on the level of individual biomarkers and for combination of biomarkers, both in terms of counts and spatial distribution has the potential to reveal hidden insights about individual tumor characteristics. The study underscores the significance of further TME analysis and demonstrates the potential of AI in extracting valuable biomarkers.

Future work will expand the use of these AI tools, scaling the study to include more patients to derive conclusive results. The ultimate goal is to enhance colorectal cancer diagnostics and identify predictive biomarkers for prognostic endpoints. This expanded research could significantly impact clinical practices and patient outcomes, providing a robust framework for integrating AI into cancer diagnosis and prognosis.

Clinical and analytical validation of an AI tumor content scoring algorithm for lung cancer

Hallie Rane, William Edelman, Rebecca Wyatt, Donna Mulkern, Hannah Moore, Peter Caie

Indica Labs, Clinical Applications, Albuquerque, United States

Email: hrane@indicalab.com

Tumor content scoring of a patient's sample prior to molecular testing is an essential part of the clinical workflow. Pathologists assess tumor content to determine if there is sufficient tumor tissue for a successful molecular test. It is a time-consuming task for the pathologist and one which suffers from inter- and intra-observer variability.

We developed an AI-based algorithm, Lung Macrodissect AI, to automatically detect areas of high tumor density and for the quantification of tumor content in digitized Hematoxylin and Eosin (H&E) stained non-small cell lung cancer tissue sections. The aim of deploying such an algorithm in the clinic is to improve turnaround times, standardize the quantification of tumor content, and reduce the number of samples with insufficient tumor sent for molecular testing.

The validation of the algorithm was twofold. Firstly, we performed clinical validation of tumor content across whole slide images (WSI) and next we analytically validated the algorithm at the cell-object level.

The clinical validation cohort consisted of 250 WSIs sourced from an institute whose data was not used in developing the algorithm. This external validation cohort was sent to 5 pathologists across independent and unaffiliated clinical institutes. For clinical validation, each pathologist annotated 1/5th of the cases for a region typically representative for macrodissection prior to molecular testing. Each pathologist scored all 250 cases for tumor content, within these annotated regions, with and without the aid of Lung Macrodissect AI, and with a 4-week wash out period between scoring.

Analytical validation consisted of a ground truth from the mode of the 5 pathologists who each annotated 9374 cells across 22 WSIs. First, we evaluated the algorithms performance at cell detection, and next its accuracy at classifying the cells as either ‘cancer’ or ‘other’.

An ICC consensus value of 0.43 was calculated for tumor content when comparing the 5 pathologist's scores across the 250 WSIs without the assistance of the algorithm. The ICC increased significantly to 0.74 when the pathologists re-scored with the algorithm's assistance. When comparing the average of the pathologists' assisted scores to the algorithm's scores, the ICC rose to 0.94. At a 20 % binary clinical cut-off for tumor content, a Fleiss Kappa interrater agreement metric of 0.23 was calculated when comparing the unassisted pathologist scores, which increased to 0.53, with a 92.8 % agreement across all pathologists, when the pathologists were assisted by the algorithm. The Kappa value for below or above 20 % tumor content further increased to 0.91, when comparing the average assisted pathologist scores to the algorithm's.

The analytical validation for cell detection was excellent across the 9374 cells with a Precision of 1.0 and an F-1 score of 0.99. Cell classification validation, as either ‘cancer’ or ‘other’ cell, also achieved excellent

agreement to the mode of the pathologists with a Precision and F-1 scores of 0.94.

Considering the significant increase in agreement between pathologists when assisted by Lung Macrodissect AI, the algorithm has the potential to reduce subjectivity in percent tumor scoring in the clinical workflow. Further potential advantages of clinically deploying the algorithm could be reducing the time spent by pathologists undertaking this task and reducing the amount of rejected molecular tests in the clinic. This study strongly suggests that AI-assisted scoring can improve interrater variability tumor content metrics and demonstrates the accuracy of the algorithm to classify tumor cells in lung H&E WSIs. Macrodissection AI may help boost clinical standardization for molecular testing through automated image analysis, AI-assisted tissue macrodissection region selection, and tumor content scoring.

The mighty mini: Small scanners in advancing digital pathology

Shaoli Sun[1], Lokman Cevik,[2] Scott Hammond,[2] David Kellough,[2] Erin Palermini,[2] Anil Parwani,[2] Wendy Frankel,[2] Jennifer Vazzano,[2] Abberly Lott Limbach,[2] Zaibo Li[2], Konstantin Shilo,[2] Sarah Reuss,[2] Swati Satturwar,[2] Mohammad Shokouh-Amiri,[2] Jesse Sheldon[2]

[1]The Ohio State University, Pathology, Columbus, OH, United States

[2]Osu, Pathology, Columbus, OH, United States

Email: shaoli.sun@osumc.edu

As the adoption of large, multi-slide scanners becomes increasingly common in hospitals for routine daily works, certain areas of digital pathology, such as frozen section evaluation, stat case review, and remote consultation, appear to have been overlooked. This is primarily because these larger scanners are not designed to handle wet slides, which are essential for immediate pathology diagnosis. To address the issues, we have been investigating the potential usage of small compact scanners in these areas.

Since the implementation of digital sign-out in 2019, most of our pathologists have become proficient in pathology diagnosis with digital slides for their daily work. The skills and confidence acquired over the past few years have prepared them to explore the use of digital imaging for frozen section analysis.

We selected the Grundium 20× single slide scanner for its rapid scanning time and browser-based operating platform. We scanned 60 routine frozen section slides, including liver and kidney donor tissue evaluation for transplant, fungal sinusitis, and various tumor and non-tumor cases. Six pathologists participated in the study. The average scanning time was 2.87 min, and the average evaluation time was 2.4 min. The image quality was rated at 3.5 out of 5, with 5 being excellent. The diagnostic accuracy was 95 %.

While the additional scanning time did indeed increase the turnaround time for frozen section diagnosis, it also enabled the on-call pathologist to digitally share slides instantly with other pathologists from different subspecialties. This not only improved diagnostic accuracy for optimal patient care but also reduced the on-duty pathologist's stress associated with the demanding frozen section evaluation.

The browser-based compact scanner we used proved to be ideal for low-volume and immediate needs, without the need for complex integration with the laboratory information system (LIS). It handled wet slides effectively and produced excellent image quality.

Encouraged by the positive experience with frozen section diagnosis, we have also begun using it for stat cases, which often occur during weekends or late in the day when large scanner personnel are not available.

Finally, pathologists at our community hospitals are now planning to use compact scanners to scan their challenging cases and directly consult with subspecialty services at the main hospital. In the past, these cases typically took days to arrive and occasionally got lost in transit which significantly delayed patients' diagnosis.

In conclusion, compact scanners are becoming increasingly popular in both large and small hospitals due to their low cost, ease of operation, and versatility in various scenarios.

A comparative study of augmented reality microscopy and quantitative image analysis algorithms

Jang Cho, Brian Smola, Sandra Camelo-Piragua, Laura Lamps, Mustafa Yousif

University of Michigan, Pathology, University of Michigan, Ann Arbor, MI, United States

Email: jangcho1991@gmail.com

Background

Tumor cell malignancy is assessed by evaluating histopathological differentiation and cell proliferation. The introduction of Ki-67 immunohistochemistry in the work-up of multiple neoplasms has opened a new approach for their diagnosis and prognostic evaluation. Ki-67, a marker indicating cellular proliferation, is integral for grading and classifying various tumors, including breast cancer, gastrointestinal (GI) neuroendocrine tumors, pituitary adenomas, and meningiomas. Traditionally, different counting methods have been employed to measure the Ki-67 index. Recently, augmented reality microscopy (ARM) has emerged, enabling real-time image analysis using glass slides. This study aims to compare traditional quantitative image analysis (QIA) methods using whole slide imaging (WSI) for Ki-67 scoring in tissue histology material with the newer ARM techniques.

Methods

Ki-67 immunohistochemical slides from 80 cases, consisting of GI neuroendocrine tumors, pituitary adenoma, meningioma, and breast carcinoma of varying grades (20 each), were analyzed. The Ki-67 index was quantified in up to three hot spots using ARM with live image analysis and WSI using the field of view (FOV) method. Intraclass correlation coefficient was calculated using two-factor ANOVA. For comparison analysis, two-tailed *t*-test was performed at the significance level of 0.05.

Results

Overall, no significant statistical difference was observed between QIA using WSI and ARM in Ki-67 detection in each tumor category. All four categories of tumor showed very high intraclass correlation coefficients (ICC): 0.99 in GI neuroendocrine tumor, 0.99 in pituitary adenoma, 0.97 in meningioma, and 0.94 in breast cancer. In GI neuroendocrine tumor, ARM demonstrated 1.26 % lower ki-67 index than WSI ($p = 0.91$). In pituitary adenoma, ARM demonstrated 0.2 % lower ki-67 index than WSI ($p = 0.8$). In meningioma, ARM demonstrated 0.97 % higher ki-67 index than WSI ($p = 0.71$). In breast carcinoma, ARM demonstrated 2.7 % higher ki-67 index than WSI ($p = 0.51$). In summary, pituitary adenomas showed the least difference and variation between the two methods (standard deviation 0.35), whereas breast tumors exhibited the greatest difference and variation (standard deviation 3.6). Meningioma results fell between these two extremes.

Conclusions

ARM streamlines and accelerates Ki-67 analysis for pathologists by overlaying image analysis on glass slides in real-time, maintaining accuracy. However, analyzing entire tissue tumors, such as in breast tumors, remains time-consuming and necessitates further optimization to improve efficiency.

Predicting TME from H&E images to characterize immunotherapy responses in lung cancer patients

Tamara Jamaspishvili[1], Sushant Patkar,[2] Alex Chen,[2] Alina Basnet,[1] Amber Bixby,[1] Rahul Rajendran,[1] Rachel Chernet,[1] Susan Faso,[1] Prashanth Ashok Kumar,[1] Devashish Desai,[1] Ola El-Zammar,[1] Christopher Curtis,[1] Sam Carello,[1] Michel Nasr,[1] Peter Choyke,[2] Stephanie Harmon,[2] Baris Turkbey[2]

[1]SUNY Upstate Medical University, Pathology, Syracuse, NY, United States
[2]National Cancer Institute/NIH, Artificial Intelligence Resource (AIR), Bethesda, MD, United States
Email: jamaspit@upstate.edu

Immune checkpoint inhibitors (ICI) have become integral to non-small cell lung cancer treatment. However, reliable biomarkers predictive of immunotherapy efficacy are limited. This work introduces HistoTME, a novel weakly supervised deep learning approach to characterize tumor microenvironment (TME) composition from routinely collected histopathology slides. In contrast to previously published computational pathology methods, which rely on detailed pixel-level annotations from expert pathologists, HistoTME harnesses recently developed digital pathology Foundation Models (FM) to robustly predict the expression levels of 30 distinct cell type-specific signatures in an end-to-end fashion from (hematoxylin-eosin) H&E slides. HistoTME is trained using matched whole slide H&E and bulk transcriptomics data of 865 NSCLC patients from The Cancer Genome Atlas (TCGA), validated using whole slide H&E and bulk transcriptomic data from 333 NSCLC patients from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) and independently tested on whole slide H&E and multiplex Immunohistochemistry (IHC) data from 82 NSCLC patients that had complete surgical resection at SUNY Upstate Medical University.

We further illustrate HistoTME's ability to identify responders to treatment with immune checkpoint inhibitors (ICI), particularly among patients with low PD-L1 expression utilizing a comprehensive clinically annotated external validation cohort of 652 lung cancer patients. Finally, we demonstrate how HistoTME can be utilized to analyze interactions among various TME components and demonstrate its effectiveness in predicting ICI responses, yielding an AUROC of 0.75 [95 % CI: 0.61–0.88] at first-line ICI treatment.

In summary, HistoTME allows for a multimodal characterization of the TME and identification of ICI responders without the need for expensive molecular tests or additional tissue stains. We believe that this cost-effective method for predicting immune responses will accelerate the discovery of novel biomarkers of treatment response and improve the prognostication and management of patients undergoing ICI treatment.

Automated AI solution for HER2 interpretation in breast cancer: A multi-site reader study

Pranil Chandra[1], Elisabeth Shearon,[2] Piotr Borkowski,[3] Rita Canas-Marques,[4] Stuart J. Schnitt,[5] Emilie Morphew,[2] Marigny Roberts,[1] Justin Cates,[1] Rawia Yassin,[3] Gaetan MacGrogan,[6] Laurent Arnould[7], Justin Poling,[1] Raz Ziv[8], Sevde Etoz[8], Maya Grinwald[8], Giuseppe Mallel[8], Manuela Vecsler[8]
[1]Pathgroup, NA, Brentwood, TN, United States
[2]Alverno Laboratories, NA, Hammond, IN, United States
[3]Quest Diagnostics, NA, TAMPA, FL, United States
[4]Champalimaud Foundation, Pathology, Lisbon, Portugal
[5]Dana-Farber Brigham Cancer Center, Pathology, Boston, MA, United States
[6]Institute Bergonié, Pathology, Bordeaux, France
[7]Center Georges-Francois Leclerc, Pathology, Dijon, France
[8]Ibex Medical Analytics, R&D, Tel Aviv, Israel
Email: pchandra@pathgroup.com

Introduction

Precision and standardization in HER2 immunohistochemical (IHC) scoring on whole slide images (WSIs) have gained tremendous importance with the availability of new HER2-targeted therapies, expanding the population of patients who can benefit from HER2-targeted therapy. This multi-arm, multi-reader study investigated the clinical utility of a HER2 scoring AI support solution for pathologists in 3 reference laboratories in the US.

Methods

The study included 558 patients who had a breast biopsy or excision between 2020 and 2023 from 3 different reference laboratories, a roughly equal number of cases from each site. Cases were selected based on sign-out HER2 scores, anonymized, and evaluated in the 2-arm with a full cross-over design (Arm A and B). Two local reader pathologists per site evaluated HER2 slides digitally without (Arm A) and with the support of the AI solution (Arm B) (6 reader pathologists in total). In parallel, three ground truth (GT) pathologists per site (8 GT pathologists), experienced in breast pathology, reviewed the same slides digitally, blinded to AI results and patient data. The ground truth HER2 score for each case was established based on the majority score (2/3 agreement of GT the breast experts). At least one GT pathologist was external to the laboratory to prevent bias.

The utility of the AI solution was evaluated based on the accuracy compared to GT and the inter-observer agreement in the study arms A and B, which were evaluated at the clinical cut-offs and all HER2 scores defined in the ASCO/CAP guidelines.

The number of discrepancies between the FISH result (positive/negative) and the reader HER2 scores in each arm was evaluated to determine the accuracy gains with the AI-supported scoring.

Major discrepancies were defined as follows: for a FISH-negative case, a reader gave 3+; for a FISH-positive case, a reader gave 0 or 1 + .

Results

The reader pathologists supported by AI showed significant improvement in overall HER2 score accuracy, showing 78.8 % [76.2 %—81.2 %] and 83.9 % [81.5 %—86.0 %] accuracy, for Arm A and B, respectively at all scores (0, 1+, 2+, 3+) ($p < 0.05$).

The inter-observer agreement significantly increased in Arm B (92.5 %) compared to Arm A (75.8 %) ($p < 0.001$) at all scores. Inter-observer agreement between the GT pathologists was 74 % at all HER2 scores combined.

The HER2 AI standalone accuracy vs. GT was 81.4 % [77.8 % - 84.6 %], 89.2 % [86.3 % - 91.7 %], and 92.4 % [89.8 % - 94.5 %] for all HER2 scores (0, 1+. 2+. 3+), 0 vs. 1+/2+/3+ and 0/1+ vs. 2+/3+ cut-offs, respectively.

Readers scored 10 cases 0 or 1+ in Arm A and B that were FISH positive. In Arm A, 11 cases were scored as 3+ with a negative FISH test. However, readers gave fewer 3+ scores to FISH-negative cases in Arm B ($n = 2$).

Various examples (borderline case scores with and without AI support) will be presented.

Conclusion

With the support of the AI solution for HER2 IHC scoring, pathologists significantly improved HER2 score accuracy and inter-reader agreement in breast cancer specimens. AI-supported reading also showed higher preliminary similarities with the FISH results; however, a more systematic approach is necessary to prove the concordance. HER2 AI solutions can pave the way for more accurate treatment selection for patients with invasive breast cancer and a more standardized approach to patient care.

