Abstract

Evaluation of an Automated Tissue Sectioning Machine for Digital Pathology

Xiujun Fu1, Pinky A. Bautista2, Veronica Klepeis2, Yukako Yagi1

1Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, 
2Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. E-mail: fux@makscc.org

Content: Automation and digital pathology are the trends for future anatomic pathology with the increasing workload in histology laboratories. While tissue processing, embedding, staining and coverslipping, and digitizing have been available for automated use, tissue sectioning appears to be the biggest roadblock to a fully automated histology process. In this study we were aimed to investigate a tissue automated sectioning machine for both clinical and research use. Technology: Tissue auto-sectioning machine AS-410 (Dainippon Seiki Co. Ltd., Japan) which has the abilities of tissue detection, barcode reading and printing, and 3-8 μm tissue preparation, was used by this study. Nanozoomer 2.0HT (Hamamatsu, Japan) scanner was used to acquire the whole slide images (WSI) of the H&E stained slides at a resolution of 0.46 μm/pixel. Design: Totally 77 surgical resection blocks of various organs embedded with standard paraffin were sectioned automatically using AS-410 at 5 μm with the default setting (Setting A). 10 slides per block were sectioned and the last 5 slides were stained with H&E, digitized with WSI scanner, and evaluated by image scientist and pathologist. The image scientist scored the images base on the extent of imperfection (Evaluation I), while the pathologist scored the images based on the clinical diagnosis purpose (Evaluation II). Both scoring systems were scored from 1 to 5, with 1 the worst quality and 5 the highest quality. Tissues with unsatisfied score were sectioned with modified setting (Setting B), and evaluated again by the same image scientist and pathologist with the same scoring systems. And the scores from the two different settings were compared. Auto-trimming and barcode reading and printing of AS-410 were also evaluated. Results: The AS-410 provided auto-trimming function to detect exposed tissue for cutting, accomplished by the installed camera and calculation software. It read sample information and printed barcode as well as input text and automatically generated slide order information. It produced good quality of sections for most cases with median score more than 4 in both Evaluation I and Evaluation II using setting A [Figure 1a and b]. The scores of the unsatisfied blocks sectioned with setting A improved significantly when sectioned with setting B [Figure 1c and d]. Conclusion: The AS-410 tissue sectioning machine

Figure 1: Evaluation scores given by imaging scientist (a and c) and pathologist (b and d)
produces high-quality sections with clinical standard paraffin tissue blocks of a variety of organs with proper settings. It promises high automation with sound sectioning quality in the era of digital pathology for both clinical and research use.

Notes: The affiliation of Xiujun Fu and Yukako Yagi was Massachusetts General Hospital when this research was performed.

### Color Image Segmentation Using Multi-level Thresholding: Applications for Computer-automated Ki-67/SOX-10 Indexing

Daniel Gonzalez, Jonhan Ho

1Pathology, University of Pittsburgh School of Medicine, Departments of 1Pathology and 2Dermatology, University of Pittsburgh, Pittsburgh, PA, USA. E-mail: dgonz14@gmail.com

**Content:** The Ki-67 labelling index has been shown to be a valuable prognostic indicator in various carcinomas including brain, breast, and skin tumors. However, adaptation into daily practice is met with challenges related to mode of assessment. “Eye-balling” is the least expensive and most widely used method but has poor reliability and reproducibility. Automated counting by image analyzers is the current gold-standard but has been shown to be expensive and impractical. In this study, we developed a custom computer-based algorithm to automatically calculate Ki-67 index in 3 Ki-67/SOX-10 dual-stained melanoma images with an emphasis on color-based image thresholding. **Technology:** An Omnyx VL120 Scanner was used for digital slide scanning. ImageJ Windows 64-bit was used for manual cell counting. The algorithm was developed and implemented using Matlab 2016b. **Design:** Ki-67/SOX-10 dual-stained slides high-grade melanomas were chosen for this study. Ki-67 is represented by red and SOX-10 by brown nuclear staining in this assay. Images were digitally uploaded using an Omnyx VL120 Scanner and magnified to 40x. Screenshots measuring approximately 1100x1700 pixels were downloaded for analysis. Cells were manually counted using ImageJ Win-64. The images were uploaded into Matlab 2016b for further automated Ki-67 indexing. The algorithm began with image upload and conversion from RGB to HSV color space. Multi-level thresholding was performed on the hue histogram to separate red/brown dual-stained (Ki-67/SOX-10 stained melanocytes) cells from brown only (SOX-10 stained melanocytes). Blob-analysis was then performed for noise reduction and a Ki-67 index was generated from the processed images. Indices calculated with the algorithm were compared to manual counts using a paired t-test. **Results:** For n=3 images we found no significant difference between the computer-automated index and a manual count (p = 0.06). **Conclusions:** Many image thresholding algorithms are based on image pixel intensity, relying on the continuity of gray-scale values for thresholds. These methods are not applicable for colored images in RGB format however, conversion to HSV places color values on a contiguous axis suitable for thresholding. Color thresholding can be a powerful tool in any application requiring quantification of colored bioassays.

### Proposed Criteria for Rapid On-site Evaluation Telecytology Validation: The University of Pittsburgh Medical Center Experience

Matthew G. Hanna, Jacqueline Cuda, Sara E. Monaco, Juan Xing, Ishtiaque Ahmed, Liron Pantanowitz

1University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: hannamg2@upmc.edu

**Content:** Telecytology is being increasingly used at many centers. The most common mode of telecytology employed for remote rapid on-site evaluation (ROSE) is real-time streaming. However, most guidelines for digital pathology validation were developed for whole slide imaging. Therefore, the aim of this study was to establish practical methods for telecytology validation using real-time live video microscopy. **Technology:** An Olympus DP71 camera attached to an Olympus BX51 microscope was used to stream 640x480 images over a 1 MB secure network. CellSens Standard 1.11 software was accessed via Citrix to remotely view streamed images. **Design:** Four cytopathologists prospectively and collectively evaluated 60 CT-guided fine needle aspiration ROSE cases first using glass slides and immediately thereafter by another cytopathologist via telecytology. Glass slide and remote digital ROSE interpretations were compared to each other and final cytology diagnoses. Difficulties experienced during the study period were recorded. **Results:** ROSE matched the final diagnoses in 98% of cases for both telecytology and glass slide microscopy [Table 1]. The single discrepant case called “negative for malignant cells” at ROSE with both glass slides and telecytology revealed malignant cells only in the cell block (sampling issue). However, there was 100% concordance for ROSE diagnoses rendered digitally vs. by glass slide. Technical difficulties were identified in 17 cases (28%), with the majority

| Table 1: Digital versus glass slide comparisons (n=60) |
|-------------------------------------------------------|
| **ROSE comparison** | **Concordance, n (%)** |
| Digital versus glass slide on-site interpretation | 60 (100) |
| Digital versus glass slide final diagnosis | 59 (98) |
| ROSE: Rapid on-site evaluation | |
We receive a surgical resection of a pancreatic mass that is suspected of harboring malignancy. We receive a consultation request from an outside institution for this resection. We receive a biopsy from a colonic mass in a patient undergoing colonoscopy. We receive a 3-part biopsy case from a screening colonoscopy. We receive a percutaneous liver biopsy upon referral of a patient with possible liver metastasis.

Workflow description

We receive a consultation request from an outside institution for a liver biopsy. We receive a biopsy from a colonic mass in a patient undergoing colonoscopy. We receive a 3-part biopsy case from a screening colonoscopy. We receive a percutaneous liver biopsy upon referral of a patient with possible liver metastasis.

Standardized Playscripts for Digital Pathology Vendor Demonstrations

Matthew Hanna1, Liron Pantanowitz1, Douglas Hartman1

1University of Pittsburgh Medical Center, Pittsburgh, PA, USA.
E-mail: hannamg2@upmc.edu

Content: There are many vendor solutions that may need to be evaluated for an institution interested in adopting digital pathology. In order to find the right fit, demonstrating that a solution can handle digital workflow processes is imperative for the vendor evaluation process. Although some pathology workflows are routine, each institution may have unique workflow requirements. Crafting playscripts that describe typical workflow scenarios has previously proven to be useful to assess laboratory information system vendor performance. Therefore, the aim of this project was to develop a standard set of playscripts that could be used to appraise a digital pathology system. Technology: Digital pathology solutions that offered hardware (scanners) and/or software (viewing and sharing of digital slides) for primary diagnosis were evaluated.

Design: Employing standard playscripts for a digital pathology workflow proved to be extremely helpful during vendor demonstrations to evaluate their system’s ability to be potentially implemented at our institution for primary diagnosis. We also recommend using such playscripts to objectively compare different digital pathology systems.

Table 1: Digital pathology playscripts

| Scenario | Workflow description |
|----------|----------------------|
| #1       | We receive a 3-part biopsy case from a screening colonoscopy. Initial H&E’s are cut and a complete diagnosis can be made from these slides |
| #2       | We receive a biopsy from a colonic mass in a patient with a prior history of malignancy. An initial H&E recut is made, but additional levels and some immunostains are required plus prior history needs to be reviewed. An intradepartmental consultation is needed for this case |
| #3       | We receive a surgical resection of a pancreatic mass that was previously biopsied by fine needle aspiration and core biopsy at our institution. Prior to gross examination, an intraoperative consultation was performed. Comparison with prior material is necessary. Gross images of the specimen were obtained and biomarker immunostains need to be ordered. This case is also selected to be presented for tumor board |
| #4       | We receive a consultation request from an outside institution to be performed on digital images from a foreign institution |
| #5       | We receive a case with invasive ductal carcinoma of the breast. The tissue is stained with estrogen receptor, progesterone receptor, human epidermal growth factor receptor 2 and ki-67. Image analysis is required and needs to be reported in the final diagnosis |

Conclusion: Telecytology using live video microscopy can be validated prospectively during routine clinical ROSE service by showing non-inferiority of diagnostic error rates and concordance rates between digital and glass modalities. Prospective collection of cases helps avoid sample bias and is easy to perform, but may take longer to complete than a retrospective study. Validation helped “stress test” the telecytology system to resolve technical issues without compromising user performance.

Recorded Pathology Didactic Lectures for Global Online Distribution: A 12-year Experience

Matthew Hanna1, Ishtiaque Ahmed1, Liron Pantanowitz1

1University of Pittsburgh Medical Center, Pittsburgh, PA, USA.
E-mail: hannamg2@upmc.edu

Content: Pathology didactic lectures are typically available at academic teaching centers, professional society meetings where participants pay for registration, or via organizations that offer paying members access to live or archived webinars and podcasts. However, there are limited free online resources to didactic lectures. As the University of Pittsburgh Medical Center (UPMC) pathology department expanded into a multi-hospital healthcare system spread over hundreds of miles, there was a need to provide remote continuing education to its department members. For this purpose, didactic lectures were shared online not just to UPMC staff, but also globally. The aim of this evaluation was to determine the impact of this free online resource. Technology: GoToWebinar (Citrix) was used to record conferences. These video files were hosted on a Microsoft Media Server and made public on the website http://pathologyconference.upmc.edu. Google analytics was utilized to track website statistics.

Design: All weekly conferences, seminars and grand rounds at UPMC were broadcast live to internal users, with the speakers’ permission recorded, and archived. Recorded videos were converted to MP4 and WebM (HTML5) files. Links to select archived videos were made available online via the department of pathology website.
Conferences were uploaded daily, indexed chronologically, and rendered searchable using a Google Custom Search Engine. Results: From 2002 to 2016, over 2,500 lectures were broadcast and recorded. For some lectures live discussion sessions and audience polling was included. There was an average 40 internal registrants per lecture. Of these, 844 videos have been shared online for public viewing. Analytics from January 1, 2005 indicate 69,656 page views and 34,650 online sessions. Average time to load our webpage was 1.51 seconds. There were an average of 9 active daily users and 200 thirty-day active users. Returning users comprised 56% of website traffic. Figure 1 depicts the global location of website visitors (color gradient portrays the number of active sessions). Desktop browsing was the preferred method for viewing presentations. Conclusions: Web hosting of didactic pathology presentations proved to be a convenient and interactive mechanism to provide remote continuing educational opportunities to all of our department members. Although the availability of these online lectures was not advertised, offering this expert educational content as a free resource to the worldwide pathology community has driven users to visit our departmental website.