Detection of stitching artifacts in whole slide images using advanced image processing techniques

Randall simonson
Mayo clinic, Computational pathology and AI (CPAI), Rochester, MN, United States
Email: simonson.randall@mayo.edu

This study introduces a novel and systematic approach to detect stitching artifacts in whole slide images (WSIs). The methodology employs advanced image manipulation, edge detection, and linear blur techniques

to enhance artifact visibility and detection accuracy. To rigorously evaluate the approach, an artificial artifact generation algorithm was developed which relocates sections within the image, simulating potential artifact regions. Two types of artificial artifact generation were investigated: pseudo-random chunk displacement and the generation of lawnmower lines. These manipulations facilitate a targeted analysis of diverse image segments. To detect the generated stitching a linear blur operation using custom kernels, and then use the Sobel operator for edge detection were applied. Thresholding the gradient images isolates prominent edges, and Hough Line Transform were applied to identify horizontal and vertical lines indicative of stitching artifacts. These results demonstrate the effectiveness of the proposed method in detecting stitching artifacts, particularly in non-homogeneous regions containing tissue structures. By integrating advanced image processing techniques, this approach provides a fast and data-efficient means to produce positively correlated signals for artifact detection, thereby ensuring improved image quality and reliability in WSIs. This research establishes a foundation for automated and precise artifact detection systems, contributing to enhanced digital pathology workflows and diagnostic accuracy.

Deep learning for diagnosis of classical Hodgkin lymphoma versus anaplastic large cell lymphoma

Daniel Rivera, Kristine Ali, Rongzhen Zhang, Brenda Mai, Hanadi El Achi, Jacob Armstrong, Amer Wahed, Nghia Nguyen

The University of Texas Health Science Center at Houston, Pathology and Laboratory Medicine, Houston, TX, United States

Email: daniel.e.riveradelgado@uth.tmc.edu

Introduction

Recent studies have shown promising results in using Deep Learning to detect malignancy in whole slide imaging. However, there is an unmet need for its applicability in clinical practice. We attempted to use deep learning with a convolutional neural network (CNN) algorithm to build a lymphoma diagnostic model aimed at distinguishing morphologically challenging cases such as classical Hodgkin lymphoma (CHL) and anaplastic large-cell lymphoma (ALCL). This tool could be practical in-service work to optimize diagnostic workups and prevent the risk of misdiagnosis.

Materials and method

Our software was written in Python language. We obtained digital whole-slide images (WSI) of hematoxylin and eosin-stained (H&E) slides of 20 cases, which included 10 cases for each diagnostic category. From each WSI, 60 image patches sized 100 × 100 pixels at 20× magnification, yielding 1200 image patches; of these, 1079 (90 %) were used for training, 108 (9 %) for validation, and 120 (10 %) for testing. For each test set of 5 images, the predicted diagnosis was combined with the prediction of five images, ensuring the robustness of our results.

Results

The distinction between CHL and ALCL using our deep learning model showed excellent diagnostic accuracy (100 %) for image-by-image prediction and 100 % accuracy for set-by-set prediction.

Conclusion

Our preliminary results provide proof of concept for incorporating this automated diagnostic tool for screening challenging lymphoma cases. They highlighted its potential to be integrated into future pathology workflows to augment pathologists' precision. Our model's high accuracy in distinguishing cell morphology between CHL and ALCL is a promising step towards improving diagnostic accuracy. Further data gathering and testing in our model will continue bringing new insights to support its broader applicability, potentially revolutionizing the field of pathology.

Prospects of a deep learning algorithm for prostate cancer diagnosis in a community hospital setting

William MacDonald[1], Himani Kumar,[1] Manisha Mishra,[1] Sarah Reuss,[1] Denise Gamble,[1] Andrea Parke,[2] Anil Parwani,[1] Swati Satturwar[1]

[1]The Ohio State University Wexner Medical Center, Department of Pathology, Columbus, OH, United States

[2]Paige AI, Clinical Science and Partnerships, New York, United States

Email: william.macdonald@osumc.edu

Background

A major challenge for deployment of artificial intelligence (AI) tools for routine clinical use include lack of real-world data supporting the clinical utility. This retrospective study was undertaken to evaluate the performance of an AI algorithm for prostate cancer diagnosis and grading (Paige Prostate) in a community hospital setting, and assess clinical utility with regards to reading efficiency, rate of deferrals and diagnostic confidence level.

Design

The study cohort comprised of 50 prostate biopsy cases to encompass 40 conventional prostatic acinar adenocarcinoma cases of different Grade groups (1–5) and 10 benign cases as per original pathology report signed out by a uropathologist. All slides were digitized using Leica Aperio AT2 scanner at 40× magnification and then processed by the AI solution. Ground truth was determined by an expert uropathologist. Three community pathologists evaluated the cases with and without AI assistance in a randomized fashion with a minimum washout period of 3 weeks between read modalities. Slide level outcomes, reading time, 5-point Likert scale for diagnostic confidence (High confidence, Confidence, Some confidence, Little confidence, No confidence) and deferrals were reported using a custom electronic case report form. Concordance was calculated for diagnosis and grading.

Results

With AI assistance, 94.0 % of cases were reported as either tumor or benign, compared to 87.2 % without AI assistance, with remaining cases being deferred. Reasons for deferral included need for immunohistochemistry (9/9 with AI assistance and 18/19 without AI assistance) and/or additional levels (1/19 without AI assistance). Pathologist binary outcomes compared to ground truth reported sensitivity and specificity of 0.99 and 0.84 for AI assistance and 0.98 and 0.85 without AI assistance respectively. Prostate AI standalone performance was calculated as 1.00 sensitivity and 0.82 specificity compared to ground truth. With AI assistance, a mean percentage reduction in reading time of 11.8 % for all 3 community pathologists was observed. Diagnostic confidence for reported cases demonstrated an 8.9 % increase in High confidence/Confidence responses with AI assistance (94.3 % vs 85.4 %). Table 1 shows overall concordance of 3 pathologists with and without AI assistance. Of note, AI showed improved agreement for benign cases from 17/19 to 26/27 with AI assistance.

Conclusion

This study provides real-world clinical utility evidence in terms of positive time savings, deferral rate and level of confidence with use of AI assistance for prostate cancer diagnosis and grading in a community hospital setting. In addition, our study indicates diagnostic workflow cost savings due to decreased requests for immunohistochemistry with use of the AI tool. Further studies with a larger data set and more pathologist reviews are needed to confirm the validity and utility of implementing AI solutions in community practice setting.

Do different digital imagers used with the AIxURO system impact performance in urine cytology?

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

[1]AIxMed, Inc., Santa Clara, CA, United States

[2]Show Chwan Memorial Hospital, Pathology, Changhua, Taiwan

Email: chu.peiyi@msa.hinet.net

Background

Digital cytopathology has recently advanced to convert cytology slides into whole-slide images (WSIs) for artificial intelligence (AI) applications, offering advantages comparable to conventional microscopy while providing quantitative analysis to assist diagnosis and improve workload efficiency. For urine cytology in bladder cancer screening and diagnosis, we developed AIxURO, an AI-assisted urine cytology system incorporating a disease-specific AI algorithm for analyzing WSIs generated by digital imagers. Current challenges for clinical adoption include the variability in images produced by different commercial imagers and the impact of user experience in digital cytology on diagnostic results. Therefore, we conducted a hybrid three-arm study at Show Chwan Hospital, a regional teaching hospital, and AIxMed, a company specializing in digital cytopathology. This study compared WSIs produced by two commercial imagers for the AI-assisted system to manual microscopy in clinical practice.

Methods

We evaluated 183 archived and de-identified urine cytology slides from Show Chwan Hospital, (SCH), including 140 Cytospin and 43 BD CytoRich (SurePath) slides. An independent expert panel rendered consensus diagnoses (ground truth) for the 183 slides of which 83 cases were Negative for High-Grade Urothelial Carcinoma (NHGUC), 45 Atypical Urothelial Cells (AUC) cases, 27 Suspicious for High-Grade Urothelial Carcinoma (SHGUC) cases, and 28 High-Grade Urothelial Carcinoma (HGUC) cases. The slides were digitized into WSIs using Leica Aperio AT2 and Hamamatsu S360 imagers. A total of 183 WSIs produced by either Leica or Hamamatsu were individually analyzed using a deep-learning-based AI algorithm following The Paris System (TPS) 2.0 guidelines. AIxURO identified candidate cancer cells in images (AIxURO on Leica or Hamamatsu WSIs), which were then reviewed by cytopathologists (CP) and cytologists (CT) using viewer software to assist in bladder cancer diagnosis.

Upon completing a formal AIxURO training program, the evaluation process began with one CP and two CTs from SCH, along with three CTs from AIxMed who independently examined the 183 slides manually using TPS criteria for result reporting (microscopy arm). Following a two-week washout period, the same reviewers used the viewer software to assess AI-identified candidate cancer cells within the corresponding Leica-generated WSIs and provide their diagnoses (AIxURO on Leica WSIs arm). After another two-week break, the reviewers used Hamamatsu-produced WSIs to make diagnoses (AxURO on Hamamatsu WSIs arm). This process yielded a total of 1098 diagnostic result sets, including three arms (microscopy + AIxURO on Leica WSIs + AIxURO on Hamamatsu WSIs = 3294 reviews). These results were then compared to ground truth for performance evaluation. We analyzed performance metrics and time taken for slide reporting in each arm to evaluate the impact of AIxURO on different imager-generated WSIs in clinical practice, and the influence of reviewers' experience levels on AI-assisted result reporting.

Results

For the binary analysis, one hundred (100) cases by ground truth were labeled positive for AUC+ (AUC, SHGUC, and HGUC) and 83 as negative (NHGUC). When assessing the performance of the three reviewers, AIxURO on Leica WSIs outperformed both AIxURO on Hamamatsu WSIs and microscopy in terms of sensitivity (85.0 % vs. 82.9 % vs. 83.7 %) and specificity (90.0 % vs. 89.6 % vs. 89.4 %). Additionally, for slide reporting duration (s = seconds), AIxURO on Leica WSIs demonstrated the shortest mean time overall compared to the other two arms (37.4 s vs. 53.1 s vs. 82.0 s).

In comparing reviewers with distinct experience levels in digital cytology, within AIxURO on Leica WSIs, the three more experienced reviewers from AIxMed demonstrated higher sensitivity (91.3 % vs. 78.7 %) but lower specificity (87.7 % vs. 92.3 %) than the three less experienced reviewers from SCH. Furthermore, AIxMed reviewers spent less reporting time than SCH reviewers (30.6 s vs. 44.4 s). Conversely, in AIxURO on Hamamatsu WSIs, the AIxMed reviewers outperformed the SCH reviewers in both sensitivity (86.3 % vs. 77.7 %) and specificity (91.2 % vs. 88.0 %) and exhibited shorter reporting times (43.6 s vs. 62.6 s).

Conclusions

(1) The two AI-assisted methods performed comparably to microscopy. AIxURO on Leica WSIs showed a 1.3 % increase in sensitivity, while AIxURO on Hamamatsu WSIs had a 1.7 % decrease. Specificity increased by 0.6 % for Leica and 0.2 % for Hamamatsu. These results indicate that both digital imagers with AIxURO effectively assist in accurate cytologic interpretations. (2) Reviewers using AIxURO on Leica WSIs reduced reporting time by 52.4 %, while those using Hamamatsu WSIs reduced it by 35.2 %, compared to microscopy. (3) AIxMed reviewers, with more experience in digital cytology, showed a 12.6 % increase in sensitivity but a 4.6 % decrease in specificity compared to SCH reviewers using AIxURO on Leica WSIs. For AIxURO on Hamamatsu WSIs, they had an 8.6 % increase in sensitivity and a 3.2 % increase in specificity compared to SCH reviewers. (4) AIxMed reviewers spent 31.1 % less time using AIxURO on Leica WSIs and 30.4 % less time using AIxURO on Hamamatsu WSIs compared to SCH reviewers. (5) These findings suggest that the AI-assisted system performs consistently across different digital sources. AIxURO showed comparable performance on Leica and Hamamatsu images and greater efficiency than conventional microscopy, indicating its potential applicability to other commercial imagers. Additionally, user experience in digital cytology significantly impacts diagnostic performance and efficiency, emphasizing the demand for training and education before clinical implementation.

Transforming bladder cancer diagnosis via AI-aided digital cytology integration in clinical workflow

Samer Khader,[1] Sigfred Lajara,[1] Daniel Geisler,[1] Jackie Cuda,[1] Caroline Horne,[1] Barbara Crothers,[2] Karen Atkison,[2] Wei-Lei Yang,[2] Chih-Yun Lin,[2] Tien-Jen Liu,[2] Rajiv Dhir[1]

[1]University of Pittsburgh Medical Center, Pathology, Pittsburgh, PA, United States

[2]AIxMed, Inc., Santa Clara, CA, United States

Email: dhirr@upmc.edu

Background

Recent advancements in digital cytopathology enable transforming cytology slides to whole-slide images (WSIs) for artificial intelligence (AI) applications. Clinical applications of AI have shown promise in improving diagnostic accuracy when compared to conventional microscope examinations, while also providing quantitative and more objective parameters and improving workload efficiency. We aim to advance digital urine cytology with AI analysis for bladder cancer screening and diagnosis, creating a hybrid clinical application. We developed an AI-assisted digital urine cytology system (AIxURO), which consists of a digital imager to digitize urine cytology slides and a disease-specific AI algorithm for analyzing WSIs. AIxURO displays cells with cancer risk as thumbnail images, including nuclear size and nuclear to cytoplasmic ratio via viewer software for review by cytopathologists (CP) and cytologists (CT), assisting in accurate diagnosis. To evaluate this new application in a clinical setting, we conducted a two-

armed study at the University of Pittsburgh Medical Center (UPMC) comparing AIxURO with current practice. We hypothesized that AIxURO integration would show non-inferior diagnostic performance and improved efficiency in clinical practice.

Methods

Two hundred (200) archived and de-identified ThinPrep urine cytology slides from UPMC were selected. An independent expert panel provided consensus diagnoses (ground truth) for the 200 slides, of which 100 cases were Negative for High-Grade Urothelial Carcinoma (NHGUC), 35 Atypical Urothelial Cells (AUC) cases, 32 Suspicious for High-Grade Urothelial Carcinoma (SHGUC) cases, and 33 High-Grade Urothelial Carcinoma (HGUC) cases. These slides were digitized to WSIs using a software-customized digital imager designed for cytology scanning (Mikroscan SLxCyto). Each WSI was examined using a specialized deep-learning-based AI model aligned with The Paris System (TPS) 2.0 guidelines.

After completing a formal AIxURO training program, 1 CP and 2 CTs began the study by independently examining the 200 study slides manually using TPS criteria (microscopy arm). After a two-week washout period, the same reviewers used the viewer software to assess AI-identified candidate cancer cells within the corresponding WSIs and render their diagnoses (AIxURO arm). This resulted in a total of 600 diagnostic result pairs, which were compared with the ground truth for performance assessment. We analyzed performance metrics with two different diagnostic positive thresholds, AUC+ (AUC, SHGUC, and HGUC) and SHGUC+ (SHGUC and HGUC), based on the 4 categories of TPS. In addition, the time needed for slide evaluation in the two arms was documented to determine the impact of AIxURO in clinical practice.

Results

For the AUC+ threshold analysis, 100 cases were categorized as positive and 100 as negative based on the ground truth. AIxURO demonstrated higher sensitivity than microscopy (85.0 % vs. 79.3 % overall; 80.8 % vs. 60.0 % for CT1; 87.0 % vs. 83.0 % for CT2; and 88.0 % vs. 95.0 % for CP). However, AIxURO exhibited lower specificity than microscopy (92.0 % vs. 98.0 % for CT1; 78.0 % vs. 89.0 % for CT2; and 87.0 % vs. 96.0 % for CP). When using the SHGUC+ threshold for positive diagnoses, there were 65 positive and 135 negative cases. AIxURO demonstrated lower sensitivity (74.9 % vs. 76.9 % overall; 50.8 % vs. 46.2 % for CT1; 78.5 % vs. 84.6 % for CT2; and 95.4 % vs. 100.0 % for CP) and lower specificity (96.0 % vs. 97.5 % overall; 98.5 % vs. 99.3 % for CT1; 94.1 % vs. 95.6 % for CT2; and 95.6 % vs. 97.8 % for CP) compared to microscopy.

Regarding slide evaluation time (s = seconds), AIxURO markedly reduced the overall mean time for the 3 reviewers compared to microscopy (37.4 s vs. 102.6 s). When comparing AUC+ and negative cases, AIxURO markedly decreased the overall mean time compared to microscopy (45.3 s vs. 116.4 s for AUC+ cases and 26.5 s vs. 88.9 s for negatives). Similarly, for SHGUC+ cases, AIxURO substantially reduced the overall mean time compared to microscopy (42.8 s vs. 108.3 s for SHGUC+ and 32.6 s vs. 99.9 s for negatives). For AUC+ cases, CTs and CP spent a comparable mean time when using AIxURO (37.2 s vs. 37.8 s), while CTs spent almost double the time compared to CP with microscopy (121.6 s vs. 64.6 s).