Optimizing a Pediatric Pathology Digital Slide Teaching Set with PathXL™ Tutor Software as an Educational Tool for Pathology Training

Oscar López-Núñez1, Cathy Chen1, Bradley Clifford1, Liron Pantanowitz2, Jennifer Picarsic1,2,3

1Department of Pathology, University of Pittsburgh Medical Center, 2Medical Student, University of Pittsburgh School of Medicine, 3Children’s Hospital of Pittsburgh of University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: lopexnunezof@upmc.edu

Content: Whole slide digital imaging (WSI) has an important impact on pathology education leading to significant changes in resident/fellow training. The scope of pediatric pathology is diverse, but the implementation of WSI can help narrow the learning gap by ensuring that all trainees see pertinent entities during their short training period. We improved the pediatric pathology WSI teaching set for trainee’s education by developing integrated modules with an innovative web-based platform. Technology: WSI files were captured at 20X resolution using an Aperio ScanScope CS scanner (Leica Biosystems Imaging, Inc., Vista, CA, USA) and uploaded into PathXL™ Tutor software (PathXL Ltd., Belfast, United Kingdom). Design: A set of 217 scanned glass slides hosted on a department server was accessible case it is not used for its own collection of slides or of anyone else’s. Rather, advantage is taken of it built-in capability of storing metadata and sorting by field. Of particular importance is the ability to include links to outside sources of digital images. Design: As a prototype, links have been established to digital slides at the University of British Columbia, the University of Western Ontario and the University of Leeds. The NYCPM collection can be accessed via its hosting site. The Digital Pathology Association has a number of additional sites which can be added to the collection. Thus, one can sort by tissue, by pathology, by institution or by other features which can be added (e.g., FISH, CISH staining). One can look at multiple slides of tuberculosis from various sources. One can look at different lung pathologies. One can look at all slides of inflammation or all slides of soft tissue tumors. It is all a function of the initial establishment of the database. Results: The website is available for multiple uses: to construct pathology courses, to use it for Tumor Board practice or to use it to construct exams at multiple levels. The essential feature is that you have here a data mining capability so that you are not limited to the storage capacity of an in-house collection. Conclusions: This is a useful tool for the many applications described above.

An Image Database Featuring Access to Decentralized Collections

Zev Leifer1

1Pre-Clinical Sciences, New York College of Podiatric Medicine, New York, USA. E-mail: zleifer@nycpm.edu

Content: There are many needs for slide collections. These include undergraduate and post-graduate medical training, Tumor Board practice, resident review and so on. There are numerous collections of slides, at particular institutions, each with their own particular method of searching and organizing. There is a need for a central site to enable such an organizing capability. This work describes such a site. Technology: The website of http://quartzy.com is used as the platform. In this

Figure 1: World map showing public usage of University of Pittsburgh Medical Center online pathology didactic lectures
using Aperio Webscope (Leica Biosystems Imaging, Inc., Vista, CA, USA) with a separate Microsoft Excel file containing associated clinical information/diagnoses for trainee review. These WSI files were copied to the PathXL image server and integrated with relevant case material using PathXL™ Tutor software. 

**Results:** Thus far, 100 cases are organized into 18 subtopics. Each case module contains a clinical vignette and 1 to 13 WSI including H&E, immunohistochemical, and special stains. The final diagnosis and educational content are revealed after clicking the diagnosis button. Selected cases also contain gross pictures and embedded broadcasted lectures, videos, or PDF articles for further study [Figure 1]. Users are allowed to make annotations on WSI files and save them into their personal profile for future review. 

**Conclusions:** An integrated web-based resource that facilitates the use of WSI to supplement traditional teaching methods provides educational benefits for pathology trainees. PathXL™ Tutor serves as an ideal platform to offer trainees consistent exposure to diverse entities in pediatric pathology. The benefit of PathXL™ Tutor includes easy access to web-based digital teaching sets integrated with relevant educational material and resources. PathXL™ Tutor is being explored for the creation of test modules and as a competency-assessment tool in residency and fellowship training.

---

**Expanding the Next Generation Sequencing (NGS) workflow to include Confirmatory Testing**

Yelina Noskin1, Emil Cobarrubia1

1Marketing, Abbott Informatics. E-mail: yelina.noskin@abbott.com

**Content:** Next Generation Sequencing (NGS) is becoming more systematically used in the clinical diagnostic laboratory, for somatic and germline genetic mutations. One of the benefits of many NGS platforms is the ability to multiplex not only the tests but the number of samples being analyzed. Many LIMS systems have developed workflows that can be followed for sample management during the NGS process, these workflows often stop at the point of data analysis or right after with a result obtained based on the NGS test. However this is often not the last step of the diagnostic process for the sample, as many NGS test results require confirmatory testing after data analysis. 

**Technology:** NGS workflows are inherently multidisciplinary and rely on the inputs of many departments, which in turn, can utilize a large array of equipment, tools and other software in addition to a LIMS. All such entities communicate to provide a unified technological ecosystem that supports the complexities of a NGS workflow. In this case STARLIMS as part of that ecosystem, manages patient and family information with relevant samples received, prepared and processed before, during and after testing. It manages short and long-term storage of biospecimens and provides the outcome to requesting healthcare professionals.

**Design:** NGS workflows were designed to follow samples that become multiplexed into single runs, and de-multiplexed for result analysis. The workflows were extended to include reflex testing based on results obtained after result analysis. Confirmatory/Reflex tests were grouped by specific methodology, example Sanger Sequencing Parameters, FISH Parameters, and PCR Parameters. Samples that contained results which required Confirmatory/Reflex testing were queued into the appropriate test group and were thus available. 

**Results:** Prior to including confirmatory testing into the LIMS workflow, samples that needed confirmatory testing, required the manual set up and batching of these test and samples. This could result in delays, and errors especially in high throughput laboratories. The inclusion of multiple possibilities of Confirmatory/reflex testing allowed for comprehensive sample tracking through the entire laboratory process and allowed to include the results of both NGS result analysis and confirmatory test results in the sample history.

**Conclusion:** Consolidating an NGS workflow with confirmatory testing into one workflow in the LIMS optimizes the laboratory process, reduces the need for documentation redundancies and increases the overall efficiency of tracking samples undergoing NGS testing in a clinical diagnostic laboratory.

---

**Multiplatform Virtual Reality Pathology**

Alexander Ringlein1, Eben Olson1, Sachith Gullapalli1, Acshi Haggenmiller1, Henry Li1, Maxwell Emerson1, Alyssa Siefert1, Joseph Zinter1, Richard Torres1

1Computer Science, Yale University, New Haven, Connecticut, USA. E-mail: alexander.ringlein@yale.edu

**Content:** New consumer virtual reality (VR) platforms provide intuitive control mechanisms for interacting with and viewing three-dimensional (3D) data. Multiphoton microscopy and optical tissue clearing allow generation of high-resolution 3D image stacks potentially providing new perspectives on histology. We have designed a VR application for viewing realistic renderings of tissue samples constructed from stacks of multiphoton microscopy images. The application was built
Immunoprofiling analysis using quantitative P

To assess and quantify the tumor associated immune cells (TAICs) and immune checkpoints in lung cancer, we optimized the need of large immunoprofiling analysis of solid tumors. Content: The advent of cancer immunotherapy has prompted the need of large immunoprofiling analysis of solid tumors. To assess and quantify the tumor associated immune cells (TAICs) and immune checkpoints in lung cancer, we optimized a multiplex immunofluorescence (mIF) technique that can be applied to formalin-fixed paraffin-embedded (FFPE) tissues. Technology: The slides were stained using Opal™ Kit (PerkinElmer, Waltham, MA, USA), scanned using a Vectra 3™ multispectral microscope (PerkinElmer) and analyzed using InForm™ 2.2.1 software (PerkinElmer). The final data was consolidated using SpotFire™ software (TIBCO Software Inc., Palo Alto, CA, USA). Design: Nineteen immune markers were optimized for mIF using standard chromogenic immunohistochemistry in control, and then we constructed four mIF panels: Panel 1: AE1/AE3 (pancytokeratin), PD-L1, PD-1 (or CD4), CD3, CD8 and CD68; Panel 2: AE1/AE3, CD20 (or CD-1), CD3 (or Granzyme B), FOXP3, CD45RO and CD57; Panel 3: AE1/AE3, PD-L1, B7-H3, B7-H4, IDO-1, VISTA and CD3; and Panel 4: AE1/AE3, ICOS, LAG3, TIM3, OX40, CD3 and CD20. FPFE tissue microarrays (TMAs) contains 256 tumors (adenocarcinomas, ADCs=156; squamous cell carcinomas, SCCs=100), 108 whole sections (WS) specimens (ADCs=61; SCCs=48) and 30 core needle biopsies (CNB) specimens (ADCs=15; SCCs=15) were analyzed. Results: Positive PD-L1 expression (>5%) in malignant cells (MCs) was observed in 23% ADCs and 31% SCCs in TMAs. 39% ADCs and 56% SCCs in WS showed higher expression of PD-L1 in chemotherapy treated than chemo-naïve cases. Furthermore, CD3+CD4+ (P=0.0030) and CD57+CD45RO+ (P=0.0001) were higher in ADCs than SCCs in all specimens. Positive significant correlation was found when we compared WS and CNB. Panel 3 that included PD-L1, B7-H3, B7-H4 and IDO-1 in TMAs showed that 20% of cases were negative in MCs for all markers, 25% expressed only one marker, 30% expressed two markers, and 25% expressed more than three markers. TAICs expressing ICOS, LAG3, TIM3 and OX40 in panel 4 showed the intricate interactions between T cells (CD3+) expression and immune checkpoints expression. Conclusion: The advantage of this technique is based on its multiplex approach generating more data per sample and allowing the co-localization of markers. The next step will be to utilize an automatic staining device for efficient mIF.

Optimization of an Efficient Algorithm for Immunoprofiling of Lung Cancer Using InForm Image Analysis Software

1Naohiro Uraoka1,2, Edwin Parra1, Jaime Rodriguez-Canales1, Carmen Behrens1, John Heymach2, Ignacio I. Wistuba1,2

Departments of 1Translational Molecular Pathology and 2Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA. E-mail: NUraoka@mdanderson.org

Content: Immunoprofiling analysis using quantitative immunohistochemistry of tumor-infiltrating lymphocytes and immune checkpoints has been helpful for predicting responses to treatment in many types of solid cancer, including lung cancer. Recently, multiplex immunofluorescence (mIF) has emerged as a powerful tool for immunoprofiling analysis, offering simultaneous detection of multiple markers. To perform the efficient analysis of immunoprofiling using mIF, we optimized

Immunoprofiling of Lung Cancer Tissue Specimens Using Multiplex Immunofluorescence and a Multispectral Imaging Platform

1Naohiro Uraoka1, Edwin Parra1, Jaime Rodriguez-Canales1, Barbara Mino1, Mei Jiang1, Jiejin Zang1, Carmen Behrens2, Cesar Moran2, Cara Haymaker4, Chantale Bernatchez2, Boris Sepesi2, Don Gibbons2, John Heymach2, Clifford C. Hoyt4, Ignacio I. Wistuba1,2

Departments of 1Translational Molecular Pathology and 2Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