Conclusions

(1) For bladder cancer diagnosis at the AUC+ threshold, AIxURO demonstrated a 5.7 % increase in sensitivity and an 8.6 % decrease in specificity compared to microscopy, suggesting while AIxURO helps reviewers identify more positive cases, it also resulted in a higher number of false positives. (2) At the SHGUC+ threshold, AIxURO exhibited a 2.0 % decrease in sensitivity and a 1.5 % decrease in specificity, indicating that reviewers may diagnose more cases as AUC than microscopy. (3) Reviewers using AIxURO saved 63.5 % of time evaluating slides compared to microscopy and achieved more consistent evaluation times. (4) CTs spent substantially less evaluation time compared to the CP when using AIxURO, with a 69.4 % time savings for CTs versus 41.5 % for the CP, indicating that CTs benefit more in terms of efficiency from using the AI-assisted system. (5) The findings highlight that our AI-assisted system AIxURO provides non-inferior performance and greater efficiency compared to conventional urine cytology. The increase in false positive cases and reclassification of AUC cases by AIxURO may be due to the quality of images displaying cytomorphological features. Further studies on image quality in AI-assisted bladder cancer diagnosis are ongoing.

Revolutionizing bladder cancer diagnosis: Insights on urine cytology from a digital pathology laboratory

Juan Santa-Rosario[1], Maria del Mar Rivera Rolon,[1] Erik Gustafson,[1] Suheidy Cardona Rivera,[1] Julyanas Ramirez Rodriguez,[1] Wei-Lei Yang,[2] Chih-Yun Lin,[2] Tien-Jen Liu[2]

[1]CorePlus Servicios, Digital Pathology, Carolina, PR, United States

[2]AIxMed, Inc., Santa Clara, CA, United States

Email: juan.santa@corepluspr.com

Background

Digital pathology has progressed rapidly, particularly in the conversion of histology slides into whole-slide images (WSIs) for artificial intelligence (AI) applications. These clinical tools show promise in achieving accuracy comparable to conventional microscopy, while also improving workflow efficiency. However, digital cytology has lagged behind histology in its development. We aim to advance digital urine cytology with AI analysis and assistance for bladder cancer screening and diagnosis. We developed an AI-assisted digital urine cytology system (AIxURO) encompassing a digital imager for urine cytology slides and a disease-specific AI algorithm for WSI analysis. The system digitizes and presents candidate cancer cells for review by cytopathologists and cytologists, facilitating accurate diagnoses. Notably, the transition to digital pathology often requires user training for optimal system utilization. We observed that prior digital pathology experience, coupled with AI assistance, can enhance diagnostic accuracy and efficiency. In this study, we compared AIxURO with conventional microscopy for urine cytology reporting in bladder cancer diagnosis at a fully digital pathology laboratory, focusing on performance and efficiency.

Methods

One hundred archived, de-identified BD CytoRich (SurePath) urine cytology slides with expert consensus diagnoses served as ground truth, including 68 Negative for High-Grade Urothelial Carcinoma (NHGUC), 11 Atypical Urothelial Cells (AUC), 7 Suspicious for High-Grade Urothelial Carcinoma (SHGUC), and 14 High-Grade Urothelial Carcinoma (HGUC) cases from a digital pathology laboratory (CorePlus). Slides were digitized to WSIs using a software-customized digital imager designed for cytology scanning (Mikroscan SLxCyto). Each WSI was analyzed with a deep-learning AI model based on The Paris System (TPS) 2.0 guidelines. After formal AIxURO training, one cytopathologist (CP) and two cytologists (CTs) independently examined the slides manually, using TPS criteria for result reporting (microscopy arm). After a two-week washout period, the same reviewers used the viewer software to examine AI-identified candidate cancer cells within the corresponding WSIs and made their diagnoses (AIxURO arm). A total of 300 resulting diagnostic pairs were compared against ground truth. Performance metrics and reporting times were analyzed.

Results

For the binary bladder cancer diagnosis (32 positive cases: AUC, SHGUC, and HGUC; 68 negative cases: NHGUC), AIxURO exhibited higher overall sensitivity (88.5 % vs. 86.5 %; 90.6 % vs. 90.6 % for CT2; 84.4 % vs. 87.5 % for CT1; and 90.6 % vs. 81.3 % for CP) but lower specificity (93.6 % vs. 97.6 % overall; 98.5 % vs. 98.5 % for CT2; 91.2 % vs. 94.1 % for CT1;

and 91.2 % vs. 100.0 % for CP) than microscopy from the three reviewers. AIxURO substantially reduced mean reporting time for all reviewers compared to microscopy (13.6 s vs. 83.3 s overall; 11.9 s vs. 92.4 s for CT2; 11.9 s vs. 127.5 s for CT1; and 17.0 s vs. 31.4 s for CP). This time reduction was particularly pronounced for negative cases (7.7 s for AIxURO vs. 72.3 s for microscopy). Additionally, AIxURO led to more consistent reporting times across reviewers (11.9 s for CT2; 11.9 s for CT1; and 17.0 s for CP).

Conclusions

(1) AIxURO demonstrates non-inferior performance compared to microscopy for bladder cancer diagnosis, with a 2.0 % increase in sensitivity and a 4.0 % decrease in specificity, suggesting that our AI-assisted system effectively supports accurate clinical diagnosis in a digital pathology laboratory. (2) Reviewers using AIxURO spent markedly less reporting time than with the microscopy method (83.8 % time saved overall; 88.5 % time saved in positive and 89.3 % in negative cases). (3) The reporting times of the three reviewers using microscopy varied notably (31.4–127.5 s), while with AIxURO, they became much more consistent (11.9–17.0 s). (4) Notably, for negative cases, reviewers using AIxURO spent only 10 % of the time required by microscopy. This is critical, as most urine cytology cases are negative, and the time saved by AIxURO can substantially improve laboratory efficiency. (5) This study demonstrates that digital cytology with AI assistance delivers the desired performance standards similar to histology, the benefit of reducing the urine cytology diagnosis workload, and the potential to be used in cytology clinical settings.

Expanding cytology AI menu with the Genius® digital diagnostics system

Shaoqing Peng, Raymond Jenoski and Sid Mayer
Diagnostic Instrument Engineering, Product Development, Hologic Inc., Marlborough, Massachusetts

ABSTRACT

The Hologic Genius Digital Diagnostics System (GDDS) incorporates advanced volumetric scanning technology, specifically designed to manage the unique depth of field challenges required for cytology.

The system was recently FDA-cleared along with the Genius Cervical AI algorithm for cervical cancer screening. This Artificial Intelligence approach is adaptable and can be extended to other cytology specimens, including urine and thyroid FNA samples, broadening the diagnostic capabilities of GDDS.

All algorithms can operate within a unified architecture and benefit from a gallery-based review workflow. Additionally, the system could generate a slide-level disease risk score to further support Cytologists and Cytopathologists in their diagnostic review.

Cytology algorithm architecture

Genius Cervical AI uses the architecture consisting of Object Location, Inspection, Presentation and Risk scoring components. The Risk Scoring component is not in the existing implementation but could be added to provide additional decision support information.

Cell location

Liquid-based cytology preparations such as ThinPrep® slides consist of a thin layer of well-separated cells or small groups, with each slide containing as many as 80,000 cells. Every cell must be identified for further analysis as indicated. Cell nuclei can be detected by a variety of image processing methods or deep learning object location frameworks. It is not necessary to precisely determine the cell boundaries or separate overlapping cells. The object location function captures a “snapshot” region around each cell to be passed on to the next stage, Inspection.

Inspection

Each cell is inspected by a convolutional neural network (CNN), which performs one or more inferences to extract salient characteristics. These may be a classification of the cell, identification of specific features (hyperchromaticity, etc.), or other outputs. Labeled data to train the CNN requires time-intensive work by multiple Cytologists or Cytopathologists. More than one subject matter expert is generally used to control variability between reviewers. Because of the limited amount of data, it is essential to start with a pre-trained CNN network such as a model trained on the ImageNet dataset, and then use transfer learning to adapt to the specific cytology problem domain.

Transfer learning was used to create the Genius Cervical AI algorithm model. In turn, this model can serve as the basis for other cytology models.

Presentation

The presentation stage creates a gallery of the most relevant cells for case interpretation. Details of ranking and selecting cells for their clinical relevance depend on the nature of the particular specimen type and disease state being assessed. This can involve combining computed inference results for the cells, model performance data such as ROC curves, and heuristic rules. The selection rules must be developed in conjunction with subject matter experts with the goal of providing the most effective decision support data. In some cases, additional AI inference results should be displayed. For example, the Genius Cervical AI displays the squamous cell count to aid in determining specimen adequacy. At this point, the algorithm has located and identified the cells for the reviewer, who can now focus on their clinical interpretation. The gallery can be reviewed at a glance, which improves efficiency considerably compared to manual microscope review of the glass slide, while also increasing sensitivity. Genius Cervical AI reviews are 2× to 4× faster than manual glass slide reviews, and similar speed improvements could apply to other cytology sample types.

Risk scoring

A second tier of deep learning inference can be applied to the entire set of cell inspection results to form a risk score for the case appropriate to the specimen type and disease being assessed. This additional information can be displayed along with the gallery information to provide additional decision support. There is a balance of providing additional contextual information to increase sensitivity without overly decreasing the specificity of the human interpretation.

Discussion

The Genius Digital Diagnostics System and the Genius Cervical AI algorithm is now in routine clinical use in laboratories all over the world. This approach for cervical cancer screening can be utilized in other diagnostic domains, expanding the menu of digital assays available on the digital cytology platform. The foundation built for Cervical AI can be adapted and transfer learning can be used to adjust algorithms quickly and efficiently with new data. The gallery presentation of diagnostic cells improves both speed and accuracy over manual glass slide review. The addition of objective slide-level risk scores can provide even further benefit for cytologists and pathologists in their diagnostic review.

Role of artificial intelligence (AI) in distinguishing atypical cribriform lesions of prostate

Zoobia Khan
UF Health - Jacksonville, Pathology, Jacksonville, FL, United States
Email: zoobia.khan@jax.ufl.edu

Introduction

Intraductal carcinoma of prostate (IDC-P) is a distinct WHO neoplastic entity that is typically associated with high-grade prostatic carcinoma and shares morphological similarities with high-grade prostatic intraepithelial neoplasia (HGPIN). Although rare, isolated IDC-P has been reported in prostate needle core biopsy (PNB) and may warrant an immediate re-biopsy. In this study, we aimed to utilize a commercially available artificial intelligence (AI) assisted analysis to distinguish between the two entities.

Methods

We selected 30 PNBs of IDC-P, with ($n = 23$) and without associated invasive cancer ($n = 7$), between 2017 and 2023. In addition, 13 PNBs each of HGPIN and benign prostatic tissue were randomly selected and digitally scanned. Histopathological diagnosis was independently confirmed by 2 surgical pathologists, including a genitourinary pathologist. AI analysis was performed on all PNBs.

Results

29 of 30 (96.7 %) PNBs of IDC-P were reported as “Suspicious” for prostatic adenocarcinoma (PCa) by AI while 1 PNB (3.3 %) was reported as “not suspicious” for PCa.

Upon review of the H&E section of the “not suspicious” PNB, the reported IDCP focus was small (<1 mm) and was clinically reported as “Suspicious for IDC-P".

In addition, this patient had 5 additional PNBs with IDC-P only.

No invasive carcinoma was reported in this patient's PNBs.

P63 or PIN4 immunohistochemical studies were performed on all IDC-P-only PNBs, supporting the diagnosis.

All “suspicious” IDC-P cores without invasive cancer (85.7 %) were graded as Gleason score $4 + 4 = 8$ (Grade group 4) by AI.

Among IDC-P cores with invasive cancer, 5 cores were graded as grade group 2, 2 cores were graded as grade group 3, 13 cores were graded as grade group 4, and 3 cores were graded as grade group 5 by AI.

All cases of HGPIN and benign prostatic tissue ($n = 26$; 100 %) were reported as “not suspicious” by AI.

Conclusion

As cribriform glands, IDC-P mimics invasive Gleason pattern 4.

By reporting PNBs with IDC-P only as “suspicious”, AI can distinguish IDC-P from HGPIN. However, the results should be interpreted with caution as the majority of IDC-P only cores were graded by AI as grade group 4, which may impact patient management.

Currently, IDC-P is not recommended to be graded by the International Society of Urological Pathology (ISUP) guidelines.

Histomorphology and immunohistochemical workup can further assist in distinguishing IDC-P from other atypical cribriform lesions, including prostatic ductal adenocarcinoma and urothelial carcinoma involving prostatic ducts.

Three-dimensional models for surgical margin assessment: A novel approach

Josselyn Andrea Hernández Chinchilla[1], Cecilia Mercedes Lopez Valdivia,[2] Carlos Alberto Zac Romero[2]

[1]Hospital Universitario y Politécnico La Fe, Pathology, HUIP La Fe Valencia, Spain

[2]HUIP La Fe, Pathology, Valencia, Spain

Email: joanherchi@hotmail.com

Introduction

Surgical resection with clear margins is crucial for the successful treatment of various oncological pathologies. Accurate margin assessment requires careful microscopic and macroscopic examination. However, macroscopic examination is limited by the inability to review and correct errors at a later time. It is important to conduct an evaluation of each surgical specimen that facilitates margin involvement determination using microscopic tools. Furthermore, providing the possibility to use macroscopy as a teaching tool is valuable. Some surgical specimens with expandable margins through surgery or potential radiotherapy treatment require precise pathological reporting regarding location and distance. Currently, there are multimedia tools available for graphical representation of specimens, such as diagrams, photographs, or videos, but these have limitations in terms of the three-dimensional perception of the specimen for pathologist and multidisciplinary team comprehension. Therefore, the use of 3D models of surgical specimens has been proposed for adequate reporting of oncological specimens with complex margins.

Materials and methods

In Hospital La Fe de Valencia, Spain, 10 head and neck cases were selected. Using the Polycam™ mobile application for Android™ and iPhone™ operating systems, three-dimensional scanning of surgical specimens from head and neck surgery (maxillofacial and otorhinolaryngology) requiring margin involvement reporting was performed. Surgical specimens were prepared and placed on carving tables with white backgrounds and adequate lighting, followed by three-dimensional scanning, focusing on the specimen in all dimensions, with a minimum of 60 shots per specimen. Subsequently, routine macroscopic and microscopic examination was conducted, reporting distances and specific areas of involvement. Image correlation between micro and macroscopic views was performed, with graphical representation of the model.

Results

The graphical representation of the 3D models was satisfactory in all cases, resulting in a total of 10 models that allowed for precise spatial reporting of margin involvement.

Conclusion

Despite the importance of obtaining standardized and meticulous samples, significant divergences in practice still exist due to technical limitations in specimen representation. Obtaining a three-dimensional understanding of surgical specimens is essential for the pathologist and treating physician. Establishing a precedent for the use of 3D models in pathology and developing technological tools that support macroscopic specimen examination are necessary. Integration of this need into multidisciplinary teams is suggested through the presentation of this technique.

Digital image-based data sharing from procured research tissues, 2019–2023

Randy Mandt, Jesse Kulewsky, Konstantin Shilo, Anil Parwani

The Ohio State University, Pathology, Columbus, OH, United States

Email: randy.mandt@osumc.edu

Background

Cooperative Human Tissue Network (CHTN) is an NCI-sponsored prospective human tissue procurement program that provides quality human tissue and clinical data to approved investigators. The Ohio State University (OSU) operates as the CHTN Midwestern Division (MWD). Tissue data collection is completed prior to shipment to investigators or as soon as available after shipment of fresh tissues. All data is maintained in a custom Research Tissue Procurement - Information System (RTP-IS).

Design

Tissue Procurement (TP) Services procures a quality control (QC) sample adjacent to the remnant tissue procured for the investigator. The QC sample is formalin-fixed paraffin-embedded. Each QC block has a slide cut, and H&E stained. This stained QC tissue is scanned (ScanScope XT, Leica Biosystems) and the image file is placed on a secure server. The Pathologist views the digital Whole Slide Image (WSI) and associated pathology report and reports the % Region of interest (ROI) of the target tissue along with % necrosis and pathologist review result (Pass, Fail, etc.). The pathologist approves the final histopathological diagnosis based on standard pathology vocabulary and other specified data (see Table 1). The WSI, pathology report and sample descriptive data with no PHI or donor identifiers are maintained in a Repository available for distribution to investigators or indefinitely in RTP-IS. A de-identifier can be used to retrieve the case record and expanded clinical data if approved.

Results

Of the 12,173 research tissue samples examined for tissue quality and data entry using WSI, 93.6 % passed and 6.4 % failed due to wrong/insufficient viable cells in the sample.12.5 % of the passed samples were *pass adjusted* on review based on specific histopathologic findings vs initial *Preliminary Diagnosis*.

Conclusion

Retrieval of tissue digital images through Informatics management provided by RTP-IS electronic records from a data centralized repository is convenient, timely and avoids the need to transport glass slides from storage for case review. The stored images, paraffin blocks and clinical data are retrieved for specific downstream investigator requests or just for digital image research. Data sharing is important to avoid duplicate efforts by the researcher and provides the investigator with the opportunity to immediately review the features of the research tissues prior to initiation of laboratory testing.