Content: The need for large immunoprofiling analysis of solid tumors. To assess and quantify the tumor associated immune cells (TAICs) and immune checkpoints in lung cancer, we optimized a multiplex immunofluorescence (mIF) technique that can be applied to formalin-fixed paraffin-embedded (FFPE) tissues. Technology: The slides were stained using Opal™ Kit (PerkinElmer, Waltham, MA, USA), scanned using a Vectra 3™ multispectral microscope (PerkinElmer) and analyzed using InForm™ 2.2.1 software (PerkinElmer). The final data was consolidated using SpotFire™ software (TIBCO Software Inc., Palo Alto, CA, USA). Design: Nineteen immune markers were optimized for mIF using standard chromogenic immunohistochemistry in control, and then we constructed four mIF panels: Panel 1: AE1/AE3 (pancytokeratin), PD-L1, PD-1 (or CD4), CD3, CD8 and CD68; Panel 2: AE1/AE3, CD20 (or CD-1), CD3 (or Granzyme B), FOXP3, CD45RO and CD57; Panel 3: AE1/AE3, PD-L1, B7-H3, B7-H4, IDO-1, VISTA and CD3; and Panel 4: AE1/AE3, ICOS, LAG3, TIM3, OX40, CD3 and CD20. FPFE tissue microarrays (TMAs) contains 256 tumors (adenocarcinomas, ADCs=156; squamous cell carcinomas, SCCs=100), 108 whole sections (WS) specimens (ADCs=61; SCCs=48) and 30 core needle biopsies (CNB) specimens (ADCs=15; SCCs=15) were analyzed. Results: Positive PD-L1 expression (>5%) in malignant cells (MCs) was observed in 23% ADCs and 31% SCCs in TMAs. 39% ADCs and 56% SCCs in WS showed higher expression of PD-L1 in chemotherapy treated than chemo-naïve cases. Furthermore, CD3+CD4+ (P=0.0030) and CD57+CD45RO+ (P=0.0001) were higher in ADCs than SCCs in all specimens. Positive significant correlation was found when we compared WS and CNB. Panel 3 that included PD-L1, B7-H3, B7-H4 and IDO-1 in TMAs showed that 20% of cases were negative in MCs for all markers, 25% expressed only one marker, 30% expressed two markers, and 25% expressed more than three markers. TAICs expressing ICOS, LAG3, TIM3 and OX40 in panel 4 showed the intricate interactions between T cells (CD3+) expression and immune checkpoints expression. Conclusion: The advantage of this technique is based on its multiplex approach generating more data per sample and allowing the co-localization of markers. The next step will be to utilize an automatic staining device for efficient mIF.

Technology: The application was developed using Unity 5.5, a popular game development engine, along with libraries for integration with both the HTC Vive VR headset and the Oculus Rift with Oculus Touch controllers. We used a high performance computer with a dedicated graphics unit (Nvidia GTX 1080) and 32 GB RAM to test the application. Design: The application was iteratively developed based on feedback from testers and pathologists. The tissue is rendered as a 3D cube using stacks of semi-transparent images along each axis. A hematoxylin and eosin transfer function may be applied to the data to simulate staining; the contrast and transparency of the images can also be altered in-application. The cube is displayed stereoscopically through a VR headset. Using motion controllers, the user may translate, rotate, and slice the data along the axis. Results: A variety of kidney and prostate samples were examined using the two virtual reality devices and the application. While the HTC Vive allows for room-scale, 360 degree tracking, this was found to be less relevant in the context of examining pathology samples. The Oculus Touch controllers provide additional button inputs for control and better tracks the individual fingers of the user. The performance of the application on both systems was comparable. Conclusions: This technology allows pathologists to view renderings of histology data in VR. It provides unique views of the surface and structure of 3D samples which are unachievable using traditional 2D microscopy. Further study comparing the effectiveness of our virtual reality technique to more traditional methods of viewing tissue samples is warranted.
Design: Ten immune markers were optimized for mIF using standard chromogenic immunohistochemistry as a control. We designed two mIF panels with these markers: Panel 1: AE1/AE3 (cytokeratins), PD-L1, PD-1, CD3, CD8 and CD68; Panel 2: AE1/AE3, CD20, FOXP3, Granzyme B, CD45RO and CD57.

Results: Using the spectra of the individual fluorophores extracted in singly stained tonsil sections, as well as an autofluorescence spectrum from an unstained tissue, multispectral libraries for two panels were constructed. Based on these libraries, the unmixed signals were extracted from mIF sections and reconstructed into composite images. To automatically divide each image into two tissue compartments: tumor and stroma, a pathologist trained the software by drawing representative training regions based on cytokeratin staining. Following tissue segmentation, cellular segmentation was performed using DAPI as a counterstain to identify the location of all cell nuclei. Cytoplasmic and membrane subcellular compartments were also determined based on the signal thresholds decided. Once individual cells were identified, the quantification and colocalization of immune markers were achieved according to the intensity of each signal in the trainable phenotyping session. The data created with InForm were also exploited to calculate the distance between tumor and inflammatory cells using RStudio (RStudio, Inc., Boston, MA, USA, https://www.rstudio.com/).

Conclusions: This algorithm offered accurate, reproducible and high-throughput data regarding the tumor microenvironment, when it was used carefully under the supervision of pathologists.

Immunohistochemical Expression of Programmed Cell Death Ligand 1 in Nonsmall Cell Lung Cancer evaluated by Digital Image Analysis: A Comparison Study Between Tissue Microarray and Whole Section

Pamela Villalobos¹, Edwin R. Parra¹, Barbara Mino¹, Ignacio I. Wistuba¹, Jaime Rodriguez-Canales²

¹Department of Translational Molecular Pathology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA. E-mail: pavillalobos@mdanderson.org

²Department of Pathology, Children’s Healthcare of Atlanta, Atlanta, Georgia, USA. E-mail: alexis.carter@chta.org

Content: Programmed cell death ligand 1 (PD-L1) is a major immune checkpoint protein and its expression is used for optimal patient selection for anti PD-1/PD-L1 therapies. Tissue microarray (TMA) technology has been widely used for the immunohistochemical (IHC) analysis of biomarkers in cancer research. However, its reliability for the assessment of PD-L1 expression has not been clarified. Our aim was to compare the scoring reproducibility of IHC PD-L1 expression between TMA and their whole section (WS) counterparts in NSCLC. Technology: The immunostained WS and TMA were digitally scanned using the Aperio® ScanScope AT2 scanner (20×) and were visualized using ImageScope™ software. The IHC PD-L1 expression was evaluated using digital image analysis (DIA) with Aperio Image Toolbox™ and Genie™ (Aperio, Leica Microsystems). Design: A TMA set (1 to 3 cores of 1 millimeter per patient) and their WS counterparts of 148 patients with NSCLC were analyzed for PD-L1 expression. PD-L1 expression (E1L3N clone) was evaluated as percentage of positive tumor cells (MCs) by IHC and DIA. Tumor was scored PD-L1 positive when >5% of cells showed membranous staining. Results: PD-L1 TMA scoring showed a positive correlation with WS scoring (r=0.592, P < 0.001). However; the TMA method revealed an overestimation of the PD-L1 positivity comparing with WS (124 vs 55 positives cases respectively, P=0.001). Thus, compared with WS, TMA-IHC had a sensitivity of 93%, specificity of 22% and positive predictive value of 41% for PD-L1. Over-scoring causes in TMA were: misinterpretation of positive macrophages as tumor cells by DIA (63%), PD-L1 tumor heterogeneity (TMA hot spots) (28%) and unspecified nuclear or cytoplasmic PD-L1 expression (6%). The discordance between TMA and WS scoring caused by PD-L1 heterogeneity; decreased significantly when 3 TMA cores per case were evaluated vs <3 cores (10% vs 33%). Conclusion: IHC PD-L1 expression analysis in NSCLC revealed important discordance between TMA and WS due to misinterpretation/ misidentification of staining and cancer cells by DIA and the PD-L1 tumor heterogeneity. To avoid the heterogeneity issue is highly recommended to evaluate at least three TMA cores to obtain a better full section representation. DIA represents an important tool for IHC analysis but requires the pathologist supervision.
Gender Gap in Clinical Informatics to Trailblaze: Bridging the Gap

An Opportunity for Pathology to Trailblaze: Bridging the Gender Gap in Clinical Informatics

**Table 1: Statistics of pathologists compared to physicians in all specialties**

| Category                              | All physicians (including pathologists) | Pathologists (%) |
|---------------------------------------|-----------------------------------------|------------------|
| Actively Licensed Physicians in USA (2014) | 916,264                                 | 17,981 (2.0)    |
| Total CI Diplomates (2013-2015)       | 1106                                    | 65 (5.9)         |
| ACM Program Directors*                | 24                                      | 3 (12.5)         |
| ACM Fellows*                          | 243                                     | 6 (2.5)          |
| Morris F. Collen Awards*              | 15                                      | 2 (13.3)         |
| CI Current Fellows*                   | 56                                      | 5 (8.9)          |

*As of March 2017. CI: Clinical informatics

Data was collected from peer-reviewed published articles from the Federation of State Medical Boards, the College of American Pathologists’ workforce analyses, from the American Medical Informatics Association and from Clinical Informatics Fellowship program data. **Design:** Multiple studies of pathologists and all physicians were analyzed to compare the proportion of pathologists among all physicians and in various CI-related professional societies and awards [Table 1]. The expected percentage of Pathologists in any given category was set to the proportion of pathologists in the general physician population. The Chi-Squared test was used to determine the p-value between observed and expected values. **Results:** Pathologists are significantly more likely to be a CI diplomate compared to non-pathologists (odds ratio 3.12; p < 0.0001). The odds ratios for a pathologist becoming a CI Fellowship Program Director or a recipient of the Morris F. Collen Award, the highest award given by the American Medical Informatics Association, were 7.14 and 7.68, respectively. The proportion of CI fellows who were pathologists was also higher than expected. **Conclusions:** The consistent strong presence of Pathologists in CI is concordant with the increasingly data intensive nature of the Pathology specialty. However, there are fewer women pathologists in CI than would be expected, so pathologists have an opportunity to further strengthen our presence and improve patient care by addressing these gaps.

**Procedural Optimization and Signal Normalization of Kinomic Peptide Microarrays**

**Alex Dussaq 1, Joshua Anderson 2, Christopher D. Willey 2**

Departments of 1MSTP Program and 2Radiation Oncology, University of Alabama at Birmingham, Birmingham, Alabama, USA. E-mail: adussaq@uab.edu

**Content:** Kinases play a role in every cellular process involved in tumorigenesis ranging from proliferation, migration, and protein synthesis to DNA repair. While genetic sequencing has identified most kinases in the human genome, it does not describe the ‘kinome’ at the level of activity of the kinases against kinase targets. The PamGene PamChip system records
Table 1: Gender statistics in clinical informatics

|                          | All physicians (pathologists and nonpathologists) | Pathologists only |
|--------------------------|--------------------------------------------------|-------------------|
|                          | Total | Female (%) | Total | Female (%) |
| Actively Licensed Physicians in USA (2014) | 916,264 | 293,565 (32) | 17,981 | 6183 (34) |
| Total CI Diplomates (2013-2015) | 1106 | 209 (19) | 65 | 7 (11) |
| ACMI Fellows* | 243 | 15 (6) | 6 | 1 (17) |
| Morris F. Collen Awards (AMIA)* | 15 | 0 | 2 | 0 |
| CI Current Fellows* | 56 | 6 (11) | 5 | 1 (20) |
| CI Fellowship Program Directors* | 24 | 0 | 3 | 0 |
| API Lifetime Achievement Awards* | N/A | N/A | 15 | 0 |
| API presidents* | N/A | N/A | 15 | 2 (13) |

*As of March 2017. N/A: Not available

and compares the phosphorylation of 144 tyrosine or serine/threonine peptides as they respond to cellular kinases utilizing a peptide microarray. This microarray platform records phosphorylation at multiple time points resulting in both end level and kinetic measurements. Similar to other microarray technologies this technology needs a thorough investigation of background and signal determination. Technology: Web Server: Node v5.6.0; Programming Language: JavaScript; Database: Mongo v3.2.6. Design: We utilized four Glioblastoma cell lines with various changes to the MARKS kinase ran in triplicate as technical replicates. Based on these results we created a level based Mongo Database that is accessed utilizing a standard web environment (JavaScript, HTML5, CSS) to perform analyses and visualize results of various reproducibility measures. Results: Current analytical techniques shift all signal – background values recorded above the 5th quantile, and log transformation all possible values. This normalization technique is applied across all samples in a given experiment and complicates any future analytical processes. Additionally this results in 5% of the data being lost. We proposed a new technique of log transforming signal / background this makes all values comparable and results in no data loss. Additionally, we observed local background and signal were highly correlated ($\rho = 0.96$), we corrected this utilizing linear regression to eliminate the effect of the signal on its own background without removing the effect it may have neighboring signal backgrounds ($\rho = 0.23$). These techniques in combination resulted in a moderate increase in signal reproducibility of $\rho = 0.9918$ to $0.9925$. Conclusions: The changes proposed here allow global comparison of results, correct the issue of missing data, correct an error that dampens the signal from high signal spots and leads to a increase in reproducibility. These analyses utilize web techniques that allow any dataset to be corrected without any downloads or data sharing.