A practical guide for scanner selection for clinical purposes

George Yousef[1], Blaise Clarke,[1] Christine Bruce,[1] Karen Weiser[2]
[1]University Health Network, Laboratory Medicine Program, University of Toronto, Toronto, Canada
[2]University health Netwrok, Laboratory Medicine Program, Toronto, Canada
Email: George.Yousef@uhn.ca

There is currently a growing wave of transformation into digital pathology practice. The choice of the scanner is a key consideration in the process. This can be tricky with the lack of in-depth understanding of the different flavors and the many vendors available in the market.

This is especially true for larger institutions with multiple locations. At the University Health Network (UHN) in Canada, we recently switched to a fully digital pathology practice in our multicenter academic institution. Scanners utilization was needed for primary diagnosis (in both main locations and satellite sites), remote teleconsultation, intra-operative consultation (frozen section) diagnosis, education/ multidisciplinary rounds in addition to AI and image analysis applications, and here we share our experience in this domain.

First, it should be noted that scanners are not equivalent to microscopes. The image that you view on a computer monitor is a reconstruction of series of images that are digitally acquired, electronically stored, and stitched together and viewed using software.

Generally, components of a modern scanner are the objective lens, scanning camera (can be of different shapes and sizes), objective lens (one or multiple magnifications can be available), robotics for moving the slide, illumination sources (brightfield and/or fluorescent) and finally the computer software for assembly and visualization.

The performance of the scanner depends on multiple factors including the scanning time (speed), slide loading capacity, image quality and resolution, ability of z stacking (scanning at multiple levels), among others. Scanner capacity varies greatly from 1 to 1000 slides per scanners. The choice of capacity should depend on workload. With moderate volumes, it might be more efficient to purchase multiple smaller capacity scanners compared to one scanner with high capacity. Most modern scanners allow autoloading functionality. Continuous loading functionality that enables the users to upload slides while others are being processed. This is a great advantage for a dynamic workload.

Most scanners accept only standard slide thickness and dimensions. Some vendors provide the ability to accommodate larger whole mount slides. Scanners also differ in their tissue finding ability and the ability to manually adjust the focal points for tissue detection. Other considerations include the scanner dimension and weight, the ability of the scanner to read multiple patient IDs (codes).

Overall, the performance of the scanner can be calculated as the combination of three main elements: speed, capacity, and automation (automatic tissue detection, autofocusing and automatic digital QA). It should be noted that automation comes at a price. It limits the ability for manual adjustment and selection of focal points.

For most surgical pathology and immunohistochemistry applications, 20× objective will do the job efficiently. 40× magnification is needed for cytology and in-situ hybridization. The ability for image compression is an additional advantage provided by some vendors.

It should be noted that the resolution of the scanner is better described in “microns per pixel” rather than the magnification power. To simplify, generally speaking a 20× magnification save image at 0.5 μm per pixel. FDA approval or equivalent can be also essential for certain institutions.

In our institution, we elected to go with multiple capacity scanner from multiple vendors, all of them are agnostic to LIS and workflow solutions. When assessing the scanner performance, it is important not to rely only on technical specifications provided by the vendor, but to also look test the performance in real life with multiple types of specimens, including archived.

Quality assurance procedures for high-throughput digital pathology

Jodi Casper, Mary Ann Klein, Sylvia Asa
University Hospitals Cleveland Medical Center, Digital Pathology, Cleveland, OH, United States
Email: Jodi.Casper@UHHospitals.org

The implementation of digital pathology requires the build of a complex information technology (IT) system to allow integration with the laboratory information system, case organization for viewing, and storage. The importance of slide quality is a potential hurdle as laboratories transition from the microscope, which allows the pathologist to compensate for artifacts, to the scanner, which does not. High quality slide preparation is critical to ensure that the whole slide images (WSIs) are complete and without flaws. We have created several novel innovations for slide preparation using simple tools to ensure that the surfaces are clean, that the coverslips are centered and cover the tissue, and that the entire tissue area is scanned. For optimal workflow one of the novel tools used during slide preparations is a mini glass container where a plastic moistener applicator covered in a piece of terry cloth is placed in the hole that is formed in the cap of the glass container; these containers are identified with a biological reagent label stating which reagent it contains. The glass dauber bottle applicator is used to apply 100 % Alcohol or Xylene on the slide to wipe off eosin residue, finger prints and paraffin wax effortlessly. A painter scraper blade is used to trim off excess label overhang and to remove mounting media residue. Last, a simple metal nail file is used to trim any overhang of cover glass on the slide. These novel innovations are simple and cost-effective tools and

allow slide preparation to be fast and simple; in our experience it takes less than 10 min per tray of 20 slides (or 15 min per rack of 30 slides). They ensure a successful and flawless workflow, saving time and money on troubleshooting errors and/or having downtime due to costly repairs to the scanners.

An AI-powered whole slide imaging for breast cancer detection: A comparative study analysis

Nada Shaker[1], Ahmad Shahin,[2] Nuha Shaker,[3] Mohammed Shaker,[2] Zaibo Li[1], Anil Parwani,[1] Noor Shaker[2]

[1]University of California, San Francisco, CA, United States

[2]Spatial X, NA, SpatialX, United Kingdom

[3]University of Pittsburgh Medical Center, Department of Pathology, Pittsburgh, United States

Email: Nada.shaker@ucsf.edu

Introduction

In recent years, the field of computational pathology has experienced remarkable growth, largely propelled by the advancement of deep-learning algorithms.

Method

We aim to evaluate the performance of the AI platform to stream the entire process, starting from preprocessing and cleaning whole-slide images (WSIs) to training deep learning models and performing scalable predictions. The novel platform encompasses over 10 distinct foundational vision models. Each training process outputs a pool of up to 40 parameterized ML models, each model is the result of hyperparameter tuning of hundreds of experimental variations. The comprehensive ensemble approach is fully data-driven, ensuring optimal performance for each specific task. The approach is generic and can be used to perform a wide range of tasks, from cell and tissue segmentation to cancer classification and tumor-micro environment (TME) analysis. Annotated WSIs are uploaded to the AI engine to start the training process. This includes clearing and pre-processing the data to ensure robust performance; data partitioning into training, validation, and testing to evaluate AI models and ensure generalizability of results and avoid overfitting; visualization of the results, and output assessment. Once the AI models are trained, they can be deployed and used at a large scale to screen WSI data.

Moreover, for tasks where the spatial arrangement of cells and tissues is important, such as analysis of the TME, the platform captures valuable spatial features by utilizing a combination of Convolutional Neural Networks (CNN) and Graph Neural Network (GNN) models, which account for the spatial organization of cells and tissues within WSIs. To capture spatial information, CNNs are trained for the classification of cell and tissue features, and this information is incorporated as nodes and features within the GNN model. This enables our models to comprehend both local and global information.

We aim to validate the performance of a novel platform in breast cancer classification by applying it to a publicly available dataset for breast cancer detection, which consists of 53 WSIs. A comparison of AI algorithm performance to other AI deep learning models was reported in the literature.

Results and discussion

Breast cancer remains a prominent global issue, characterized by its high prevalence worldwide and its status as a leading cause of cancer-related deaths among women.

Diagnostic applications of artificial intelligence in breast pathology include tumor detection, metastatic deposits detection in lymph nodes, breast cancer grading and subtyping, assessment of tumor microenvironment, mitotic figures counting, and receptor status assessment. These tools aid in identifying patients at higher risk of recurrence or metastasis, guiding healthcare professionals in tailoring personalized treatment plans for better outcomes.

The presented AI platform exhibits a comprehensive and robust workflow for WSI analysis in breast cancer detection, encompassing various critical components from preprocessing and training deep learning models to scalable predictions. The incorporation of over 10 distinct foundational vision models and the generation of up to 40 parameterized machine learning models through data-driven, hyperparameter-tuned processes highlight the platform's versatility. The spatial arrangement of cells and tissues is a pivotal aspect of tasks like TME analysis, and the platform addresses this effectively through a combination of CNNs and GNN models. The utilization of CNNs for local feature classification and their integration into GNN models.

In this comparative study evaluating various models for breast cancer detection, the EXPLORE model also showcased a high accuracy rate of 99 %. It stood out with a Dice coefficient of 0.79, a Jaccard index of 0.76, and precision and recall rates of 0.79 and 0.80, respectively. Both the FCN and DeepLab models demonstrated a high accuracy performance of 99 %. The FCN model exhibited a Dice coefficient of 0.75, a Jaccard index of 0.72, and precision and recall rates of 0.84 and 0.78, respectively. Similarly, the DeepLab model delivered a Dice score of 0.77, a Jaccard index of 0.71, and precision and recall values of 0.77 and 0.78, respectively. Overall, The comparative analysis emphasized the EXPLORE model's superiority across all metrics, highlighting its particularly strong performance in the recall, a critical metric in healthcare applications.

Conclusion

The innovative platform offers a comprehensive AI engine that seamlessly manages WSI data from analysis and pre-processing to training and deployment. Its user-friendly interface streamlines the workflow, providing ease of use alongside superior performance in breast cancer detection compared to other AI deep learning models with a particularly advantageous position in the recall, a crucial metric in healthcare applications, emphasizing its potential clinical utility.

Evaluating the use of digital image analysis for quality control in histopathology

Shakti Kumar Yadav[1], Riya Singh,[2] N. Kapoor[2]

[1]All India Institute of Medical Science Bhopal, Pathology and Lab Medicine, Bhopal, India

[2]All India Institute of Medical Sciences Bhopal, MBBS Student, Bhopal, India

Email: Nada.shaker@ucsf.edu

Background

Histopathological diagnosis is critically dependent on various technical factors, among which the quality and completeness of histological sections are paramount for accurate interpretation. The current process of ensuring the quality of these sections is predominantly manual and highly dependent on the expertise and experience of histotechnicians. Given the increasing demand for accuracy and efficiency in medical diagnostics, there is a pressing need for more objective and automated methods to assess the quality of histological sections.

Histopathology laboratories have long focused on the analytical aspects and staining procedures for quality control. However, processes such as tissue embedding and sectioning remain manual even in advanced laboratories with automated instruments. These manual steps are subjective and rely heavily on the observation skills of histotechnicians, leading to variability in the quality of histological sections. This study aims to introduce and evaluate the application of digital image analysis for quality control in histopathology, presenting a proof-of-concept for a more objective and automated approach.

Methods

This cross-sectional observational study was conducted at a tertiary care hospital and medical college. A total of 1000 tissue blocks and their corresponding slides were selected for analysis. To standardize the image capture process, a customized box with a fixed setup was designed. An iPhone 13 Pro was used as the image capture device, and images were stored in JPEG format.

Images of the tissue blocks (referred to as Digiblock) and their corresponding slides (referred to as Digislide) were captured using this setup. The images were analyzed using the ImageJ 1.53 K application (an open-source software from NIH) to measure the area of the tissue sections on both the block and the slide. The DigislideQC score was calculated by dividing the area of the tissue on the slide by the area of the tissue on the block. This score was then compared with the number of recuts requested for incomplete sections to determine its accuracy and effectiveness.

Results

The study analyzed 1000 tissue blocks and their corresponding slides, with a total of 249 (24.9 %) tissue blocks being sent for recutting. The mean area of the Digiblock was 7.12 cm^2 (SD 2.35 cm^2), ranging from 0.59 to 13.43 cm^2, while the mean area of the Digislide was 4.99 cm^2 (SD 1.68 cm^2), ranging from 0.19 to 9.44 cm^2. A significant difference was found between the size of the tissue in the block and the slide ($P < 0.0001$).

The DigislideQC score ranged from 0.1 to 0.99. The receiver operating characteristic (ROC) curve for the DigislideQC score showed an area under the curve (AUC) of 98.8 %. A cut-off value of 0.65 yielded a sensitivity of 99.6 % and a specificity of 96.7 %, indicating that the DigislideQC score can accurately predict the need for recuts before slides are submitted for histopathology reporting.

Discussion

The findings of this study highlight the potential of digital image analysis in improving the quality control process in histopathology. The significant difference between the tissue areas on the block and the slide underscores the need for an objective method to evaluate section completeness. The high sensitivity and specificity of the DigislideQC score suggest that it can be a reliable tool for identifying slides that require recuts, thereby reducing the time and effort required by histopathologists to screen slides.

Incorporating Digiblock and Digislide images into routine histopathology workflows can streamline the quality control process. The DigislideQC score provides a quantitative measure that can be used to set thresholds for automatic recuts, reducing the subjectivity and variability associated with manual assessment. Additionally, this method can be integrated with laboratory information systems to provide real-time feedback to histotechnicians and pathologists.

The implications of this study extend beyond traditional histopathology workflows. With the advent of telepathology and digital pathology, where slides are scanned and reported remotely, an objective method to assess section completeness becomes even more critical. The DigislideQC score can be incorporated into these digital workflows, providing pathologists with valuable information about the quality of sections and potentially reducing the turnaround time for diagnosis.

The study also opens avenues for further research and development. For instance, automated systems can be developed to capture and analyze images of tissue blocks and slides in bulk, minimizing the need for manual intervention. Advanced technologies like 3D imaging and photogrammetric methods can be explored to enhance the accuracy of tissue area measurements. Custom-designed tissue cassettes and slides with pre-printed scales can standardize the calibration process, further improving the reliability of the DigislideQC score.

Conclusions

The study demonstrates that digital image analysis, specifically the calculation of the DigislideQC score, can effectively enhance the quality control process in histopathology. By providing a quantitative measure of section completeness, this method can identify slides requiring recuts with high sensitivity and specificity, reducing the workload for histopathologists and improving overall efficiency. The integration of Digiblock and Digislide images into routine workflows and laboratory information systems can streamline the quality control process, making it more objective and reliable. This technology has the potential to significantly benefit future digital pathology and telepathology workflows, ultimately contributing to better patient care through more accurate and timely diagnoses.

Impact of slide scanner calibration on perceived color variation

Peter Ouillette[1] pouillet@med.umich.edu, Jonas De Vylder,[2] Bart Diricx,[2] Johan Rostang,[2] Tom Kimpe,[2] Louise Collins,[3] Richard Salmon,[3] Mustafa Yousif[1]

[1]Michigan Medicine Department of Pathology, Ann Arbor, MI, United States

[2]Barco, NV, Barco Healthcare, Kortrijk, Belgium

[3]PathQA, London, United Kingdom

Email: pouillet@med.umich.edu

Background

Whole slide images (WSI) can exhibit color variations due to differences in scanners. The amount of color variation between scanner brands and scanner types is impacted by the stains used in the WSI. Calibrating scanners with calibration slides can help reduce these differences by generating color transforms that standardize all slides to a consistent color representation, regardless of the scanner. These transforms are stored in ICC (International Color Consortium) profiles that can be applied to the original WSIs without duplicating the pixel data. High-end digital slide viewers and image analysis software typically support ICC profiles. This study quantifies the effectiveness of such calibration methods in mitigating perceived color variations introduced by different scanners and the extent to which the calibration's effect differs among various stains.

Methods

248 whole slide images (WSI) were used to assess color variability, encompassing 62 unique renal biopsy slides, each scanned by four different scanners. Renal biopsies were chosen due to their extensive analytical potential, provided by a variety of stains that span a broader color spectrum compared to classical hematoxylin and eosin (H&E) staining alone. The study included 10 different stains to comprehensively evaluate stain-induced color variability. Among the scanners, three were from the same manufacturer, with two being identical models. Thirteen slides were excluded from further statistical analysis due to issues related to automatic tissue detection or blur.

All four scanners were calibrated using a commercially available calibration slide designed for pathology. The calibration slide features 55 color patches simulating the spectral absorption of commonly used pathology stains. The corresponding ICC profiles were generated by the software provided with the calibration slide.

Tissue pixels were identified using Otsu Thresholding. Background pixels were excluded from the color measurements. All WSIs were aligned to the scan acquired with the reference scanner using an automated image processing method. This alignment step results in spatially consistent slides, i.e., pixel in a scan under evaluation corresponds to the same spatial location in the corresponding reference scan. Per pixel pair, the perceived color difference is expressed using the Delta-E (CIE 2000) metric (ΔE). All color measurements were performed both on the unprocessed raw WSIs and calibrated WSIs with the applied ICC profiles from the four scanners.

Results

Consecutive scans of the same slide can introduce small artifacts such as local blur, fingerprints or dirt introduced between scans, or minor tissue displacement, especially at the tissue boundary. To reduce the impact of corrupted pixel pairs due to such artifacts, statistical metrics robust to outliers have been used, e.g., median and median absolute deviation.

Same model scanners

No significant impact was observed between scanners of the same model. The raw images from these scanners were already similar, with color variations being barely noticeable. These scanners had similar age and usage patterns. An evaluation of how scanners might drift over time was beyond the scope of this study.