Sharing CellaVision Blood Smear Images with Clinicians via the Electronic Medical Record

Matthew G. Hanna1, Lydia Contis1, Roseann Hyman1, Christine Garcia1, Liron Pantanowitz1

1University of Pittsburgh Medical Center, Birmingham, Alabama, USA.
E-mail: hannahmg2@upmc.com

Content: The CellaVision automated system digitizes peripheral blood smears and performs blood cell differentials. These images are available for local review in hematology laboratories on computer monitors or they can be transmitted for remote review via telehematology. Clinicians such as hematologists who typically review their patient’s peripheral blood smears often need to contact an offsite hematology hospital laboratory for this purpose, frequently during evening or weekend hours. To overcome this barrier, our institution sought to make CellaVision blood smear images available to clinicians and trainees in the electronic medical record (EMR). The images could therefore be viewed at any time and from any computer or mobile device which could access the EMR. Technology: CellaVision DM96 (CellaVision AB, Lund, Sweden) digital hematology analyzers. Images hosted on a Microsoft SQL Server. Microsoft Internet Information Services 7 web server and ColdFusion 9 middleware. Sunquest laboratory information system. Cerner PowerChart EMR. Design: Image sharing was accomplished in 2 phases. For phase 1, a secure web portal (called “HemaVue”) was created to allow clinicians to access CellaVision digital images with associated metadata (i.e., patient unique identifiers, blood counts). CellaVision image data was archived daily and stored for 6 months. For phase 2, CellaVision instrument results were tagged with the test code “Image Differential Performed”. This code appeared in the patient result to inform the clinician this type of differential had been performed, as opposed to a manual or automated differential code. The HL7 message router transmitting results with this tag created a link in the EMR to the HemaVue website. This allowed clinicians to access CellaVision images from a patient’s chart within the EMR. Results: The CellaVision image database contained 3.38 million images for 23,000 patients, comprising 12TB of data on per 6-month temporal storage. Database and website maintenance had to be internally supported. While clinicians were pleased with having access to blood smear images, they expressed dissatisfaction that images were not always available immediately, not retained indefinitely, not offered

Journal of Pathology Informatics  
http://www.jpathinformatics.org

S31
on all patients, and that the images were of isolated white blood cells only. **Conclusion:** CellaVision images were able to be successfully unlocked from the vendor’s server and subsequently made accessible for remote viewing by linking patients’ laboratory results in the EMR to a custom built website. While such enterprise image sharing of digital blood images offered some improvement for clinicians seeking to review peripheral blood smears, enhancements will be required to make this more readily available for all patients at our institution.

**PathXL in an Academic Hospital Setting: Lessons of Implementation**

Matthew G. Hanna¹, Brad Clifford¹, Matthew O’Leary¹, Douglas J. Hartman¹, Anthony Piccoli¹, Jon Duboy¹, Liron Pantanowitz¹

¹University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: hannahmg2@upmc.com

**Content:** Our department has used whole slide images (WSI) for many educational uses such as digital teaching sets and unknown conferences. However, most of these offerings required homegrown solutions, were not standardized across our department, and thus made it difficult to manage and search content. Therefore, we acquired PathXL to address pathology education and training needs. We required a web-based platform that allowed multiple users to easily incorporate new and archived WSI with media content (e.g. gross photos, documents, videos), publish virtual teaching sets online, and create tests. Our aim was to deploy PathXL on internal servers instead of hosting it on the vendor’s cloud environment. **Technology:** PathXL Tutor web-based educational management system. Microsoft Azure and Microsoft OneDrive for Business cloud solution. **Design:** PathXL tutor was deployed in a private Azure cloud behind a demilitarized zone (DMZ) using an application server, database server, and network storage. Development, test, and production virtual environments were created. The production instance is built on-premise to avoid latency issues while transferring large image files to the cloud server [Figure 1]. Implementation addressed computing architecture, user authentication, data transfer, as well as roles and permissions. Allocation of 7TB for image storage was made available. **Results:** The default Client-Server model of PathXL was reconfigured to employ the Business-To-Consumer (B2C) Azure Active Directory to provide increased security, as well as allow internal and external user registration, authentication, and authorization. Internal institutional accounts were authenticated at the application level in PathXL, whereas external accounts required external authentication via Azure. OneDrive was used for data storage and uploading files to be accessed via PathXL. Large WSI file transfers were mediated by Quality of Service (QoS) to maximize network performance and decrease latency of data uploading. Institution wide accounts were granted access to private and public content while external users were only granted access to public content. **Conclusions:** PathXL was successfully deployed in our institution’s on-premise and Azure cloud platform that allowed us to easily store, sync, and share digital files. Implementation in Azure was challenging, required significant local information services and vendor input, and delayed going live for this project. The benefits of hosting PathXL on our own data center include running multiple virtual environments such as test and production systems, secure mediation of external user authentication, and secure file upload and storage.

**Federated Laboratory Information Systems Trending Away from the Tower of Babel**

Matthew G. Hanna¹, Liron Pantanowitz¹, Colleen O’Hare¹, Douglas J. Hartman¹

¹University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: hannahmg2@upmc.com

**Content:** The laboratory information system (LIS) depends on data tables and dictionaries for relational database management. Large, integrated hospital systems may have various dictionary elements to describe identical entities for each individual LIS. Standardization of these data tables may be difficult to ascertain due to the complexity of testing at different laboratories and varying clinical workflows that demand different lab test naming conventions. Our institution aimed to review and standardize all anatomic pathology part types for 14 internal hospital sites. **Technology:** The anatomic pathology LIS utilized at 14 hospitals within our healthcare system is Cerner CoPath Plus. The American Medical Association Current Procedural Terminology (CPT®) was used to determine specimen part types. **Design:** The specimen...
part type dictionary was reviewed in conjunction with the CPT manual by specimen part type level (from I to III of VI). All part type entries were iteratively identified. Entries were either merged into a standardized part type, inactivated, or newer part types added for all laboratory sites. Results: A total of 388 original part type dictionary entries were identified and assessed. Levels I, II, and III part type dictionaries were consolidated by 94%, 67%, and 69%, respectively [Table 1]. The most significant reduction occurred with bone fragments, which were reduced from 18 to 2 entries. Related billing errors were also resolved, resulting in a 10% overall correction.

Conclusions: Ongoing maintenance to standardize dictionaries within the LIS database for an integrated delivery network is a worthwhile endeavor that may significantly reduce redundant dictionary entries and related billing errors. Education of stakeholders at each hospital is required to employ consistent data elements and maintain consolidated LIS dictionaries.

### Analysis of Influence of Additional Diagnostic Clues during Pathology Diagnosis

**Mikhail Kovalenko**, Richard Hammer, Dmitriy Shin

1Department of Pathology and Anatomical Sciences, University of Missouri, Columbia, Missouri, USA. E-mail: kovalenko@health.missouri.edu

**Content:** Traditional pathology diagnostic process routinely relies on disease-specific diagnostic clues. While in majority of clinical cases these clues result in accurate diagnosis, in other cases they can lead to diagnostic pitfalls. Our ongoing research is focused on revealing diagnosis-related details and heuristics that can be used to quantify and potentially improve the diagnostic process in pathology using whole-slide imaging (WSI) and analytical tools. Specifically, in this work we propose an informatics pipeline to identify and quantify additional diagnostic clues that can improve diagnosis, we extended association rule mining techniques to measure information gain of the additional diagnostic clues. We induced association rules based on the collections of cell types covered by a user’s gaze track and the resulting diagnostic decision made by pathologist in a simulated session.

**Results:** When interesting rules for a particular diagnosis were considered, we observed that an additional tissue feature (diagnostic clue) increases the likelihood of a correct diagnosis. Rules using only two diagnostic clues showed better support (were used more frequently), as would be expected. However, adding another diagnostic clue to the same rule would increase confidence that that rule would result in a correct diagnosis up to 79% with a similar improvement in lift. To validate our findings, we computed Kullback-Leibler divergence that indicates information gain generated by additional diagnostic clues.

**Conclusions:** We believe that with such an increase the chances of an occurrence of a diagnostic pitfall are correspondingly reduced. Further research is needed to ensure our findings are consistent across a wider range of tissues and diagnoses. This may lead to a closer look at current recommended diagnostic procedures and more specific recommendations for pathologists on diagnostic pitfall avoidance.

### Adaptive Pathology Laboratory Teaching System

**R. R. Lindquist**

1Pathology, University of Connecticut School of Medicine, Farmington, CT, USA. E-mail: lindquist@uchc.edu

**Content:** In order to use the allotted time more efficiently, we adapted lab time by devoting less time to concepts that students easily grasped and more time to difficult concepts by using a web based clicker system to assess student comprehension. Initially Uniform Server (www.uniformserver.com) and later Bitnami web stacks (https://bitnami.com.Stacks) were employed. The Uniform Server, a Microsoft Windows operating system web stack, is portable and can run from a thumb drive or any USB storage device. It requires no installation. Thus the system can simply be used on any computer with windows operating system and a static IP address. The Bitnami web stack can run in Windows, Mac OS X or Linux environments as well as VMware or VirtualBox virtualized environments, and popular cloud platforms such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform.

**Design:** A web based clicker system in conjunction with Aperio virtual microscopes (http://www.leicabiosystems.com/digital-pathology/apero-digital-pathology-slide-scanners) was developed to actively engage second year medical students in pathology laboratories and
to evaluate their comprehension. No more time was devoted to concepts that students easily grasped and more time was devoted to difficult concepts until most students had a good understanding of the concept. Thus the flow of the pathology laboratory adapted to how well students understood the objectives of the laboratory session. Results: We were unable to determine the effect adaptive teaching had on student performance. However, student attitudes of the use of the web based adaptive approach system in laboratory were very favorable. From a faculty point of view, this approach highlighted areas that needed more emphasis and use the time more efficiently. Conclusion: By adapting to student comprehension of the learning objectives, the allotted time was used more efficiently and more effectively.

Evaluation of a Natural Language Processing Platform for Rapid Development of Medical Information Retrieval Resources

Richard Long1, Laura Monovich2, Maria-Teresa Ramirez2, Nancy Single1, Mark Lloyd1, Anil Parwani1, Ulysses G. J. Balis1

1Algorithm Development, Inspirata, Inc., Tampa, Florida; 2The Ohio State University Comprehensive Cancer Center, Columbus, Ohio, USA. E-mail: MLloyd@inspirata.com

Content: Critical information in cancer case surgical pathology, cytopathology and molecular testing reports are primarily contained in narrative “free text” sections that limit the ability to implement similarity based case retrieval systems used in modern pathology informatics. This work describes a flexible Natural Language Processing (NLP) framework based on open-source tools for the reliable identification and extraction of biomedically relevant information in cancer case reports. Technology: We evaluated a web services-oriented NLP platform, the Inspirata NLP Server (INS), developed at Inspirata, Inc., in Tampa, Florida. The architecture is open-source and pipeline-based with annotators, operating on a wide variety of document source formats, for linguistic manipulation, named-entity recognition, negation detection and classification. Design: We created six RESTful resources for processing surgical synoptic, cytopathology and molecular testing reports in PDF, RTF and Excel formatted documents. We evaluated two of these resources against 3144 de-identified cases: a cancer/non-cancer classifier resource and a resource for retrieving 12 concepts in breast and lung specific surgical synoptic reports, “Case Number”, “Primary Site”, “Procedure”, “Diagnosis”, “UMLS Code”, “Location, “Laterality”, “IHC Test”, “IHC Result”, “IHC%”, “IHC Intensity”, “IHC Score”. The classifier resource performance was compared to human performance for 1092 cases, consisting of 429 surgical pathology, 619 cytopathology and 44 molecular testing reports. The synoptic report resource performance was compared to human performance for 250 breast and lung cases. Individual points were awarded for each correctly identified concept and negation of concept. Results: The INS cancer classifier resource matched human performance in all 1092 cases, showing 100% accuracy. The synoptic report resource was able to retrieve data in 223 cases, thus showing 89% accuracy in retrieval. Of these, 95% returned the correct diagnosis and IHC data accuracy was 97% overall. Conclusions: INS is able to facilitate rapid development of biomedical information retrieval resources with high accuracy. We find that this system is able to correctly match a high percentage of pathology concepts, using INS architecture to integrate new annotators and tune performance of the negation annotator. We plan to add more general pathology annotators to address the issues that caused matching failures in the current analysis.