Same brand, different model scanners

Scanners of the same brand, but different models, exhibited significant color variation, especially for uncalibrated slides. Even after ICC color correction the color variation remained higher than for other scanners. This variation might result from advancements in scanner technology or scanner aging; the reference scanner had been operational for 2 years, while the other scanner had been in use for 9 years. However, color calibration significantly reduced color variation, from ΔE 10.81 (± 3.91) to ΔE 4.21 (± 1.51). The extent of color variation depended heavily on the stain used. For instance, C4D and IGG4 stains showed a ΔE below 5 in both raw and calibrated slides, whereas PASME showed a ΔE 16.61 (±4.39) in uncalibrated slides. Calibration benefited stains with strong color variation, such as PASME, where variation dropped to ΔE 4.98 (±1.92).

Different brand scanners

Scanners from different brands showed significant color variation on uncalibrated slides across all stains (ΔE 12.02 (±2.31). ICC calibration significantly reduced this variation to only 30 % of its original value, with multiple stains dropping to levels that were barely noticeable.

Conclusion

WSI scanners from different scanner brands and models capture slides with slightly different color appearances. Calibrating scanners with specific calibration slides can standardize color representations across different scanners, significantly reducing color variations. This study assessed 248 renal biopsy slides, scanned by four different scanners using 10 stains. Calibration notably decreased these variations, especially for stains with initially high variability.

Application of preexisting deep learning cancer subtyping algorithms to multiphoton microscopy

Charles Robbins[1], Faisal Mahmood[2], Kelsey Moore[3], Sudhir Perincheri[1], Richard Torres[4]

[1]Yale School of Medicine, Department of Pathology, New Haven, United States

[2]Harvard Medical School, Division of Computational Pathology at Brigham and Women's, Boston, United States

[3]Applikate Technologies, Inc., Director of Market Development, Fairfield, CT, United States

[4]Applikate Technologies, Inc., CEO Applikate Technologies, Inc., Fairfield, CT, United States

Email: jack.robbins@yale.edu

Background

Deep learning algorithms applied to digitized images of physical tissue slides can enhance the precision and accuracy of neoplastic histology interpretations. The clustering-constrained-attention multiple-instance learning (CLAM) method effectively pinpoints diagnostic sub-regions on whole slides for accurate classification. However, these algorithms often underperform when applied to datasets that differ from the training set and are sensitive to the variability inherent in physical slide preparations. Employing ultrafast multiphoton microscopy offers benefits to workflow efficiencies and image quality compared to traditional slide preparation by eliminating embedding and sectioning artifacts and generating high-resolution images comparable to whole slide imaging (WSI). However, few existing algorithms previously trained on physical slides have been applied to multiphoton datasets. Here, we assess the efficacy of CLAM algorithms in classifying renal cell carcinoma (RCC) and lung cancer (LC) on pseudo-colored multiphoton WSI images of a clinical sample cohort.

Methods

Clinical RCC and LC surgical samples were processed and imaged with Clearing Histology with MultiPhoton microscopy (CHiMP, Applikate Technologies, Fairfield, CT). This technique produces digital images of intact tissues that resemble H&*E*-stained optical slices. The multiphoton images were downscaled to a resolution of 0.5 μm/pixel to align with the resolution required by the CLAM models. These models were previously trained on TCGA and CPTAC datasets using WSI from physical slides for subtyping RCC (chromophobe, clear cell, papillary) and LC (squamous cell, adenocarcinoma) and were applied directly to CHiMP multiphoton images without any adjustments. Classifications with those from physical and digital slides were compared to validate the subtyping.

Results

The CLAM models were effective in subtyping the RCC and LC categories included in the training dataset when applied to multiphoton WSIs, achieving high prediction accuracy without the need for stain normalization or network modifications. However, subtypes not included in the training set, such as oncocytoma for RCC, showed low prediction scores (below 0.85), underlining the models' specificity. Analyzing images at multiple levels of slide depth illustrate the relevance of heterogeneity for machine learning classification.

Conclusion

Preliminary findings suggest that CLAM models originally trained on standard H&E WSIs successfully translate to pseudo-H&E multiphoton WSIs without requiring any domain-specific adaptations. This indicates that both the CHiMP multiphoton data is highly analogous to H&E WSI and that these models have effectively learned and can recognize diagnostic histological features in digital images produced.

Festival of dermpath: A more accessible conference platform

Charles Herndon[1], Jonhan Ho[2], Michael Cardis,[3] Timothy McCalmont[4]

[1]Touro University Nevada, College of Osteopathic Medicine, Henderson, United States

[2]University of Pittsburgh, School of Medicine, Pittsburgh, PA, United States

[3]Georgetown University Hospital, Department of Dermatology, Washington DC, DC, United States

[4]University of California, San Francisco, California, Department of Pathology, Walnut Creek, CA, United States

Email: herndonc@g.ucla.edu

Attending conferences is an irreplaceable opportunity for practicing pathologists to learn from colleagues and experts in the field. Despite advances in digital pathology, however, digital content is either absent or difficult to access before or after the event. We present a free, novel quarterly event named the Festival of Dermpath on the KiKo platform in

which contributors from around the world share highly interesting cases with each other to further each other's knowledge. The event occurs in a private group of **615** medical students, trainees, and pathologists every three months on a set date, and contributors can post text, links, static images, videos, and digital slides for others to peruse. In the most recent events, members attend a live Zoom meeting where faculty discuss selected cases in depth. Members can "attend" the festival in both a synchronous and/or asynchronous manner, allowing them to consume the content at their own convenience. Finally, members always have access to prior events. In the original Festival occurring in **September** 2019, **4** contributors shared **35** cases viewed **1314** times. Since then, **14** Festival of Dermpaths have been held, where **845** cases have been shared with **40,999** views from **57** countries. In conclusion, the Festival of Dermpath has enabled pathologists to collectively see and learn from each other's interesting cases on their own time and at no cost.

Quantitative multimodal anisotropic imaging for fibrosis assessment in usual interstitial pneumonia

Shuo Niu[1], Haodong Xu,[1] Xiaoxian Zhao,[1] Eric Hsi,[1] Kelly M Credille,[2] Aaron Gruver,[2] Aya Ryuzoji,[2] Nhat Le,[3] Guillaume Chhor,[3] Jun Zhang,[3] Jacqueline Brosnan-Cashman,[3] John Paul Hernandez Alcala,[3] Julia Varao,[3] Veronica Rivera,[3] Robert Egger,[3] Limin Yu,[3] Samuel Vilchez[3]

[1]Wake Forest University School of Medicine, Pathology, Winston-Salem, NC, United States

[2]Eli Lilly and Company, Diagnostic and Experimental Pathology, Indianapolis, IN, United States

[3]PathAI, Biomedical Data Science, Boston, United States

Email: sniu@wakehealth.edu

Introduction

Usual interstitial pneumonia (UIP) is characterized by histologic patterns that are heterogeneous and reflect the progressive nature of the disease, including honeycomb change, fibroblast foci, and paraseptal and subpleural fibrosis, alternating with areas of lung parenchyma that are normal in appearance. While significant progress has been made towards understanding the extracellular matrix composition of UIP, quantitative characterizations of this intricate histologic pattern remain largely unexplored. Quantitative multimodal anisotropy imaging (QMAI) is a novel imaging modality that allows individual collagen fibers to be visualized from a hematoxylin and eosin (H&E)-stained whole slide image, resulting in the ability to extract and compute features of collagen fibers from these images. To better understand patterns of fibrosis in UIP, we utilized QMAI to compare fibrosis features between UIP and normal lung tissue.

Methods

H&E-stained glass slides of samples demonstrating an advanced UIP pattern of injury ($N = 20$) and normal lung tissues ($N = 5$) as controls were imaged using QMAI and registered to H&E WSI of the same slides. Single fibers were extracted from the QMAI intensity heatmap, and fiber area, length, and tortuosity were quantified for each fiber. Slide-level QMAI outputs were compared among diseased and control samples. WSI from adjacent sections stained with Masson's trichrome were used for visual assessment of fibrosis in each sample. Two sets of regions of interest (ROI) were defined using pathologist-directed annotations of relevant tissue substances on H&E WSIs: 1) pulmonary arterial structures ($N = 430$), bronchiolar structures ($N = 267$), and alveolar spaces ($N = 85$) were annotated in both normal and UIP WSIs, and 2) dense collagen deposition ($N = 458$), honeycombing ($N = 74$), and fibroblastic foci ($N = 339$), representing diseased phenotypes were annotated only in the UIP WSIs. Collagen fiber-associated features (QMAI intensity, fiber density, mean fiber area, length, tortuosity, width) were extracted and compared at both the whole slide level and ROI level.

Results

QMAI features were compared between UIP and normal specimens at the slide level. Significant enrichment of two features – overall QMAI signal intensity and density of collagen fibers – was observed in UIP samples ($p = 0.007$ and $p = 0.02$, respectively). Fiber width, length, area and tortuosity were also compared between UIP and normal, but no significant differences were observed ($p = 0.2$, 0.5, 0.2 and 0.2, respectively). QMAI features were also compared in annotated ROI of relevant tissue substances. QMAI signal intensity and fiber density were both significantly elevated in regions of dense collagen deposition in UIP ($p < 0.0001$ for both; Kruskal-Wallis test) compared to other annotated ROI. Additional QMAI features, including fiber length ($p = 0.03$), fiber width ($p < 0.0001$), fiber area ($p < 0.0001$) and fiber tortuosity ($p < 0.0001$), were observed to significantly vary between ROIs in UIP samples. Notably, fiber length, width, area, and tortuosity are higher in UIP dense collagen deposition than fibroblast foci in UIP ($p < 0.05$; Dunn's test), consistent with observations from manual assessment of UIP histopathology. No significant differences in collagen features were observed between pulmonary arterial structures, bronchiolar structures, or unaffected alveolar spaces between diseased and normal samples.

Summary

Fibrosis has long been appreciated as a key component of UIP. However, the ability to characterize this histologic pattern in a quantifiable manner has not previously been feasible in routine H&*E*-stained specimens. Here, we demonstrate the utility of QMAI for the quantitative analysis of fibrosis features in UIP. Notably, QMAI allows the extraction of collagen features that cannot be obtained from manual assessment of WSI stained with H&E or other special stains alone. Features extracted from QMAI are able to be analyzed both at the slide level and in specific pathologist-annotated ROI, illustrating the ability of QMAI to reveal relevant collagen-related features at both a global and local level, allowing a granular understanding of the contribution of fibrosis to UIP. As such, quantification of QMAI features reveal specific alterations in UIP: in diseased cases, QMAI signal intensity and collagen fiber density were both enriched in UIP at the slide-level, as well as in certain UIP-specific ROIs. Additional work is needed to further understand how digital pathology approaches, such as QMAI, may be applied to better understand the spectrum of fibrotic lung disease, to differentiate between neoplastic and idiopathic fibrogenesis, and to provide insight for pharmacodynamic models of fibrosis progression.

Quantitative modeling of maternal inflammatory response in placental membranes

Teresa Chou, Karolina Senkow, Megan Nguyen, Payal Patel, Kirtana Sandepudi, Lee Cooper, Jeffery Goldstein

Northwestern University, Pathology, Chicago, IL, United States

Email: teresa.chou@northwestern.edu

Problem

The placental membranes are a key barrier to fetal and uterine infection. Inflammation of the membranes, diagnosed as maternal inflammatory response (MIR) or alternatively as acute chorioamnionitis, is associated with adverse maternal-fetal outcomes. MIR is staged 1-3, with higher stages indicating more hazardous inflammation. However, the diagnosis relies upon subjective evaluation and has not been deeply characterized. The goal of this work is to develop a cell classifier for 8 placental membrane cells and quantitatively characterize MIR1-2.

Method of Study

H&E-stained placental membrane slides were digitized. A convolutional neural network was trained on a dataset of hand-annotated and machine

learning-identified cells. Overall cell class-level metrics were calculated. The model was applied to 20 control, 20 MIR1, and 23 MIR2 placental membrane cases. MIR cell composition and neutrophil distribution were assessed via density and Ripley's cross K-function. Clinical data were compared to neutrophil density and distribution.

Results

The classification model achieved a test-set accuracy of 0.845, with high precision and recall for amniocytes, decidual cells, endothelial cells, and trophoblasts. Using this model to classify 53,073 cells from healthy and MIR1-2 placental membranes, we found that 1) MIR1-2 have higher neutrophil density and fewer decidual cells and trophoblasts, 2) Neutrophils colocalize heavily around decidual cells in healthy placental membranes and around trophoblasts in MIR1, 3) Neutrophil density impacts distribution in MIR, and 4) Neutrophil metrics correlate with features of clinical chorioamnionitis.

Conclusions

This paper introduces cell classification into the placental membranes and quantifies cell composition and neutrophil spatial distributions in MIR.

Key Words

placenta, placental membrane, chorioamnionitis, clinical chorioamnionitis, histologic chorioamnionitis, maternal inflammatory response, cell classification, machine learning, pregnancy, inflammation

The critical role of standards for AI in digital and computational pathology

Chhavi Chauhan[1], Joe Lennerz,[2] Joachim Schmid[3]
[1]American Society for Investigative Pathology, MD, United States
[2]BostonGene, MA, United States
[3]Illumina, CA, United States
Email: chhavich@gmail.com

ABSTRACT

As the field of digital pathology embraces artificial intelligence (AI), two key challenges remain to be fully addressed: interoperability and the lack of standards. A few members of the Digital Pathology Association's Education Committee along with key opinion leaders embarked on an arduous journey to highlight the importance of missing standards and propose a set of recommendations to advance our field further. This presentation will highlight the proposed recommendations to realize computational advances and enable integration of AI solutions to improve patient care.

Session description

This presentation aims to raise awareness regarding the critical roles of standards in digital and computational pathology, share recommendations that will enable the efficient use of artificial intelligence (AI) tools in digital pathology, and their impact on improved patient care.

Specifically, the presentation will explore data standards in the context of digital pathology, focus on the use of clinical data for research and development as well as integration of AI-products into clinical workflows, explore interoperability challenges with specific examples, and review the existing standards.

In addition, the experts will discuss emerging trends in AI that may require standards in our industry going forward.

Attendees will benefit from an improved understanding of standards to integrate them in their own research and development, clinical practice, and as they approach the emerging global regulatory needs for their AI tools.

How to A.C.E unanticipated events encountered during Telecytology Rapid on-site evaluation

Handy Oen, Oscar Lin, David Kim
Memorial Sloan Kettering Cancer Center, Department of Pathology and Laboratory Medicine, New York, NY, United States
Email: oenh@mskcc.org

Background

Rapid on-site evaluation (ROSE) with cytology preparations plays a critical role in minimally invasive procedures. Studies have consistently shown that ROSE improves quality care and reduces healthcare costs by decreasing non-diagnostic specimens, unnecessary passes, and repeat procedures. Since 2014, our institution's regional sites have employed robotic microscopes to facilitate Telecytology ROSE in the absence of on-site pathologists or cytotechnologists. Non-pathology physicians or laboratory technicians can prepare and stain smears, then load the slides into this equipment. The robotic microscopes allow full control of the field of view, including magnification, focus, and selection of points of interest. For this study, we will discuss several unanticipated events encountered during Telecytology ROSE over the past few years, including operating system updates, network connectivity problems, and inactive hardware/software applications. We will explore recommendations on how to prevent or solve these events using the A.C.E. methodology (Assess the root cause, Coordinate with IT and vendor support, Eliminate any future barriers).

Material and methods

We reviewed all technical issues and events documented during clinical Telecytology ROSE at our institution's regional sites over the past three years. This review included email exchanges with both main campus and regional IT staff, vendor support communications, and any relevant entries in our QA troubleshooting activities log.

Discussions

In the past three years, we assessed 3734 regional Telecytology ROSE cases in our institution. We identified thirteen technical issues (six were network issues, four were hardware issues, and three software update glitches). While most were resolved within twelve hours using vendor instructions (IFU) or our internal Standard Operating Procedures (SOPs), a few unforeseen events presented greater complexity. These events involved a combination of desktop malfunctions, network connectivity problems, and/or equipment failures that couldn't be resolved with standard troubleshooting.

Conclusions

Telecytology ROSE offers significant benefits for minimally invasive procedures, but technical issues can disrupt workflow and potentially impact patient care. While most issues are resolved quickly using standard procedures, unforeseen events involving software incompatibility, hardware malfunctions, and network connectivity problems require a more comprehensive approach.

This study highlights the importance of clear communication between IT departments and vendors for effective issue resolution. Implementing an A.C.E. (Assess, Coordinate, Eliminate) approach can address these challenges:

- Assess: Identify the root cause of complex technical issues to prevent recurrence.
- Coordinate: Collaborate effectively with IT staff and vendors to ensure timely resolution.
- Eliminate: Implement preventative measures such as health check alerts for network connectivity and software compatibility checks before Windows updates to eliminate future barriers.

By addressing these technical challenges, institutions can ensure the continued success of Telecytology ROSE and the quality of patient care it provides.

References

1. Lin O. Telecytology for Rapid On-Site Evaluation: Current Status. J Am Soc Cytopathol. 2018 Jan-Feb;7(1):1-6. doi: 10.1016/j.jasc.2017.10.002. Epub 2017 Oct 12. PMID: 29546032; PMCID: PMC5846691.

2. Sirintrapun SJ, Rudomina D, Mazzella A, Feratovic R, Lin O. Successful Secure High-definition Streaming Telecytology for Remote Cytologic Evaluation. J Pathol Inform. 2017 Sep 7;8:33. doi: 10.4103/jpi.jpi_18_17. PMID: 28966833; PMCID: PMC5609353.

Tumor growth and aggressiveness in lung cancer: An analysis of pathology slides and CT imaging data

Saanie Sulley, Yan Zhou
Boston University, Pathology & Laboratory Medicine, Boston, MA, United States
Email: sulleys@bu.edu

Background

Accurate assessment of tumor growth and aggressiveness is critical for the effective management and treatment of lung cancer. This study leverages machine learning (ML) techniques to classify tumor growth patterns and predict the aggressiveness of lung cancer using data from the National Lung Screening Trial (NLST), which includes pathology slides, CT imaging data, and clinical information.

Methods

We integrated multiple datasets from the NLST, focusing on key features such as tumor characteristics (sct_ab_gwth, sct_ab_preexist, sct_ab_code), imaging quality (ctdxqual_breath, ctdxqual_motion), and clinical data (de_grade, clinical_stag). An SVM model was trained to classify tumor aggressiveness, defined by high-grade (de_grade $\geq$ 3) or advanced stage (clinical_stag $\geq$ 3). Model performance was evaluated using metrics such as precision, recall, F1 score, and ROC-AUC.

Results

The SVM model demonstrated strong performance with a high ROC-AUC score of 0.9605, indicating robust overall accuracy. The model achieved high precision (0.95) and recall (0.99) for aggressive tumors, though it showed lower recall (0.45) for non-aggressive tumors. Visualizations, including the confusion matrix and ROC curve, highlighted the model's effectiveness and areas for improvement.

Conclusion

This study illustrates the potential of ML in assessing tumor growth and aggressiveness in lung cancer. The high accuracy of the SVM model in predicting aggressive tumors underscores its utility in aiding personalized treatment planning. Future work should focus on improving data quality, enhancing feature engineering, and performing extensive model tuning to further enhance performance.

Keywords

Lung cancer, Machine learning, Tumor aggressiveness, Pathology slides, CT imaging, National lung screening trial, SVM model, Precision-recall curve, ROC curve.

Building a FAIR digital pathology repository for pre-clinical safety assessment: Tackling storage

Benjamin Freiberg
Genentech, gCS Computational Catalysts, South San Francisco, CA, United States
Email: freiberg.ben@gene.com

Pre-clinical safety assessment plays a critical role in the drug development process by observing the effects of drug candidates at multiple concentrations on organ systems of laboratory animals. A critical component of this assessment is based on histopathology whereby the organs can be interrogated at high magnification to look for multiple pathologies. Harmonization of terminologies and metadata is key to the useability of these data for primary review, peer-review and downstream analytics. FAIR data principles enable this harmonization of data and metadata and are at the center of the pre-clinical data strategy at Genentech. Data and metadata from multiple sources can be brought together to facilitate peer-review while simultaneously enabling advanced AI algorithms to be trained and executed on the same data without interfering in either workflow. Our cloud based ecosystem contains over 1.5 PB of raw WSI binaries from slide scanners and grows by ~300 TB annually - making this a very expensive repository to maintain. One way to decrease storage costs is to pursue compression methodologies that do not degrade important workflows. Here we present the second phase of our system implementation where we discuss the implementation of a novel image compression/decompression solution on system performance as measured by pathologists and AI scientists. Additionally, we demonstrate the short and long term cost savings of our solution and efforts to make this compressed format open to read.

Securing patient data in clinical pathology: Integrated deidentification workflow

Ercan Alp Serteli[1] ercan.serteli@virasoft.com.tr, Gizem Solmaz Yılmaz,[1] Sercan Çayır,[1] Yakup Budancamanak,[1] Samet Ayaltı,[1] Fatma Tokat,[2] Kerem Kayhan,[2] Ümit İnce,[2] Burak Uzel,[3] G. Evren Keles[1]
[1]Virasoft Corporation, New York, United States
[2]Acıbadem University, Department of Pathology, Istanbul, Türkiye
[3]Çamlık Hospital, Department of Internal Medicine, Istanbul, Türkiye
Email: ercan.serteli@virasoft.com.tr

Introduction

There is a need for deidentifaction of Whole Slide Images (WSIs) for protecting the sensitive health information of the patients. We therefore integrated a deidentification tool to our existing image management system (IMS). We used the WSI Anon library, which is compatible with multiple vendors and formats.

Materials and methods

We integrated a deidentification tool, the WSI Anon library, into our IMS solution used in clinical pathology laboratories. Users can select WSIs, monitor deidentification progress, and access the anonymized images all within the application. It supports multiple WSI formats and includes features for secure access and template-based file renaming to easily organize deidentified images. The supported formats include 3DHistech Mirax, Roche Ventana, Aperio SVS, Hamamatsu NDPI and Philips iSyntax files.

Results

The integrated deidentification solution was successfully implemented and validated. It succeeds in removing all protected health information from WSIs in an efficient manner, supporting WSIs from different vendors.

An iterative process of user trials and feedback resulted in a user-friendly interface which allows easy selection and tracking of WSIs in the deidentification process, and fits easily into existing laboratory workflows. Template-based renaming was shown to enable flexibility when it comes to varying institutional needs and regulatory changes. Finally, security validation of the system has proven that access rules strictly prevent limited users from viewing non-deidentified WSIs.

Conclusion

Our image management system integrated deidentification tool has successfully removed protected health information from all of the whole slide images. These WSIs can be consulted, or be used in research without compromising patient privacy.

Keywords

Whole slide images, Deidentification, Digital pathology, Patient privacy, Image management system, WSI anon library.

Deep learning for breast cancer: Semantic segmentation of invasive regions in whole slide images

Ahmed Imam Shah[1], Süha Berk Kukuk,[1] Fatma Tokat,[2] Berrak Gumuskaya,[3] Gizem Solmaz Yılmaz,[1] Eren Tekin,[1] Muhiddin Enes Yavuz,[1] Ercan Alp Serteli,[1] Sercan Çayır,[1] Samet Ayaltı,[1] Kerem Kayhan,[2] Ümit İnce,[2] Burak Uzel[4]

[1]Virasoft Corporation, New York, United States
[2]Acıbadem University, Department of Pathology, Istanbul, Türkiye
[3]Lokman Hekim University, Department of Pathology, Ankara, Türkiye
[4]Çamlık Hospital, Department of Internal Medicine, Istanbul, Türkiye
Email: ahmed.shah@virasoft.com.tr

Background

Other than skin cancer, breast cancer is the most common cancer in the world, and the second leading cause of cancer related mortality in women. The histologic grade of breast cancer is an important factor in determining the aggressiveness of breast cancer, prognosis, and survival of the patient. The Nottingham Grading System is widely used by pathologists for histologic grading of breast cancer. This system assigns a score of 1 to 3 for each parameter such as tubular formation, nuclear pleomorphism, and mitosis. The evaluation of each parameter should only be performed in the invasive regions of whole slide images (WSIs). However, it can be difficult to precisely differentiate between invasive and non-invasive regions within breast cancer tissue. Recent scholarly investigations have unveiled the potential of deep learning algorithms when applied to the analysis of WSI for the accurate identification and grading of various cancer types. In this study, we propose and evaluate a system aimed at semantic segmentation of the invasive regions in the WSIs from the breast cancer tissue.

Materials and methods

The proposed methodology involves a two-stage model. In the first stage, a multi-branch, multi-resolution, convolutional neural network (CNN) based classifier detects the patches with invasive regions from the WSIs. The second stage of the model uses the positive patches from the first stage to semantically segment invasive regions. Hematoxylin and eosin histopathological whole slide images for breast cancer from Breast Carcinoma Subtyping (BRACS) dataset are used for training and validation of the classifier. The dataset is divided into two classes: invasive carcinoma (IC) and others. The model was trained on 395 slides (100 invasive and 295 others), and 65 slides are used for validation (12 invasive and 53 others). The Breast Cancer Semantic Segmentation (BCSS) dataset is used for training and validation of the semantic segmentation model. The dataset contains over 20,000 segmentation annotations of tissue regions from The Cancer Genome Atlas (TCGA). The dataset is annotated at pixel-level into two classes: tumor and others. The segmentation model was trained on 121 slides, and 30 slides are used for validation. Lastly, the end-to-end model was tested using 15 test slides.

Results

Our proposed deep learning model achieved remarkable results in classifying and semantically segmenting invasive carcinoma in breast cancer WSIs. Notable metrics include a sensitivity of 87.84 %, demonstrating the model's effectiveness in identifying invasive regions, and a specificity of 97.45 %, showcasing accurate identification of non-invasive areas. The precision reached 82.47 %, indicating the model's reliability in positive predictions, while the negative predictive value was 98.32 %, highlighting accuracy in correctly identifying non-invasive areas. Overall accuracy stood at 96.29 %, and the F1 weighted average was 85.07 %, emphasizing a balanced performance between precision and sensitivity. The Intersection over Union (IoU) for semantic segmentation was 74.01 %, indicating a substantial overlap between predicted and ground truth segmentation masks.

Conclusion

Our two-stage deep learning model demonstrates good performance in breast cancer WSI classification and segmentation. These results suggest its potential as a valuable tool for automatically finding invasive regions for breast cancer. As a next step, we are planning to run our Nottingham Histologic Grading algorithm within the regions that are outlined by this algorithm.

Evaluating the robustness and functionality of an open, end-to-end digital pathology solution

Jake Eden[1], Alan Byrne,[1] Rob Turner,[2] Callum Arthurs,[2] Luca Caneparo,[3] Stephan Fromme,[3] Arun Ananth,[3] Martin Kristensson[4]

[1]Agilent Technologies, Denmark
[2]Hamamatsu, United Kingdom
[3]Proscia Inc., Netherlands
[4]Visiopharm A/S, Denmark
Email: jake.eden@agilent.com

Background

The rise in cancer incidence and the steep increase in testing complexity is putting pathology laboratories under pressure. Along with a decreasing number of pathologists many laboratories are experiencing a demand for a change in work routine. Digital technologies are a proven way to overcome challenges in resources and complexity and pathology is beginning to transform into a digital process. In addition to alleviating common issues, providing better opportunities for remote work, and easy and faster peer review of cases, digital pathology improves efficiency and may also help attract new pathologists. However, proven digital pathology end-to-end workflows from staining to AI still need to be validated for scaled usage.

Aim

This study aims to substantiate the readiness of a multi-vendor, open, agnostic Digital Pathology Solution for scaled use in today's digital pathology laboratories. Specifically, it intends to demonstrate its capability to consistently manage the typical output of a high-volume immunohistochemistry (IHC) laboratory within an appropriate processing time window, across multiple runs.

Materials and methods

The study utilized two hundred coverslipped glass slides stained on Dako Omnis and Dako CoverStainer, representing the biomarker ratio

and daily volume typical of a high-volume workflow. These slides were stained within a week prior to the study. Anonymized patient profile metadata was generated for each unique barcode, including details such as name, accession date, specimen type, date of birth, case ID, and medical record number (MRN).

The study was conducted in the Dako training site in Copenhagen. The digital pathology setup consisted of the Hamamatsu NanoZoomer S360MD, Proscia's Concentriq pathology platform and Visiopharm Applications. Both software components were set up in a cloud environment, and utilized the seamless integration between Concentriq and Visiopharm AI applications.

Stained slides were loaded into the Hamamatsu NanoZoomer S360MD racks and scanned in Auto Mode (40×) with predefined scan profiles. Performance metrics and quality control were reported through Hamamatsu software. Whole slide images (WSIs) were automatically ingested into Concentriq, with patient data assigned to each image by barcode. Functional and timing assessments were performed within the software environment. WSIs were analyzed by the appropriate Visiopharm applications for each slide type, as determined by the metadata, and then returned to Concentriq for viewing with embedded AI results. To test the repeatability and robustness of the solution, the protocol was repeated three times over three days, with an overnight washout period using the same slides.

Results

The study confirmed that all of the functionalities were performed as expected. All slides were successfully scanned and processed through the complete digital pathology solution. Scanning with the Hamamatsu NanoZoomer S360MD was consistent over the three days with an average scanning time per slide of 40 s on each day. During the three days there was a 100 % success rate of scanning slides. All slides were handled correctly, zero scanner errors were reported and no maintenance was performed on the scanners in between runs. All tissue was detected on all scans fully automated, i.e. no human interaction was needed to amend any scan settings during the batch.

Correct patient data was assigned to each slide automatically by barcode upon ingestion into the IMS, and the appropriate AI applications were run on each slide. Visualization of AI overlays were accurately presented in the IMS through cross-vendor integration. All IMS functionalities including viewing, panning, annotations collaboration performed as expected.

The end-to-end workflow, comprised of scanning, data assignment, AI analysis, and final viewing, was completed within the appropriate processing time window, consistently within the same working day on each run. These results demonstrate the system's efficiency and reliability, even under high-volume conditions.

Advancing breast cancer diagnosis: A deep learning approach to predicting HER2 FISH scores

Daniel Macaulay, Christopher Garcia, Mark Zarella, WENCHAO HAN, David McClintock, Thomas Tavolara

Mayo Clinic, Laboratory Medicine and Pathology, Rochester, MN, United States

Email: d.o.macaulay@gmail.com

Breast cancer, a prevalent malignancy affecting millions worldwide, requires precise biomarker testing and quantification for effective treatment selection. Human epidermal growth factor (HER2) is of particular importance in guiding therapeutic strategies, especially for patients that require targeted anti-HER2 therapy. With the advent of novel Antibody-Drug Conjugate (ADC) therapies such as Trastuzumab deruxtecan, attention has shifted towards identifying breast cancers with low levels of HER2 expression (HER2 low breast cancers) as there is evidence to show that this subgroup of breast carcinomas responds well to ADC therapy.

HER2 testing employs Immuno-Histochemistry (IHC) and In-Situ Hybridization (ISH) assays to quantify HER2 expression. Despite their widespread use, both methods present unique challenges. IHC is a semi-quantitative assay that is prone to significant intra and inter-pathologist variability. Fluorescent ISH is a more accurate test that is also more expensive and time intensive. These issues underscore the necessity for more reliable and cost effective HER2 testing methodologies.

We proposed a deep learning model aimed at reducing the reliance on reflex FISH tests. CLAM, a weakly-supervised attention model was modified and trained on over 5000 HER2 IHC images. This includes 592 cases upon which reflex FISH testing had been performed, and FISH scores were thus available. The goal was to predict the binary class output of FISH results (positive or negative) using only HER2 immunohistochemistry images. To achieve this, we trained 2 CLAM models; one to predict base HER2 scores from IHC images (i.e. 0, 1+, 2+, 3+), and the other was trained to predict FISH scores from equivocal cases (i.e. 2+). Images were divided into training, validation and test sets, in an 80/10/10 split with 10-fold Monte-Carlo cross-validation. Individual images were broken up and tiled in small (256 × 256) patches, and then trained on CLAM's attention-based architecture. Class predictions were done at the slide level.

With this approach we obtained a 91 % overall accuracy and ROC AUC of 98 % (SD ± 0.002) on our HER2 IHC score prediction model. The FISH score prediction model had an ROC AUC performance of 84 % (SD ± 0.07) with sensitivity = 0.37 (SD ± 0.13), and specificity = 0.96 (SD ± 0.03) on the test IHC image set. External validation analysis was done with consult cases obtained from several outside institutions. On these our model showed an overall performance accuracy of 91 %, with an ROC AUC of 98 %, for the base model.

Our AI model represents a significant advancement in overcoming the inherent subjectivity and variability of current HER2 scoring methods. However, it is important to acknowledge that the lower accuracy and sensitivity metrics of the FISH score prediction model makes it unsuitable for use exclusively. Thus, the need for reflex FISH testing for HER2 amplification in the setting of breast cancer diagnosis and therapy, cannot be completely ruled out. We do think however, that our model offers a unique cost benefit for resource-limited settings, where routine reflex FISH testing may be unavailable or prohibitive. With a reasonable high specificity, we believe negative score predictions can be ruled out for testing with a high level of confidence. This approach addresses variability, but also seeks to enhance the accuracy and efficiency of breast cancer diagnosis and treatment selection.