A Sandboxed Web Development Environment for Trainee-driven Software Design and Prototyping

Thomas Pearce1, Michelle Heayn1, Russell Silowash1, Michelle Stram1, Nicholas Barash1, Marie DeFrances1, Douglas Hartman1, James Beiber1, Ruth Craig1, Liron Pantanowitz2

1University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: pearctem@upmc.edu

Content: Pathology informatics draws from a broad set of information technology disciplines. This includes web-based, cloud-computing solutions which are being increasingly utilized for medical informatics purposes. Today, web technology is ubiquitous and it provides a relatively accessible entry point into the world of computer programming and user interface design. An abundance of free online tools and tutorials are available, enabling anyone with a browser to begin web development with minimal overhead costs. Our aim was to establish a sandbox server that encourages residents to develop pathology-related applications. Technology: A networked server and database, equipped with a suite of resource management software, was attained. An integrated development environment was installed on a virtual machine to provide users with remote access to software development tools. Design: Desirable characteristics we included in this trainee-driven software development platform were: 1) support for a variety of client-side, server-side and data management tools; 2) full containment behind the institutional firewall, with a mechanism to limit potential access to protected health information; and 3) isolation from other information services resources, to ensure any mistakes in server management by clinical trainees do not have the ability to adversely impact the larger organization. Results: With the support of our Information Services Division, a “sandboxed” web
Implementation of a Whole Slide Imaging Repository and Platform Using Open Source Software for Undergraduate, Graduate, and Continuing Medical Education

Sahr Syed1, Brian Paterson1, Nicole Cipriani1, David McClintock1

1Pathology, The University of Chicago, Chicago, IL, USA. E-mail: Sahr.Syed@uchospitals.edu

Content: Over the past 15 years, whole slide imaging (WSI) has proven itself an invaluable tool in pathology education. At our institution, a need arose for a WSI repository that could satisfy the needs of undergraduate, graduate, and continuing medical education. The goals of this repository were to provide 1) a secure, access-controlled online platform for viewing slides on multiple platforms (iOS, Android, Windows, MacOS, etc.), 2) an ontology for creating use-appropriate metadata for each slide, and 3) an easy-to-use import tool for adding slides to the repository. Finally, our aim was to organize and filter the slides for both directed teaching and open discovery. Technology: Multiple software tools were used, including: VIPS (free image processing system, vips.ecs.soton.ac.uk); OpenSeadragon (web-based viewer for high-reszomable images, openseadragon.github.io); Visual Studio, .NET Framework, and MS SQL Server (Microsoft), Groupier (enterprise access management system) and Shibboleth (enterprise federated identity system). Design: An agile software development methodology was used to create our WSI repository composed of: 1) a database server for course groups, a slide ontology, and guest ticket information; 2) a slide importer allowing WSIs to be imported by different course managers; 3) a slide manager for editing WSI metadata; 4) a slide category manager for organizing course groupings and editing the slide ontology; and 5) a course manager for editing user roles and issuing guest tickets for outside access. Results: Typical use for the WSI platform consisted of authorized users either 1) accessing slides from external course sites (direct links), 2) reviewing course content by filtering for course/section, or 3) reviewing content by body system, organ, condition, sub-condition, species, stain, magnification, or diagnosis. This system, although only very recently implemented, has thus far been received well and has increased both course manager and end-user satisfaction (verbal reports). Conclusion: As compared to vendor-based solutions, we believe that the low development and operating costs, in addition to the flexibility of being able to customize our WSI repository to our undergraduate, graduate, and CME course needs, has vastly improved the WSI experience at our institution.

Peripheral Blood Smear and Bone Marrow Templates

Ngoc Tran1, Rufei Lu2, Christopher Williams3, Teresa Scordino4

1Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma. 2Medical Student, College of Medicine, University of Oklahoma, Norman, Oklahoma. 3Department of Pathology, University of Michigan, Ann Arbor, MI, USA. E-mail: ngocthybao-tran@ouhsc.edu

Content: The microscopic description component of peripheral blood smear (PBSM) and bone marrow (BM) reports is often written in lengthy paragraphs, which may not easily convey important information. Written paragraphs are also time-consuming to compose and, additionally, manifest in a wide variety of reporting styles that are to be read by practicing clinicians and physicians at various levels of training. These aforementioned issues can be reduced with easy-to-use PBSM and BM templates that generate list format reports. Technology: TML (hypertext markup language), CSS (cascading style sheets) and JavaScript were used to design PBSM and BM templates. Design: The PBSM and BM templates are integrated into one HTML page on the end-user side. The input elements, located in the left half of the page, compose of four types: text, checkbox, radio button and submit button. The input contains the entire structure of typical PBSM and BM reports in our institution, as well as comments, most common ICD-10 codes, resident/pathologist signature, and current date. The output, located in the right half, displays the exact appearance of the final report in our laboratory information system so that additional format-related changes are not needed. Elements of the output can also be uniquely edited for convenience. A “Hematology” navigation button allows easy switching from PBSM and BM templates. Subsections of the BM template can be accessed via that navigation button as well to avoid scrolling. Input subsections
Figure 1: Peripheral blood smear and bone marrow templates in hematopathology

can also be minimized so that the text output aligns with its input counterpart. Normal findings were designed to appear as default. Results: The template with user-friendly appearance and easy-to-use features is shown in the Figure 1. Conclusions: Implementing PBSM and BM templates in hematopathology daily work will increase efficiency for pathologists and help clinicians find details more easily. This format also creates a uniform report from different pathologists. This reduces clinician reviewing time and allows pathology trainees to appreciate the most important elements of the final report. Additionally, this synoptic-like reporting will facilitate structured data extraction for future research.

Economics of Cloud-based Graphical Processing Unit

Christopher Williams¹, Ulysses G. J. Balis²
¹Pathology, Michigan Medicine, University of Michigan, Ann Arbor, MI, USA. E-mail: chriswl@med.umich.edu

Content: As whole slide imaging (WSI) enters clinical workflow, the opportunities for integrated image-based analytics will increase. To satisfy the need for increasing scale of high-performance hardware supporting complex algorithms, Graphical Processing Units (GPUs) will become essential. GPUs designed expressly for high performance computing can accelerate processing significantly, but such hardware is significantly more expensive than standard components and harder to support. To mitigate this incremental burden, we propose the utilization of GPU hardware residing in either the local datacenter or the cloud. Technology: We utilize GPU hardware from Nvidia Corporation (Santa Clara, CA) to realize a scaleable and generalizable solution. Operating system virtualization, provided by Docker (San Francisco, CA), allowed for further simplification of the algorithm scale-out process, with this approach enabled by use of the Amazon Web Services platform (AWS; Seattle, WA). Design: A high-performance image analysis algorithm was implemented as a typical web service. Slides were initially stored in AWS and subsequently made available to the algorithm. Requests for analysis from the web client, along with specific user parameters and the specific field of view under interrogation were returned in the form of a superimposed overlay channel of highlighted pixels, in near-real-time, providing decision support information. Results: Execution of high-throughput analytical image-based computational pipelines and associated machine vision analytics on the web, as opposed to use of a local server, is an improved approach that reduces or even eliminates user response delay, while at the same time providing an elevated level of computational capability. When compared to a stationary work station that is heavily utilized, using cloud resources is unquestionably the more expensive option. However, for average computational load settings, cloud resources utilized by multiple users can be a more economical, and thus attractive solution. Conclusions: While there is retained misconception that resources realized in the cloud are inherently less expensive, due to “economies of scale,” we have found that this is not necessarily the case. Nevertheless, convenience and ability to scale to demand offer significant benefits that are worth the cost, in some settings.

Content-based Electronic Templates and Autopsy Reporting: A Perfect Match!

Keren I. Hulkower¹, Dylan V. Miller², Jody E. Hooper², Philip R. Foulis⁴, Christine Bashleben¹, Samantha Spencer¹
¹College of American Pathologists, Northfield, IL, ²Pathology, Intermountain Central Laboratory, Salt Lake City, UT, ³Johns Hopkins University School of Medicine, Baltimore, MD, ⁴Pathology, James A. Haley Veterans Hospital, Tampa, FL, USA. E-mail: khulkow@cap.org

Content: The Autopsy Committee of the College of American Pathologists (CAP) is developing content in response to pathologist requests for electronically implementable templates to be used as a reference and a tool for standardized autopsy reporting. The Pathology Electronic Reporting (PERT) Committee and CAP electronic Cancer Checklist (eCC) Team were challenged with creating data entry forms from this content according to structured data modeling rules.

Technology: The CAP eCC Template Editing tool was used for visual content modeling of question and answer sets for the data entry form, with models and metadata stored in a SQL Server database. A custom .NET tool generated an eCC schema-compatible C# object model to serialize the database records into the eCC XML format. Additional XSLT programs transformed the XML files into eCC HTML files suitable for data entry form review and incorporation into eFRM software (mTuitive, Centerville, MA). Design: The eCC Team and the Autopsy Committee members developed question and answer set content in an appropriate relational format for structured data capture and reporting. We reviewed structured data capture rules, item types, and metadata attributes to reach a consensus for modeling specific use-case sections for external and internal...
exams. Results: A proof-of-concept adult gross autopsy template for synoptic reporting was created using structured data modeling principles. Initial steps involved determining the template content using standard anatomic terms so that relationships between question and answer sets (QAS) and metadata attributes were understood for eCC modeling. This included designating questions as single or multi-select, and as core, conditional, or optional. Instructional notes to pathologists were added to provide additional information to assist with completing data entry forms. Metadata requirements were captured for question fill-in and answer fill-in items, especially units of measure, min/max values, and decimal places. Conclusions: Understanding and providing standardized guidance regarding item type options, applied metadata attributes, and modeling rules to content experts improved content development and layout of standardized electronic data entry forms for autopsy reporting.

Transition of Tumor, Node and Metastasis Staging Definitions in College of American Pathologists Electronic Cancer Checklists to Reflect American Joint Commission on Cancer Eighth Edition Content

Keren I. Hulkanow1, Jaleh Mirza1, Richard Moldwin1, Sanjeev Baral1, Thomas P. Baker2, Michael A. Berman3, Philip R. Foulis4 and Samantha Spencer1

1College of American Pathologists, Northfield, IL; 2Pathology, The Joint Pathology Center, Silver Spring, MD; 3Jefferson Hospital, Allegheny Health Network, Jefferson Hills, PA; 4Pathology, James A. Haley Veterans Hospital, Tampa, FL, USA. E-mail: khulkow@cap.org

Content: The College of American Pathologists (CAP) electronic Cancer Checklist (eCC) Team is transitioning the eCC templates to reflect the updated American Joint Commission on Cancer (AJCC) 8th Edition tumor, node and metastasis (TNM) prognostic staging classification content for implementation on January 1, 2018. Technology: The CAP eCC Template Editing tool was used for visual content modeling of question and answer sets and metadata for the data entry form, with models and metadata stored in a SQL Server database. A custom .NET tool generated an eCC schema-compatible C# object model to serialize the database records into the eCC XML format. Additional XSLT programs transformed the XML files into eCC HTML files suitable for data entry form review and to compare original and revised template XML file versions as a QA review. Design: The eCC Team reviewed changes between the AJCC 7th and 8th edition definitions of TNM prognostic staging categories. A table was prepared for each template to facilitate the comparison of TNM definitions between editions. Results: The eCC Team has over 50 templates that require updates to TNM categories to comply with the AJCC Cancer Staging Manual 8th edition. Changes in TNM definitions fall into three categories, each requiring separate modeling rules to maintain the process by which staging results are queried across different editions. These include: 1) visible text changes to definitions that do not affect semantics, 2) elimination of definitions which require deprecation of both the answer item and parent question, and 3) new definitions that must be added along with new parent questions. Additionally, the eCC Team faced the challenge of a Q1 2017 interim update to the histologic type terminology for genitourinary resection templates according to latest WHO standards while maintaining the current AJCC 7th edition staging definitions. This was achieved by storing 8th edition definitions as hidden content in the Template Editing tool without displaying them in the XML files. Conclusions: Updating eCC template content to comply with AJCC 8th edition TNM prognostic staging definitions required the adoption of different modeling strategies. This ensured the preservation of staging queries across different AJCC editions.