Storage challenges and solutions for digital pathology workflow: A single institution's perspective

William MacDonald, Aaron Chow, Anil Parwani, Wendy Frankel, Zaibo Li, David Kellough, Erin Palermini, Giovanni Lujan, Swati Satturwar

The Ohio State University Wexner Medical Center, Department of Pathology, Columbus, OH, United States

Email: william.macdonald@osumc.edu

Background

Current digital storage consists of local in-house physical hard drives for most digital pathology laboratories. With ever increasing volume of scanned slides and demand for digital pathology workflows, storage remains an important issue. We share our experience with storage issues associated with clinical digital pathology workflow.

Design

The usage of hard disc space for digital slide image storage was tracked. Cost for updates and upgrades were measured, and the feasibility of digital image storage was compared between local storage in an established server space in-house and cloud-based storage.

Results

As of May 2024, 3.978 million whole slide images were captured. These images represent cancer only cases at our institution since 2011, all cases

since 2018, and all consult cases since 2017. In total, these images amount to 5 petabytes of required storage. The number of archived cases and storage demands continues to grow daily, with approximately 2800 new whole slide images scanned daily accounting to 3.65 terabytes of new data each day. Due to our data storage needs, we contemplated moving from local in-house hard disk-based storage to a cloud-based platform. We looked at various factors such as cost, ease of access, data-backup issues. On-site storage allows for the control of physical storage media but requires regular maintenance and replacement of physical hard drives every 5 years and is very expensive. Additionally, storage expansion with physical hard drives is typically limited by physical server size or building infrastructure. In contrast, cloud-based storage methods offer easy and straightforward storage expansion at a fraction of the cost of on-site storage. However, cloud-based storage takes significantly longer to back-up and restore data. Therefore, any issues that requires back-up or restoration of data would significantly interfere with usability and system stability in day-to-day sign-out and access to digital slides. We have been working diligently with our cloud storage provider to minimize those potential setbacks and establish a storage model that fits our growing demands and provides efficient back up and a recall time that it is acceptable to all end users. Our results show that cloud-based storage would be more beneficial for larger digital pathology laboratories.

Conclusion

Archiving images from clinical cases requires a significant amount of digital storage and comes with a high price tag. With physical and logistical limitations of large server installations, large institutions able to work with their provider to address possible drawbacks would likely benefit from cloud-based storage.

A two-stage approach for breast cancer lymph node metastasis detection

Ahmed Imam Shah[1], Süha Berk Kukuk,[1] Fatma Tokat,[2] Berrak Gumuskaya,[3] Gizem Solmaz Yılmaz,[1] Eren Tekin,[1] Muhiddin Enes Yavuz,[1] Ercan Alp Serteli,[1] Sercan Çayır,[1] Samet Ayaltı,[1] Kerem Kayhan,[2] Ümit İnce,[2] G. Evren Keles,[1] Burak Uzel[4]

[1]Virasoft Corporation, New York, United States

[2]Acibadem University, Department of Pathology, Istanbul, Türkiye

[3]Lokman Hekim University, Department of Pathology, Ankara, Türkiye

[4]Çamlık Hospital, Department of Internal Medicine, Istanbul, Türkiye

Email: ahmed.shah@virasoft.com.tr

Background

Lymph node metastasis has a significant impact on treatment decisions and patient survival in breast cancer patients. Recent advances in computational pathology have facilitated the development of automated methods for detecting lymph node metastasis in breast cancer.

In this study, we applied two different machine learning techniques, Multiple Instance Learning (MIL) and Clustering-constrained Attention Multiple Instance Learning (CLAM), to breast cancer lymph node metastasis detection.

Materials and methods

We used three publicly available datasets to train and validate the algorithms. First, a tumor segmentation model was trained using the Breast Cancer Segmentation and Survival (BCSS) dataset to delineate tumor regions in pathological images. From these segmented regions, patches were extracted to capture information relevant to the tumor. The extracted patches were then fed into CLAM and the MIL model. CLAM was tailored to classify whole slide pathology images based on patch-level data, distinguishing between metastatic and non-metastatic regions. MIL is particularly well suited for tasks where only slide-level labels are available, making it suitable for weakly supervised learning scenarios common in computational pathology. Training and validation were performed using the Camelyon 17 and Camelyon 16 datasets, respectively. We used a two-stage algorithm for breast cancer lymph node metastasis classification.

Results

The BCSS dataset was utilized for training and validating a semantic segmentation model aimed at distinguishing between tumor and non-tumor tissue regions. This dataset includes over 20,000 pixel-level annotations sourced from The Cancer Genome Atlas (TCGA), classified into two categories: tumor and others. The model's training phase involved 121 histopathological slides, while its validation was conducted using 30 additional slides. To assess the model's performance comprehensively, 15 separate test slides were employed. Camelyon 17 included 500 samples with 182 positive (metastatic) and 318 negative (non-metastatic) cases. Camelyon 16 contained 270 samples, with 112 positive and 158 negative instances.

The evaluation of the algorithm's performance, performed on the Camelyon 16 validation set, resulted in 43 true positives, 80 true negatives, 0 false positives and 6 false negatives (one of them is macrometastasis and 5 of them are micrometastasis). It showed a sensitivity of 0.879 and a specificity of 1, meaning that the algorithm was able to correctly identify all true negative cases without any false positives. This specificity underscores the robustness of the algorithm in distinguishing non-metastatic lymph nodes from metastatic lymph nodes. In addition, the algorithm demonstrated a precision score of 1, indicating that all positive predictions made by the algorithm were true positives.

The overall performance of the algorithm was further validated by its F1 score of 0.935, a metric that balances precision and recall, highlighting its effectiveness in breast cancer lymph node metastasis classification.

Conclusion

In conclusion, the results of this study demonstrate the effectiveness of the proposed algorithm in breast cancer lymph node metastasis classification. Using patches from tumor segmentation with multiple instance learning techniques, the algorithm achieved high sensitivity, specificity, precision and F1 score when evaluated on the Camelyon 16 validation set. These results demonstrate the value of the model as a useful tool for pathologists to reliably and accurately identify lymph node metastases. The only limitation of this algorithm is that it may miss some micrometastases. In the future, further validation and integration into clinical practice could contribute to improved outcomes in breast cancer management.

Refining hotspot identification in KI67 proliferation index assessment: QuPath algorithm vs manual

Ruben Oganesyan[1] oganesyan.rv@gmail.com, Osman Yilmaz,[1] Yevgen Chornenkyy,[1] Kanchan Kantekure,[1] Yuho Ono,[1] Buthaina Hakamy,[1] Nandan Padmanabha,[1] Vikram Deshpande,[1] Raul Gonzalez,[2] Monika Vyas[1]

[1]Beth Israel Deaconess Medical Center, Harvard Medical School, Pathology, Boston, MA, United States

[2]Emory University School of Medicine, Pathology, Atlanta, United States

Email: oganesyan.rv@gmail.com

Background

Ki67 is a crucial biomarker used in the grading of well-differentiated neuroendocrine tumors (WD-NETs), providing essential prognostic information and guiding therapeutic decisions. Traditionally, the assessment of Ki67 proliferation index (PI) is performed by pathologists through manual counting of positive cells in hotspot areas, a process that is inherently subjective and thus prone to inter-observer variability, potentially affecting patient management. To address this issue, we developed an algorithm

designed to automatically determine Ki67 hotspots in WD-NETs. The algorithm then calculated the Ki67 PI in these areas, aiming to provide a more objective and reproducible quantification and tumor grading.

Methods

Hotspots for estimation of Ki67 PI of Gastroenteropancreatic (GEP) WD-NETs ($n = 20$) were manually assessed and annotated by three pathologists and compared with the algorithm's hotspot detection on whole slide images (WSI). QuPath software with a custom classification algorithm was used to perform automated Ki67 scoring on hotspot areas of WSI. The three pathologists were given the same slides to manually determine hotspots on WSI, replicating their routine practice. They were then asked to submit a camera captured image, which was then used to determine PI and grade by positive cell detection feature of QuPath. Fleiss' Kappa statistics was calculated to assess agreement reliability among pathologists. Friedman Test was used for the assessment of the area size variability and the differences in cell detection numbers. The overlaps of pathologist-identified hotspots with QuPath-selected areas were also evaluated.

Results

The Ki-67 index was translated to grade as determined by the current WHO classification (G1: <3 %, G2: 3–20 %, G3: 20 %). There was moderate agreement between pathologists (Fleiss' Kappa = 0.47) and fair agreement between pathologists and the QuPath algorithm (Fleiss' Kappa = 0.40). In 8/20 (40 %) cases, there was an overlap in the area selected by the pathologist and the software. In 65 % (13/20), all 3 pathologists assigned concordant grades, though in 4 of these cases, the algorithm disagreed and upgraded the tumor (G2 vs. G1). The grades assigned by pathologists were equal to or lower than those determined by image analysis; however, the difference was not statistically significant ($p = 0.07$). In 1 case, there was a clinically important discordance between pathologists, where only 2/3 pathologists and the algorithm assigned G3 to the tumor. There was a statistically significant difference in hotspot area size and the number of cells included in the hotspot analysis among pathologists ($p < 0.001$), whereas the algorithm always picked the area of a similar size and cell number.

Conclusion

Our study reveals significant variability among pathologists in manually assessing the Ki-67 proliferation index in well-differentiated neuroendocrine tumors due to subjective hotspot detection, affecting tumor grading and patient management. The difference in the size of hotspots marked by pathologists highlights lack of standardization in this process and explains the large inter-observer variability in the current practice. Adoption of automated, algorithm-driven hotspot detection may provide objectivity and help reduce discrepancies.

HOST-factor: A platform for highplex quantitative analysis of the tumor microenvironment

Jeni Caldara[2], Janusz Franco-Barraza,[1] Fabian Tobias Schneider,[2] Daniel Winkowski,[2] James Mansfield,[2] Edna Cukierman[1]
[1]Fox Chase Cancer Center, Marvin & Concetta Greenberg Pancreatic Cancer Institute, Philadelphia, PA, PA, United States
[2]Visiopharm, Research, Horsholm, Germany
Email: JCA@visiopharm.com

Solid tumor complexity extends beyond the genetic and functional landscapes of heterogenous cancer cells, encompassing the tumor microenvironment (TME). Elucidating the TME's complexity requires a comprehensive assessment of its cellular composition, functional states, and spatial distributions. We developed the Harmonic Output of Stromal Traits (HOST) to identify TME cells, and the HOST-Factor to quantify their functional states. The HOST-Factor is a numerical value that reflects the relative contribution of cancer-associated fibroblasts (CAFs) to tumor-suppressive or tumor-promoting functions.

Our workflow combines automated cycling highplex immunofluorescent microscopy with artificial intelligence (AI)-guided image analysis. This generates HOST-Factor values for each identified TME cell within selected regions of interest, providing spatial distribution data. The TME signature encompasses 15 immune cells and 14 CAF antibody-detected biomarkers.

We applied our workflow to ten human pancreatic cancer specimens, generating OME-TIFF output images. This cancer model was used due to its significant TME makeup. The 29 highplex AI-based digital image analysis was conducted using the Phenoplex™ workflow from Visiopharm. The workflow included deep-learning-based tissue morphologic and cellular segmentations, cellular phenotyping, and integration of spatial/location data. Biomarker subsets were visualized, and a user-trained algorithm was used for tissue segmentation. Nuclear segmentation was done using a pre-trained algorithm on a DAPI-labeled DNA channel. Cellular phenotyping was performed using thresholds based on visual assessment and confirmation of positivity. Spatial neighborhood plots and metrics, heatmaps and partitioned t-SNE plots were generated for the dataset for downstream analysis. Importantly, the workflow's visualization templates, pre-trained nuclear/cytoplasm segmentation tools, and neighborhood plots and metrics, are reusable and fully customizable for new datasets.

Using HOST-Factor values, we successfully classified cancer and TME cells, along with their functional states, and spatial distributions.

This AI-based computational approach and user-friendly workflow provides a simple and effective way to obtain single-cell information from multiplex immunofluorescence images, making spatial phenotyping of several cell populations in tissues more accessible to researchers, providing a fully amendable means for future clinical translation.

Color standardization in whole slide imaging: A method to reduce color variability

Richard Heale, Louise Collins, Richard Salmon
PathQA, Research and Development, London, United Kingdom
Email: richard.heale@pathqa.co.uk

Introduction

Whilst the uptake of digital microscopy and whole slide imaging has brought numerous advantages to the field of pathology, it has also increased the need for color management for the accuracy and reproducibility of color for consistent diagnosis. This requirement for reliable color imaging is likely to be compounded by the development of Artificial Intelligence (AI) algorithms, which can be more easily perturbed by uncontrolled digital artifacts. In the digital pathology workflow, one area where color variability could be introduced is within the deployment of Whole Slide Imaging (WSI) scanners, due to the use of different manufacturers as well as variance within the same scanner. The present study aimed to investigate variation between the different WSI scanners and the use of color standardization to reduce the color variability in addition to bringing the imaged color closer to the “real world” color.

Methods

Using data taken from a variety of scanner models, color standardization was achieved using 55 color patches that mimic stained tissue and a proprietary algorithm to generate an interpolated ICC profile to accurately adjust the color values according to their associated spectral values. Tone profiling was extracted from these values and compared before and after standardization and across multiple scanner models and manufacturers. In addition, the color gamuts from these scanners were compared to the ground-truth color before and after standardization, by applying the

generated ICC profile. WSI scanner profiles were also compared to the sRGB gamut commonly used by displays.

Results

Data demonstrated different WSI scanner manufacturers typically employ significantly different linear or curved tone profiles to their scanned data, as well introducing variability in the overall color gamut achieved by the different scanners. The color standardization method used reduced the differences between scanners in terms of the tone profile and color gamut. Additionally, there was a significant reduction in the difference between the color values collected by the scanner and their true spectral values and a reduction in difference between absolute color errors between scanners of the same and different models that would otherwise create human and AI-perceivable differences if using factory settings. Moreover, it was possible to use the data in the profile generated to identify scanners that were potentially in incorrect calibration states ahead of tissue imaging, meaning that a user could take valuable corrective action before it created an impact on the workflow. It was also observed that there is a difference between the color gamut of the WSI scanners and the sRGB gamut commonly used by display manufacturers.

Conclusions

It is possible to reduce the impact of WSI scanner imaging color variability within the digital pathology workflow using ground-truth color standardization. There is also an identified need for routine metrics to determine the ongoing quality of WSIs being produced. Furthermore, there is a disparity between the color data potentially collected by a scanner and the sRGB gamut commonly used by displays which requires further work to address. These discrepancies in how color is processed by the scanner and presented for analysis could create a confounding variable, highlighting the importance of solutions that are able to aid color consistency regardless of scanner manufacturer both at the points of image collection and presentation to the users via commercial or in-house viewer platforms. As the digital pathology field attempts to validate AI the differences in the presentation of color on a display compared to what is presented to AI are vital to be addressed as human validation and an AI to be validated should have the same data presented to them, for which ICC profiling provides the foundation.

Scalable image analysis workflow for high-throughput Ki67 quantification in colorectal polyp samples

Gabriel Reines March[1], Karen McClymont,[1] James Going,[1] Finley Pinkerton,[1] John Waller,[1, 2] Aula Ammar,[3] Noori Maka[4], Stephen McSorley,[3, 4] Gerard Lynch,[3] Joanne Edwards[3]
[1]OracleBio Ltd., Glasgow, UK
[2]Sciento Technology Ltd., Glasgow, UK
[3]University of Glasgow, School of Cancer Sciences, Glasgow, UK
[4]NHS Greater Glasgow and Clyde, Glasgow, United Kingdom
Email: gabriel.reinesmarch@oraclebio.com

The Integrated Technologies for Improved Polyp Surveillance (INCISE) project is a triple-helix collaboration between the University of Glasgow, NHS Greater Glasgow and Clyde and an array of industrial partners. This multidisciplinary effort is aimed at improving bowel cancer surveillance in the UK by developing a multi-omics prediction model to stratify patients based on their risk of developing additional polyps.

OracleBio, as the project's Digital Pathology work package lead, was in charge of developing an image analysis pipeline to quantify Ki67 biomarker expression on whole slide colorectal polyp tissue samples. The proposed workflow, built using Visiopharm AI, followed a top-down approach and was composed of four stages:

- Tissue detection and slide quality control: detection of viable tissue and removal of large stain and scanning artifacts at low magnification.
- Adenoma tissue classifier: classification of viable tissue into adenoma and non-adenoma regions.
- Epithelium tissue classifier: segmentation of epithelial glands from surrounding stroma.
- Cell detection and quantification: single cell segmentation and phenotyping at full magnification.

Pathologist input was used to validate the adenoma tissue classifier and cell detection algorithms by comparing ground truth annotations against classifier outputs.

The end-to-end workflow was composed of 8 Visiopharm apps and was run on $n = 2211$ Ki67 immunohistochemistry stained whole slide polyp samples. We leveraged our proprietary cloud-based, scalable IT infrastructure to launch 75 GPU-enabled processing nodes to analyze the images in parallel, with a turnaround time of 45 h to complete analysis on the full cohort. The successful execution of this workflow demonstrated that complex image analysis pipelines can be run in a robust and efficient manner for future potential support of diagnostic healthcare services.

Color calibration of pathology scanners benefits AI by removal of digital and temporal variation

Richard Salmon[1], Xiaoyi Ji,[2] Anders Blilie,[3] Louise Collins,[1] Kimmo Kartasalo[2]
[1]PathQA, R&D, London, United Kingdom
[2]Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Stockholm, Sweden
[3]Stavanger University Hospital, Department of Pathology, Stavanger, Norway
Email: rick.salmon@pathqa.co.uk

Background

The full potential of artificial intelligence (AI) in digital pathology is limited by technical inconsistencies in scanning of whole slide images (WSIs), posing a challenge for widespread clinical application as fine-tuning algorithms for each new site is impractical. Concerns have also been raised over 'AI aging' as scanners on which AI is initially validated change over time and use, impacting AI post-validation performance and potentially compromising patient safety. We evaluated whether physical color calibration of scanners can standardize WSI presentation to AI and remove temporal variation to enable robust and reliable AI.

Methods

To assess all scanner color performance we employed a color calibration technology that quantifies and corrects WSI to ground truth color of real glass slides using ICC profiles. The method consistency was confirmed across 18 scanners models and 55 real pathology colors. To assess impact of removing digital color variation, a calibration slide was scanned in four different laboratories with scanners of varying models and performance of an AI system for prostate cancer diagnosis was evaluated on 1161 WSIs. This was compared to computational normalization by cycleGANs and Macenko.

To assess impact of temporal color variation, 119 prostate core needle biopsy slides with a balanced ISUP grade distribution were scanned alongside a color calibration slide on a scanner every 14 days for one year. Any change in stained glass tissue color was baselined. The impact of scanner-induced temporal variance on AI model performance was evaluated on a deep learning model trained on over 46,000 WSI for prostate cancer diagnosis.

Results

Highly variable, inaccurate and inconsistent color across the scanner market was demonstrated to be standardized and corrected to the color of

real tissue using the intended method. Color standardization resulted in consistently improved prostate AI model calibration and significant improvements in Gleason grading performance from multi-scanner sources. Color standardization was also more reliable when AI had less test data than computational normalization methods. Full temporal data collection is still underway, however diagnostic performance metrics (sensitivity and specificity, Cohen's kappa) will be presented and pathologists will evaluate the slides with discrepant AI results to provide human insights into the potential causes of any variation.

Conclusions

The study demonstrates that physical color calibration provides a solution to the variation introduced by different scanners from both cross-vendor WSI quality and scanners drifting in performance, making AI-based cancer diagnostics more reliable and applicable in clinical settings. Maintenance of reliability by physical color standardization even when AI is provided with less test data makes AI more feasible for deployment in smaller clinics and research groups. This study pioneers a real-time quality assurance approach for ensuring stable and scalable performance of scanners and AI over time.

Artificial intelligence-based tool for detection of oral cancer

Garima Jain[1], Aarti Yadav,[2] Abhijeet Patil,[3] Amit Sethi[4]
[1]National Institute of Research in Digital Health & Data Scie, Pathology, ICMR-NIRDHDS, New Delhi, India, India
[2]ICMR-NIRDHDS, ICMR, New DElhi, India
[3]IIT, Bombay, Electronics and Communication, IIT, Bombay, India, India
[4]Indian Institute of Technology, Mumbai, Department of Electrical Engineering, Mumbai, India, India
Email: garima.j@icmr.gov.in

Introduction

Breast and oral cancer rank as the first and second most common cancers in India, respectively, with an annual incidence of 1.78 & 1.35 lakh new cases. Most cases of breast and oral cancer (70 %) present to healthcare facilities in rural settings and at advanced stages resulting in poor outcomes. Early diagnosis with the help of regular screening has demonstrated a decline in mortality rates of up to 34 % in oral cancer and 39 % in breast cancer. Digital pathology (DP) encompasses the process of transforming cytopathology slides into digital images via whole-slide scanners. DP has picked up pace and has evolved from being the center of research and development at medical/ technological universities, and healthcare startups to the edge of mainstream adoption in clinical practice in the last decade. Digitization of slides offers many benefits viz. decreased risk of case or patient misidentification, efficient case tracking, streamlined storage and retrieval of archival cases, better time per slide workflow efficiency, reduced risk of tissue loss or damage, etc. Machine learning (ML), a branch of AI has been extensively explored and employed for pathological predictions. Various models for primary tumor detection, detection of metastatic deposits, grading, subtyping, assessing tumor heterogeneity & receptor status, prognostication, and correlation of morphology with response to treatment have been developed for breast, prostate, colon cancers, and cervical and urinary cytology. With the scarcity of AI algorithms on cytopathology and that too not on Indian data, there is an urgent need to develop the same. A successful algorithm would assist cancer screening across the nation and help in early diagnosis thereby decreasing cancer mortality for these two malignancies.

Rationale

The utilization of AI algorithms aims to improve the classification accuracy of microscopic cytopathological images, distinguishing between benign and malignant categories, and ensuring alignment with the final histopathological diagnosis. Also, AI systems can be trained on vast datasets of annotated images, allowing them to recognize subtle patterns and features indicative of cancerous cells with a high level of precision. This capability enhances the diagnostic accuracy of pathologists, aiding in the early detection of cancerous lesions. Further, the implementation of AI tools facilitates early detection efforts for both new and suspected cases of oral and breast cancer. By rapidly analyzing large volumes of cytopathological images, AI algorithms can flag suspicious findings for further investigation, prompting timely interventions and reducing the risk of disease progression.

Objectives

As mentioned in the aforementioned section 6.

Methodology

As mentioned in the aforementioned section 7.

Results

As mentioned in the aforementioned section 9.

Outcomes and translational potential

The present study received a total of 1022 cases for breast FNAC and 755 cases for oral brush cytology. Analysis of demographic details revealed that the urban population was predominantly affected by oral and breast cancer. FNAC samples indicated a predominance of benign tumors, with 65 % classified under.

Bi-RADS category II and 25 % showing malignancy (category V), suggesting a prevalence of non-cancerous conditions. Also, lump development was predominantly observed in the upper outer quadrants (UOQ) of both breasts. The study also highlighted common habits among participants, with daily consumption rates of cigarettes/bidis (43 %), khaini (22 %), and tobacco (20 %) reported at approximately 10–15 times per day. Further, the developed model's training accuracy was evaluated using a validation dataset, revealing that the smallest model achieved the highest accuracy with an average precision of 0.80. In comparison, the medium and large models scored 0.78 and 0.72, respectively, thus indicating lower error rates. The translation potential of this project suggests that the successful development of an AI algorithm could pave the way for diagnostic and prognostic models for oral and breast cancer. These models could assist pathologists and clinicians in improving diagnostic workflows and ensuring error-free, accurate, and timely diagnoses for patients.

Digital cytopathology: Navigating the opportunities and challenges of the digital frontier

Poombal[1], Noor Shaker,[2] Nuha Shaker,[3] Nada Shaker[4]
[1]UMass Chan Medical School-Baystate Campus, Pathology, Springfield, MA, United States
[2]SpatialX, Pathology, London, United Kingdom
[3]University of Pittsburgh Medical Center, Pathology, Pittsburgh, United States
[4]University of California, San Francisco, CA, United States
Email: poombalsk@gmail.com

Background

Digital cytopathology revolutionized the way cellular specimens are analyzed and interpreted. The development of digital cytopathology has been fueled by a growing demand for more efficient and accurate diagnostic techniques in the face of increasing caseloads, the globalization of healthcare, and the need for remote consultation and collaboration. As healthcare systems strive to improve patient care while managing costs,

digital cytopathology offers a promising solution by enhancing workflow efficiency, facilitating interdisciplinary communication, and enabling the integration of advanced computational techniques into routine practice.

Aim

This study aims to investigate the potential advantages and challenges in the application of digital cytopathology.

Methods

We conducted a comprehensive literature review on digitalization within cytopathology to evaluate its impact and integration. The search on PubMed from 1990 to 2024 yielded 53 relevant articles to discuss. The search focused on relevant key terms such as "**Digital Cytopathology**," "**Virtual Microscopy**," "**Artificial Intelligence in Cytopathology**," "**Digital Cytopathology Atlases**," and "**Whole Slide Imaging**."

Results

The search highlighted a robust concordance rate of 98.7 % between digital and analog diagnoses, affirming the reliability of digital pathology. Despite these promising results, challenges persist with scan failure rates reaching 10 % due to technical issues such as pauci-cellular samples and smear thickness. These challenges highlight the need for enhancements in slide preparation and scanning technology.

The studies collectively demonstrated the potential of digital cytopathology to improve diagnostic accuracy, streamline workflow processes, enhance educational tools, and facilitate international collaborations. Digital cytopathology provides invaluable continuous learning resources, including interactive modules, virtual microscopy platforms, and extensive case repositories. Additionally, digital pathology facilitates remote collaboration and telecytopathology, enabling remote consultations, tumor board meetings, and real-time interaction among specialists conducting multidisciplinary meetings and consultations; this integration ultimately improves access to expertise and streamlines workflows. Moreover, quicker access to digital slides enhances efficiency and reduces the average diagnostic turnaround time by up to 40 %, significantly enhancing patient care. However, several challenges need to be addressed for successful implementation. These challenges include technical complexity, the need for consistent quality control, adaptation to digital workflows, resistance to change, and integration issues. Other challenges involve technological limitations, high initial investment, ongoing development needs, variability in image quality, data transmission issues, and dependency on digital infrastructure and reliable internet connectivity. Furthermore, there are concerns about the complexity of AI algorithms, ethical considerations, regulatory challenges, standardization across institutions, and the ongoing need for updates, maintenance, and professional training.

Pre-analytic and analytic factors such as inconsistent slide preparation, staining quality, and scanning technology can lead to significant image quality errors, necessitating careful optimization of these processes. Ensuring high concordance between digital and traditional glass slides is essential for maintaining diagnostic accuracy, requiring rigorous validation studies. Anderson et al. (2022) study highlighted the challenges in slide preparation and staining. A study by Johnson et al. (2019) study showed that 30 % of pathology staff resisted transitioning to digital systems. The median cost for a mid-sized pathology lab to transition to digital systems can exceed $500,000, including hardware, software, and training costs.

Conclusion

The digitization of cytological specimens offers numerous advantages over conventional microscopy. Digital slides can be easily archived, retrieved, and shared among pathologists and clinicians, facilitating remote consultations, collaborative discussions, and educational activities. Furthermore, digital pathology platforms can support the implementation of automated image analysis algorithms and artificial intelligence (AI) tools, which have the potential to enhance diagnostic accuracy, streamline workflow processes, and improve patient outcomes. However, addressing the potential challenges is crucial for the successful integration of digital cytopathology into routine clinical practice.

From data to discovery: PathologyMap™ and the integration of AI to accelerate pathology discoveries

Akash Parvatikar, Christopher Lee
Histowiz, Inc., Long Island City, NY, United States
Email: akash@histowiz.com

Background

The rapid growth of digital pathology has highlighted the need for robust quality control and advanced image analysis tools to ensure the reliability of histopathological data. HistoWiz's PathologyMap™ platform addresses these challenges by integrating state-of-the-art AI technologies with high-quality histology workflows. Trusted by over 3500 researchers from 550 organizations, PathologyMap™ has become a cornerstone in preclinical and clinical research, providing a comprehensive suite of tools for seamless tissue processing, data management, and AI-driven analysis.

Approach

The PathologyMap™ platform leverages a vast, curated database of over 350,000 digital slides annotated by experts. This extensive dataset fuels the development of AI models for automated quality control (Auto-QC), slide tagging, and content-based image retrieval. Our Auto-QC system identifies common quality issues such as blurriness, folds, tissue tearing, and air bubbles, achieving high accuracy (AUC 0.97) and significantly reducing manual oversight. Additionally, the integration of third-party AI applications enhances the platform's analytical capabilities, ensuring consistent excellence in slide preparation and enabling precise diagnostic and research outcomes.

Results

The implementation of AI-driven Auto-QC in PathologyMap™ has resulted in an 86 % reduction in quality control time compared to manual processes. Our deep learning models, trained on tens of thousands of annotated patches, demonstrate exceptional performance in detecting various quality defects. Furthermore, the platform's advanced features, including ultra-fast whole slide image viewing, dynamic group management, and comprehensive LIS integration, empower researchers to conduct collaborative and insightful data exploration with ease.

Conclusion

PathologyMap™ sets new standards in digital pathology by combining efficient sample processing, dynamic data management, and cutting-edge AI technologies. By providing a holistic, end-to-end solution, the platform not only enhances lab efficiency and quality control but also drives the future of histopathological analysis. Join us to explore how PathologyMap™ is transforming the landscape of digital pathology and accelerating discoveries in biomedical research.

Taking the temperature: A large academic center status post digital pathology integration

Matan Kadosh[1], Nhat-Phuong Nguyen,[2] Mehrvash Haghighi[3]
[1]Icahn School of Medicine at Mount Sinai (West) Progam, Pathology, New York, NY, United States
[2]Mount sinai West, Pathology, Mount Sinai West New York, NY, United States

[3]Mount Sinai, Pathology, Molecular&Cell Based Medicine - ISM, Mount Sinai/New York, United States
Email: matan.kadosh@mountsinai.org

Background

The integration of digital pathology ushers institutions' pathology workflows into the modern era. Few institutions have taken the plunge to promote a full digital pathology sign out. In our large academic institution, the image management system (IMS), Phillips Image Management System (Amsterdam, NL), is fully integrated with the laboratory information system (LIS) (Sunquest PowerPath, Tucson, AZ). Since the integration, the institution has troubleshooted various problems related to different factors. In this survey, we examined the pathologist community's attitudes towards digital pathology.

Methods

A survey was designed using Google Forms (Mountain View, California, USA). The survey was distributed via departmental email. The participant's responses were anonymous, and no linking information was recorded. The respondents could only submit their survey one time. The multiple choice questions directed the 113 participants to reflect on their attitudes towards the themes of efficiency, operability, and accuracy. The participants varied in status from resident to attending as well as years of experience, from 0 to 11 + years of pathology experience. The answer choices allowed the participants to answer questions based on frequency, ease of use, satisfaction, and favorability. All data collected was anonymous, with no name or email address information being tied back to the respondent.

Results

The majority of those who responded were residents, with a few fellow/faculty respondents. Of the 34 out of 113 participants, most were residents at 61.8 % and 32.4 % were attending pathologists. 63 % of respondents reported digital pathology being very easy to use and 77.7 % found the interface with the image viewer to be either intuitive or very intuitive to use. While most people found the remote use of the image viewer to be favorable or very favorable at 62.9 %, 17.6 % of respondents found the remote use of the image viewer to be unfavorable, very unfavorable, or sometimes works well. Interestingly, 14.7 % of the respondents never use the image viewer remotely.

The most surprising answers pertained to the quality of digital slides when compared to glass with 47.1 % of respondents reporting that the two methods are about the same. 2 people found using digital pathology made them less diagnostically confident when compared to glass slides. A concerning finding was that 41.2 % of respondents felt that retrieving archived slides was either difficult or very difficult. Most respondents, at 47.1 %, occasionally had technical issues with the image viewer.

When given the opportunity to freely respond, respondents commented on several issues, including lag time with slide navigation, and IMS and LIS integration. Overall, an overwhelming majority of respondents were satisfied or very satisfied with the image viewer at 91.2 %.

Conclusions

Digital pathology has the power to revolutionize traditional pathology workflows. Many institutions are now investing and feverishly developing both digital and physical infrastructure to further usher in their digital pathology integration. While digital pathology integration can be achieved, consensus has still not been reached about digital pathology satisfaction and acceptance.

The participants of this survey agreed that three major issues that plague digital pathology operability are retrieval of archived images, remote use, and technical issues. Our institution archives images 6 months after creation. Retrieval of archived images is useful when comparing morphology, from primary to recurrence of cancers. For frozen section slides, one can retrieve the images the day before to be familiar with the patient's tumor biology. Remote use can be challenging due to both the monitor used and the internet connection. The network connection should have enough bandwidth to handle the image loading and the monitor should be compliant with FDA approved requirements for monitors. It is also difficult to discern the technical issues that participants faced with the image viewer, however one participant did note that the system will “glitch when bringing up two slides at once”. This favors the notion that the image viewer still has issues yet to be addressed concerning data storage and image retrieval.

Since implementation of digital pathology at our institution in early 2020, many trainees and attendings have been trained in digital pathology practice. This survey intended to study the current attitudes towards our digital pathology climate. The respondents' answers to questions about usability, efficiency, and satisfaction show that while most are in favor of digital pathology integration, there are still several operational issues that require future troubleshooting efforts.