Case Identification through Full-text Analysis of Pathology Reports

Wade Schulz1, Thomas Durant1, Peter Gershkovich2

Departments of 1Laboratory Medicine and 2Pathology, Yale School of Medicine, New Haven CT 06520, USA. E-mail: wade.schulz@yale.edu

Content: Biomedical research often relies on the identification of patient cohorts that meet specific pathologic criteria. Similar queries are also used to identify cases for quality review and trainee education. LIS-based full-text searching capabilities allow only basic queries and often degrade performance when used directly on production databases. On another hand, specialized search engines require significant investment of time and resources. Modern indexing and searching technologies along with the growing set of Unified Medical Language System (UMLS) tools may present a lightweight alternative to fill the gap. Technology: We assessed two approaches for the text-based search of historical records: CoPath (Cerner, Kansas City, MO) and Elasticsearch (Elastic, Mountain View, CA), a distributed, NoSQL search platform. UMLS Metathesaurus was used for concept-based tags within Elasticsearch. In addition, clinical laboratory data was obtained from the Cloverleaf integration engine (Infor, New York, NY) to allow for joined queries that included laboratory data. Design: Pathology reports were extracted from CoPath and loaded into Elasticsearch. A method to apply basic concept-based search based on terms in UMLS Metathesaurus was integrated into a custom Python/Elasticsearch platform. A set of selected queries based on common diseases was run on
Divergent Requirements for Tumor DNA Sequencing Software

Wade L. Schulz1, Thomas Durant1, Neil Mutnick1, Peter Gershkovich2

Departments of 1Laboratory Medicine and 2Pathology, Yale School of Medicine, New Haven, CT 06520, USA. E-mail: wade.schulz@yale.edu

Content: The absence of Next Generation Sequencing (NGS) support in traditional laboratory information systems (LIS) has forced many laboratories to develop custom software to streamline NGS workflow, data analysis, and reporting. There is hope among pathology groups that LIS vendors will pick up speed and create integrated software modules that can broadly address the varying needs of NGS-based assays. However, rapid changes in molecular medicine and divergence of user requirements continue to foster the development of isolated systems for NGS support. Technology: Anatomic Pathology and Laboratory Medicine at Yale created three separate standalone applications to support tumor DNA sequencing within three separate labs. All systems are web-based and build using either the Java programming language or .NET framework. One system in Anatomic Pathology had been integrated with CoPathPlus LIS and another with NCI-Molecular Analysis for Therapy Choice program, while Laboratory Medicine is integrating with the Beaker LIS. Design: All systems incorporate common functionality that assist bioinformaticians in uploading genomic variants and help molecular pathologists to access previously discovered variants, view patient information, choose from a list of variant annotation archetypes, etc. At the same time, the analytic workflow, loading of data, presentation of variants, annotation mechanisms, and, ultimately, reporting of results, differ sufficiently enough that separate implementation and support were required. Results: Three different systems have been developed to support tumor DNA sequencing in one institution and continue to evolve addressing different, and continuously divergent needs of their respective user base. We identified that the differences are mainly related to bioinformatics pipeline management, systems integration, and reporting of results. To assist with the integration of clinical variant data, we also developed a JSON-based data format that can be used to exchange and store results. Conclusion: Despite the fact that tumor DNA sequencing is performed using similar platforms and despite the fact that the ultimate goal of all three systems is to reduce the complexity of clinical reporting, there are multiple divergent requirements that require different implementation. Persistent pressure to add, remove, or modify existing features supports the approach of rapid and agile development of segregated systems.

Logic and Complexity of Electronic Alerts in Pathology

Peter Gershkovich1, Thomas Durant2, Wade Schulz2
Departments of 1Pathology and 2Laboratory Medicine, Yale School of Medicine, New Haven, CT 06520, USA. E-mail: peter.gershkovich@yale.edu

Content: LISs often provide basic mechanisms to create rule-based alerts associated with predefined logic for ordering tests, pathological findings, and workflow events. However, the LIS-based mechanisms for alerts rely on primitive logic and lack the ability to integrate with external systems. Such alerts are hard to refine and evaluate regarding users’ reactions. Accordingly, there are often no configuration mechanisms in a common LIS to create the simple task of sending an email or text message to a pathologist when slides for a STAT case are ready or to further escalate such an alert, based on elapsed time). Technology: We developed the Repetitive Task Scheduling Engine (RTSE) – a Java-based, in-house solution – to create alerts with associated functionality for monitoring and escalation mechanisms. RTSE is a Web-based application that runs on the Apache Tomcat server. The open source Quartz Scheduler is the core component of RTSE. We also used a MySQL database to store data and configuration parameters. Design: Individual tasks in RTSE are defined as “jobs”. Each job encapsulates processing logic and has its own trigger and configuration parameters. Jobs rely on common utility classes and share the data model that is largely borrowed from the LIS. Jobs are independently executed in their own threads and do not interfere programmatically. A web interface is used to control and observe the execution of these jobs. Results: The system demonstrates a method of encapsulating complex logic in a separate, custom-built system. Since the alerts are configured through the use of a programming language, we are able to create very specific and highly nuanced electronic alerts. In addition, it provides capabilities to monitor user reactions to messages and can escalate alerts based on elapsed time and other discretionary criteria. Conclusion:
Electronic alerts in clinical settings are ubiquitous. At the same time, the fatigue associated with inundating, nonspecific, and misdirected alerts may offset their utility. It is important to realize that the accuracy and precision of alerts require complex programming logic and often needs external data. The RTSE system at Yale demonstrates the complexity of electronic alerts and offers a viable solution to address the intricacy of accurate, electronic notifications.

**Multivariate Modeling of Urine Culture Results Based on Urinalysis Parameters**

Michelle Stram1, Jansen Seheult1, Tom Pearce1, William Pasculle1

1Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: strammn@upmc.edu

**Content:** Urinalysis is often the first test in the workup of a possible urinary tract infection. Urine culture may be ordered concomitantly or as a reflex test if the urinalysis results suggest possible infection. The purpose of this study was to develop an algorithm for a urinalysis reflex to urine culture protocol, to reduce the number of urine cultures performed. **Design:** Results from urinalyses (Sysmex UF-1000i, Sysmex America, Lincolnshire, IL, and Clinitek Atlas, Siemens Washington, DC) and urine cultures were extracted from the laboratory information system from Nov. 1, 2016 through Feb. 11, 2017. Only patients who had a urinalysis and urine culture performed on the same day were included in the analysis. **Technology:** Statistical analysis was performed in R v3.3.3 (R Core Team, Vienna, Austria) and Stata v14 (Statacorp, TX). Four urinalysis parameters (bacteria, leukocyte esterase, nitrite, white blood cells (WBCs)) were assessed for their ability to determine urine culture positivity. Logistic regression analysis, receiver-operating curves and recursive partitioning classification tree analysis (complexity parameter of 0.01) were done for exploratory analysis. An algorithm was developed to generate confusion matrices for all logical combinations of cutoffs for the urinalysis parameters, utilizing a leave one out cross-validation. Model selection was based on maintaining sensitivity above 90% and maximizing specificity. **Results:** The logistic regression prediction area under the curve was 0.8299 using all four factor variables. The developed algorithm identified 10,578 different combinations of the four urinalysis variables and cutoffs. A subset (337) of models had a 90% or higher sensitivity and specificity greater than 40% [Figure 1]. The criteria for the model with the best performance were: bacteria="Many" OR Leukocyte Esterase="Large (P3)" OR WBCs≥11 (sensitivity=90.8%, specificity=52.3%, positive predictive value=69.4%, negative predictive value=82.6%). The pruned recursive classification tree failed to generate a better model (maximum sensitivity=86%, specificity=53.2%). **Conclusion:** We developed an intuitive algorithm to consider all possible combinations of urinalysis parameters and cutoffs in order to predict urine culture positivity. The selected model should allow an almost 50% reduction in the number of negative urine cultures performed by the laboratory, with a false negative rate of less than 10%. Prospective validation will be performed on the top five models identified.

![Figure 1: Classification models for urine culture prediction](http://www.jpathinformatics.org)
A Genomic Information System: A New Information System Architecture for Molecular Pathology and the Integration of Genomic Data into the Electronic Health Record

Chesley Leslin

Bioinformatics, Northeastern University, Standard Molecular, Inc. E-mail: chesley@standardmolecular.com

Content: As Molecular Pathology and next generation sequencing transitions from the academic laboratory to clinical practice there is a need for purpose-engineered tools to meet the systems integration, workflow management, variant annotation, analysis, and signout needs of the clinical molecular pathology laboratory. Furthermore, once the molecular pathology report is signed out, it is crucial that molecular results are integrated with the electronic health record AND clinical data repository as discrete data, enabling clinical decision support, population health studies, pharmacogenomics, and clinical trials recruitment. Technology: Continuity GIS is a Java back-end application, with browser-based front-end client, a PostgreSQL database, and HL7/XML interface capabilities. The application is initially an on-site data center-installed system, but the architecture is cloud-compatible. Design: The system has the following components: 1. User Interface for sign-out 2. User Interface for the interpretation knowledgebase 3. Gateways for communication: Bi-directional Anatomic Pathology, Bi-directional Sequencing LIMS, Inbound Sequence Pipeline, Outbound to EHR: PDF & Discrete Data, and Outbound Data Repository PDF & discrete data including full variant file. Results: As one of the first tools to integrate discrete molecular pathology results with the electronic health record, GIS enables personalized molecular medicine. Surgeons, internists, and oncologists have access to discrete genomic data in the EHR, enabling genetics-based clinical decision support and clinical trials recruitment. Conclusion: Molecular pathology as a separate silo, in laboratory medicine, apart from the EHR represents a severe limitation on the clinical utility of genetic sequencing, and is not in alignment with health-systems’ long-term goals of integrated patient genomic data. As next generation sequencing transitions from niche test to broadly utilized diagnostic, the tools supporting workflow and data integration must meet the demands of clinical practice. With the development of dedicated Genomic Information Systems, molecular diagnostics are poised to follow previous diagnostic regimes (MRI/Digital Radiography-PACS, Digital Angiography-PACS, and Anatomic Pathology Informatics) originally developed in academia as they are adopted for broad clinical use with data continuity throughout the electronic health record.

Living Biobank: Informatics Methods to Improve Efficiency of Obtaining Clinical Specimens for Medical Research

Christopher L. Metts, James E. Madory, Steven L. Carroll, Andrew M. Cates, Doug Williams, Leslie A. Lenert

Pathology, Medical University of South Carolina, Charleston, South Carolina, USA. E-mail: mettsch@musc.edu

Content: Biospecimen demand is growing at a rate of fifteen percent per year, driven primarily by precision medicine research and associated biomarker discovery and validation. It is becoming clear that current manual methods of obtaining clinical specimens for research do not allow for the economics of scale necessary for cost-effective precision medicine research. In this project, we attempt to design an informatics approach to automate the process, including informing potential study participants of research opportunities, screening for appropriate populations, obtaining informed consent for data and specimen collection, and configuring electronic medical record systems, pathology laboratory information systems, and laboratory robotics for specimen retrieval. Technology: REDcap (Application and API), ResearchKit, iOS, Google App Engine, Objective-C, Python. Design: Potential study participants are encouraged to download our mobile app to learn about the project. Interested participants verify their identity through interactions with the Health Sciences South Carolina’s FHIR interface with a state-wide master patient index of four million individuals. Upon successful verification, participants are screened and informed consent is obtained.
for collection of clinical data and specimens. Participants completing this process will have a research ID assigned and registered with iSpecimen, a marketplace where bioresearchers indicate the type of specimens they need for their research. The research ID is also registered with our REDCap project and Epic, such that subsequent blood draws for laboratory testing at participating hospital will result in an extra tube being drawn for research purposes if the phenotype matches the needs of a researcher registered with iSpecimen. 

Results: We have designed an informatics approach to automate the process of obtaining clinical specimens for medical research. The process preserves participant privacy. 

Conclusions: Informatics methods can automate many of the steps traditionally handled manually to obtain clinical specimens for research purposes. Further development is needed to complete the workflow and pilot the process at our institution.

### Computational Pathology for Promoting Diagnostic Concordance of Atypical Ductal Hyperplasia Detection in Breast Biopsies

Akit Burak Tosun1, Luong Nguyen1, Nathan Ong1, Olga Navolotskaia2, Jeffrey Fine2, Gloria Carter2, D. Lansing Taylor2, S. Chakra Chennubhotla1

1Department of Computational and Systems Biology, University of Pittsburgh, 2Department of Pathology, Magee Womens Hospital of UPMC, University of Pittsburgh, 3Department of Computational and Systems Biology, Drug Discovery Institute, University of Pittsburgh, Pittsburgh, PA, USA. E-mail: tosun@pitt.edu

Content: Atypical Ductal Hyperplasia (ADH) diagnosis is an important source of discordance and uncertainty among pathologists. Accurate diagnosis is crucial as ADH is associated with an increased risk of invasive breast cancer development. Using computational pathology, we would like to begin a standardization process that delineates key quantifiable measurements, such as roundness and monomorphism of nuclei and architectural patterns such as uniform between-nuclei spaces. Herein we propose a computational pipeline for detection of ADH, using such features that would serve as an intelligent, expert guide for pathologists. This should increase diagnostic concordance and lead to faster, more accurate diagnostic decisions. 

Technology: Whole slide images (WSIs) were created from 93 slides from breast biopsies with ADH at 0.5 microns per pixel resolution (Aperio ScanScope XT, Leica Biosystems, Buffalo Grove IL). Computational pathology pipelines analyzed entire WSIs that were typically 30k x 30k pixels, using available cloud computing resources within our institution. A web-based tool was used to display a total of 1759 regions of interest (ROIs) to pathologists. 

Design: Our pipeline includes: WSI color normalization; WSI segmentation for ROIs with ducts; and detection of ROIs with ADH by encoding cellular morphology. Based on opponent color spaces, our color normalization method is scaled to work with large WSIs. Spatial density of epithelial nuclei were used to segment ducts in the WSIs. To encode cellular morphology, we computed nuclear features of segmented nuclei from the duct ROIs. Finally, feature selection was performed to classify the ROIs as either ADH or Non-ADH.

Results: Our end-to-end computational pathology pipeline required about 5 minutes per WSI. 1009 ROIs out of 1759 were analyzed by most experienced pathologist (P1), forming the training set, the remaining 750 labeled by all three expert pathologists (P1, P2, and P3), forming the consensus test set. Computational results on 93 WSIs of breast tissue images showed good performance on consensus of expert pathologists with an F-score of 0.89, where average F-score for single pathologist vs. consensus is 0.54. 

Conclusions: Our experiments highlight the challenge of diagnosing ADH in clinical world, where we observed only 10% unanimous concordance of three pathologists on ADH ROIs. One main reason behind this discordance is the subjectivity in the definition of atypia that may lack reproducibility of diagnosis. This study is the first attempt in improving the diagnostic concordance of ADH by computational methods. It is very likely that a combined approach of more specimens (i.e. more ADH or other atypical examples), larger numbers of pathologists, and consensus diagnoses would improve the reliability and usability of ground truth in developing computational methods.

### Open Algorithm Standard

Christopher Williams1, Ulysses Balis1

1Pathology, Michigan Medicine, University of Michigan, Ann Arbor, MI, USA. E-mail: chriswi@med.umich.edu

Content: Image analysis algorithms for pathology images tend to be either standalone applications or embedded solutions within proprietary platforms. As whole slide imaging integrates into pathology workflow, there will be an expanding need for integrated algorithms within the viewing platform. The current model is underdeveloped and inefficient, in that each integration represents a vertical solution, which must be individually supported. To facilitate algorithm interoperability, we propose adoption of an Open Algorithm Standard (OAS) to define a common application programming interface. The resultant “plugin architecture” supporting image-based analytics will benefit investigators, algorithm developers and digital pathology vendors alike. 

Technology: The key to supporting a heterogeneous mix of architectures is an abstraction layer, allowing for isolation of the algorithm from the underlying operating system and dependencies. OAS leverages operating system level virtualization, colloquially referred to as containerization, to accomplish this. Containers traditionally have exclusively supported the Linux operating system. However, one of the largest providers, Docker (www.,
Extraction of Clinically Useful Laboratory Data from Digitized Red Blood Cell Histograms Produced by an Automated Hematology Analyzer

Amrom E. Obstfeld1, Priya Velu1

1Pathology, Children's Hospital of Philadelphia, Philadelphia, PA, USA.
E-mail: obstfelda@email.chop.edu

Content: The RBC histogram produced by hematology analyzers is used by the instrument to calculate the mean corpuscular volume (MCV) and hematocrit of a sample. By extracting the data underlying this plot, it may be possible to infer additional actionable information. The two aims of this study were to (1) assess the feasibility of extracting this information in an accurate manner, and (2) determine whether the method may be applied to produce clinically meaningful information beyond the MCV of a sample. Technology: Screenshots of histograms were copied from the Sysmex XN computer and data was extracted using an online application, WebPlotDigitizer. Further analysis was performed using the tidyverse and mixtools packages within the R statistical program. Design: Twenty histograms were digitized, and the MCV calculated and compared with the original XN analyzer MCV using linear regression. Two additional histograms were digitized from patients who had bimodal RBC populations. The MCV of the entire RBC population, and the MCV of each individual population was obtained. A mixture model was used to estimate the mixing weights, the proportion of RBC in each of the individual populations. Results: The MCV calculated from the digitized histograms compared well to the analyzer MCV with an adjusted R² of 0.98, a slope of 0.9, and a P-value of 9.9 * 10⁻¹⁸. In samples without unimodal RBC populations the analyzer algorithm fails and the MCV results are unreliable. To assess whether these methods can be used in this scenario two bimodal RBC histograms from patients with severe iron deficiency anemia treated with transfusion were obtained. The calculated MCV of these samples replicated the analyzer MCV. Using the digitized histogram data, the MCV of each individual RBC population was calculated. Finally, using the mixtools package we were able to produce reasonable estimations for the proportion of the two RBC populations in the patients. Conclusion: Analysis of raw CBC data produced by hematology analyzers, in this case with a digitization step, may provide clinically useful information for patient care, especially in patients who have multiple RBC populations that result in spurious MCV results reported by standard analyzers.

Three-dimensional Reconstruction and Observation of Atypical Mitosis

Kazuhiro Tabata1,2, Jiujun Fu1, Yukako Yagi1

1Department of Pathology, Pathology Digital Imaging, Memorial Sloan Kettering Cancer Center, New York, USA, 2Department of Pathology, Nagasaki University Hospital, Nagasaki, Japan. E-mail: tabatak@mskcc.org

Content: Histological decision of mitosis plays a great role in pathological diagnosis, especially, atypical mitosis is well known as good deciding factor for malignancy. Multipolar mitosis is popular as typical picture of atypical mitosis, however, it is merely observed that the figure of atypical mitosis on the histological specimen is one of cross section in steric structure. It is difficult to observe entire image of multipolar mitosis in detail with existing pathological technique. The purpose of this study is detecting of atypical mitosis using 3 dimensional histological reconstructed imaging. Technology: Automated sectioning machine, AS410 (Dainipponseiki, Kyoto, JPN); Whole slide imaging scanner, Aperio AT2 (Leica, IL, US); Software, Voloom (MicroDimensions GMBH, Munich, GER). Design: We prepared 50 specimens of lung carcinoma tissue in formalin-fixed paraffin embedded block. They were sectioned thickness of 5um sequentially by automated sectioning machine. After hematoxylin and eosin staining, all histological slides were digititized to whole slide images in 20x (0.5um/pixel) by Aperio AT2 scanner. We observed the 3 dimensional image...
Results: Following standardized protocols for prospective case selection, enrollment, validation and intra-system, inter-system, and inter-site precision assessment, we report the agreement rates for three arms of the study: intra-system (92.0%), inter-system (93.8%), and inter-site (90.2%). The results demonstrate that evaluation of digitized slides is highly reproducible within and across systems as well as laboratories. Conclusion: These data make an important contribution towards regulatory clearance of digital pathology and provide a benchmark against which individual laboratories can compare their digital performance. With a precise medical device in place, further efforts to drive digital adoption in pathology can now focus on achieving financial sustainability.

Comparison of Different Digital Image Resolutions for Identifying Acid Fast Bacilli

Matthew Hanna1, Liron Pantanowitz2
1University of Pittsburgh Medical Center, Pittsburgh, PA, USA.
E-mail: hannamg2@upmc.edu

Content: Image resolution is essential for the reliable interpretation of digital images in pathology. Advancing technology continues to enhance sensors (e.g. 4K microscopy cameras), offering us ultra high resolution imaging. The diagnosis of microscopic features by whole slide imaging versus conventional light microscopy has demonstrated high concordance. However, the ability to reliably resolve microorganisms with whole slide images has been problematic. This study aimed to compare several image resolutions for the optimal identification of acid fast bacilli (AFB). Technology: Images were acquired using different microscope-mounted digital cameras (Spot Insight 2, Prosilica GT, Olympus UC90, Olympus DP27) and at 0.25 micron/pixel with whole slide scanners (Aperio ScanScope XT, Sakura VisionTek M6) [Table 1]. Design: 10 surgical pathology glass slides with Ziehl–Neelsen stained tissue sections containing positive AFB were selected. They were scanned using two different WSI scanners at 40x (0.25micron/pixel). Glass slides were also used to manually capture images of positive AFB at 40x magnification using four different digital cameras each mounted to an Olympus BX45 microscope. Acquired images from identical fields

Table 1: Comparison of digital camera sensors

| Sensor type       | Olympus UC90 | Olympus DP27 | Prosilica GT | VisionTek M6 | Spot Insight 2 | ScanScope XT |
|-------------------|---------------|---------------|--------------|--------------|----------------|--------------|
| Megapixels        | Color CCD     | Color CCD     | Color CMOS   | Color CMOS   | Color CCD      | Color CCD    |
| Resolution (max)  | 3384×2708     | 2448×1920     | 1920×1200    | 1876×1150    | 1600×1200      | 1712×1150    |
| Pixel size (μm)   | 3.69×3.69     | 3.45×3.45     | 5.86×5.86    | 5.5×5.5      | 7.4×7.4        | 14.0×14.0    |
| Camera size (cm)  | 8.3×8.1×4.4   | 7.7×6.9×4.25  | 3.0×2.9×2.9  | 9.2×5.3×3.3  | 14.2×9.5×7.1  | 3.81×6.2×6.2 |

of atypical mitosis which was reconstructed by Voenum, and search the figure of atypical mitosis. Results: We found some large cells larger than 30 microns in diameter with atypical mitosis using 3 dimensional reconstructed images easier than using single slide images. Thereafter, we extracted 15-20 whole slide images including atypical mitosis with surroundings and reconstructed 3 dimensional images of atypical mitosis. 3 dimensional images helped to detect atypical mitosis easier than using single slide images. And it was able to observe that multipolar transferring of chromosomes in 3 dimensional reconstruction image even though it had some degrees of gap and malalignment. Conclusions: 3 dimensional image reconstructed from multiple whole slide images of hematoxylin and eosin staining enabled us to observe the steric structure of atypical mitosis.
Utility of Google Image Search for Diagnostic Surgical Pathology

Matthew G. Hanna¹, Liron Pantanowitz²
¹University of Pittsburgh Medical Center, Pittsburgh, PA, USA.
E-mail: hannamg2@upmc.edu

Content: Content-based image retrieval is a powerful application of computer vision. Currently, there are limited algorithms available that perform content-based image retrieval in anatomic pathology. Reverse image lookup using the Google Image Search engine (GISE) is a popular mechanism to identify images based on a similar digital image. Our aim was to determine the diagnostic accuracy of using GISE for histopathology.

Technology: Reverse image search using Google (https://images.google.com) and Google Chrome internet browser Figure 1. Design: A control set of 11 well-known landmarks (e.g. Eiffel tower) was used to of view for all imaging modalities were compared. Pathologists blinded to the capture device reviewed each image and subjectively ranked (highest to lowest) their diagnostic confidence for detecting AFB. Objective image quality data (noise, luminosity, acutance, color, saturation) were recorded with Adobe Photoshop 2017 and the open source software VQEG Image Quality Evaluation Tool (VIQET). Results: All imaging modalities were able to capture diagnostic AFB fields of view. Figure 1 illustrates the diagnostic confidence of each imaging modality in descending order. Of note, two false positive AFB images on WSI were correctly interpreted only on the UC90 4K microscopy camera. Subjective image quality for viewing AFB was ranked (in descending order) as follows: Proscilla GT, Olympus UC90, VisionTek M6, Olympus DP27, Spot Insight 2, and Scanscope XT. Conclusions: This study demonstrates that microscopic pathology image quality directly correlates with camera sensor resolution (i.e. megapixels) and that high resolution images are better for confidently diagnosing AFB in histopathology slides. We recommend using WSI scanners with high resolution sensors or 4K microscopy cameras when imaging slides for the purpose of detecting mycobacteria.

Table 1: Examples of the top three related images returned using Google image search

| Original diagnosis (histopathology) | Google diagnosis (image search only) | Adding “Cancer” term | Adding “Organ” term | Adding both “Organ” and “Cancer” terms |
|------------------------------------|--------------------------------------|----------------------|---------------------|----------------------------------------|
| Invasive ductal carcinoma           | Granite                              | Serous ovarian carcinoma | Invasive ductal carcinoma | Invasive ductal carcinoma |
| Prostate carcinoma                  | Musical organ                        | Lung adenocarcinoma    | Prostate carcinoma    | Prostate carcinoma               |
| Clear cell renal carcinoma          | Twisted rope                         | Clear cell renal carcinoma | Clear cell renal carcinoma | Clear cell renal carcinoma |
| Giant cell tumor                    | Flowers                              | Osteosarcoma           | Giant cell tumor      | Giant cell tumor                 |
| Osteosarcoma                        | Roses                                | Breast carcinoma       | Bone fracture         | Bone fracture                     |
| Leiomyoma                           | Flowers                              | GIST                  | Leiomyoma            | Leiomyoma                        |
| Melanoma                            | Soil                                 | Serous ovarian carcinoma | Melanoma             | Melanoma                          |
| Meningioma                          | Fur                                  | Meningioma            | Meningioma           | Meningioma                        |

GIST: Gastrointestinal stromal tumor

Figure 1: Diagnostic confidence of acid fast bacilli images for each imaging modality (larger area represents higher confidence). The ranking list shown to the right is in descending order of confidence.

Figure 1: Examples of Google reverse image search. (a) Target image search (without descriptors) returned; (b) floral pattern fabric (target image: Mucinous adenocarcinoma); (c) Target image of leiomyoma searched with descriptor (“uterus”) accurately matched to an online image of, (d) leiomyoma
assess baseline accuracy. Representative screenshots were subsequently captured from whole slide images of various benign histology (n=14) and neoplasms (n=23). File names excluded diagnoses. Each image was searched with GISE and the best match recorded. Searches were refined by adding anatomic location and tumor type as search terms. Images from “tumor of unknown origin” cases (n=23) with known molecular studies were also searched. Results: Normal histology accurately matched 50% of images. Table 1 provides examples of retrieved histopathology image based diagnoses. Adding tumor descriptors, organ sites, and their combination accurately matched 4%, 43%, and 83%, respectively of cases using the top 3 most likely images to corresponding diagnoses. Tumor of unknown origin case searches returned only 13% of pathology diagnoses that matched molecular findings when adding tumor type. Conclusions: Employing GISE to find related images on the web using histopathology images alone is unreliable to make a histopathology diagnosis. However, adding organ and the term “cancer” with an image search increases diagnostic specificity by allowing the search engine to narrow retrieved image results. Future algorithms that allow content-based image retrieval of curated pathology image databases will likely play a greater role in computer assisted diagnosis for anatomic pathology practice.

Evaluation of Diagnostic Concordance between Manual Mitotic Figure Counting on Glass Slides versus Whole Slide Images

Matthew Hanna1, Juan Xing1, Sara E. Monaco1, Douglas Hartman1, Liron Pantanowitz2

1University of Pittsburgh Medical Center, Pittsburgh, PA, USA.
E-mail: hannamg2@upmc.edu

Content: Cancer patient outcomes and survival are usually contingent on disease grade and stage. Pathology staging criteria are often dependent on mitotic counts for neoplastic conditions. With digital pathology on the near horizon, validation of diagnostic concordance has been a large focus in the literature, however little attention has been given to manual mitotic counting with whole slide images (WSI). This study aims to compare visual detection, time, and clinical concordance of mitotic counting for sarcomas and neuroendocrine tumors between glass slides and WSI. Technology: Glass slides were scanned on an Aperio ScanScope XT. WSI files were stored in a storage area network hosted by a VMware Virtual Machine (Windows Server 2008 R2 Enterprise). Image data collection was performed on ImageScope (v12.3) software using Hewlett-Packard 32-bit computers. Design: Glass slides of sarcomas (n=50) and neuroendocrine tumors (n=30) were selected to reflect a range of low to high mitotic counts. Using an Olympus BX51 microscope with a 40x/0.65 PLAN objective and 10x/22 eyepiece, the highest mitotically active focus of tumor was marked with ink on these glass slides. Two board certified practicing pathologists were asked to identify all mitoses in the designated area for each slide. Glass slides were scanned on an Aperio ScanScope XT at 40x (0.25 µm/pixel). After a two-week washout period, the pathologists were given the equivalent digital slide and asked to evaluate the same area and generate mitotic counts for each case [Figure 1]. Results: Pathologists’ review of the replicate tumor foci on glass slides and WSI generated an average of 2936 and 2706 total mitoses, respectively. Time to generate mitotic counts in a 2mm2 tumor area averaged 41 seconds and 148 seconds for glass and WSI modalities, respectively, per case. Tumor-specific mitotic grade differed for glass slide and WSI mitotic counts [Table 1]. Compared with the original diagnosis grading schema, glass slide counts were discordant in 12 (15%) of 80 cases, whereas

| Grade* | Glass slide (%) | WSI (%) |
|---------------|-----------------|--------|
| Sarcoma (n=50) | | |
| 0-9 | 14 (28) | 18 (36) |
| 10-19 | 9 (18) | 10 (20) |
| ≥20 | 27 (54) | 22 (44) |
| NET (n=30) | | |
| <2 | 13 (43) | 14 (47) |
| 2-20 | 12 (40) | 13 (43) |
| >20 | 5 (17) | 3 (10) |

*Criteria based on FNCLCC for sarcomas, and WHO criteria for NET. NET: Neuroendocrine tumor, FNCLCC: French federation of cancer centers sarcoma group, WHO: World Health Organization, WSI: Whole slide image

Figure 1: Representative whole slide image of a sarcoma showing annotation of mitotic figures. Pathologist mitotic figure annotations (green crosshairs) shown inside framed tumor area (black ink marker) with designated 2 mm2 area (white rectangle). Inset: High power view of mitotic figure annotated by pathologist
WSI counts differed in 8 (10%) cases. Conclusions: WSI have sufficient resolution to adequately visualize and assess mitotic figures, comparable to glass slides. Digital counting of mitotic figures outperformed counts based off glass slides. However, counting of mitotic figures was 3.6 times longer using WSI, albeit that manual annotation time may be a confounding factor. Therefore, reliance on automated mitotic figure image analysis algorithms instead of pathologists may be required to perform mundane tasks such as counting mitotic figures.

**Automated Vascular Smooth Muscle Segmentation, Reconstruction, Classification and Simulation on Whole-slide Histology**

Yiwen Xu1, J. Geoffrey Pickering1,2, Daniel Goldman1, Zengxuan Nong2, J. Sachi Elkerton1, Aaron Ward1,3

1Department of Medical Biophysics, Western University. 2Robarts Research Institute, Western University. 3Department of Oncology, Western University, London, ON, Canada. E-mail: yxu225@uwo.ca

Content: Histology of the microvasculature depicts detailed characteristics relevant to tissue perfusion. One important histologic feature is the smooth muscle component of the microvessel wall, which is responsible for controlling vessel caliber. Abnormalities can cause disease and organ failure, as seen in hypertensive retinopathy, diabetic ischemia, Alzheimer’s disease and improper cardiovascular development. Technology: We have developed a software platform for automated (1) 3D vascular reconstruction, (2) detection and segmentation of muscularized microvessels, (3) classification of vascular subtypes, and (4) simulation of function through blood flow modeling. Design: Vessels were stained for α-actin using 3,3’-Diaminobenzidine (Vector Laboratories, Burlingame, CA, USA), assessing of function through blood flow modeling. Results: A Dice coefficient of 0.88 (compared to manual) for the segmentations, a 3D reconstruction target registration error of 4μm, and area under the receiver operator curve of 0.89 for vessel classification. We found 24% and 18% decreases in the blood flow through the network for the regenerated vasculature during increased oxygen demand as compared to the normal vasculature, respectively for 14 and 28 days post-ischemia. Conclusion: We developed an automated system to assess the arteriolar smooth muscle in the microvasculature, which allows for high throughput analysis of digital histology images and flow simulation. With microvascular measure visualization methodologies in 3D and automated segmentations, we are now capable of locating focal pathologies on a whole slide level using 3D histology reconstruction, and performing separate analyses on the arteriolar side of the microvascular tree.

**Digital Pathology for Automated Quantification of Tumor Microenvironment for Metastasis and MenaCalc: Biomarkers for Metastasis**

David Entenberg1,2, Joseph Sparano3,4, Maja Oktay1,2, Timothy D’Alfonso5, Brian Robinson6, Paula Ginter7, Thomas Rohan1, Xiaonan Xue5, John Condeelis1,2, Joan Jones1,2

1Departments of 1Anatomy and Structural and 2Biology, Gruss Lipper Biophotonics Center, Integrated Imaging Program, Montefiore Medical Center, Einstein College of Medicine, Departments of 3Medicine and 4Obstetrics and Gynecology and Women’s Health, Montefiore Medical Center, Einstein College of Medicine, 5Department of Pathology and Laboratory Medicine, Montefiore Medical Center, Einstein College of Medicine, 6Department of Epidemiology and Population Health, Einstein College of Medicine, Bronx, 7Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, USA. E-mail: David.Entenberg@einstein.yu.edu

Content: Our group has discovered and clinically validated two mechanistically linked biomarkers for metastatic progression: MenaCalc, a multiplex immunofluorescence assay quantifying isoforms of the protein Mena (panMena–Mena11a), associated with the ability of tumor cells to exhibit resistance to chemotherapy, and disseminate from the primary tumor; and TMEM, a microanatomic structure composed of the direct juxtaposition of a Mena overexpressing tumor cell, a macrophage, and an endothelial cell. TMEM is the portal through which MenaCalc high tumor cells disseminate hematogenously. Technology: We use digital pathology
and whole slide analysis to mimic pathologists’ workflow and automate the morphological identification and scoring of TMEM in immunohistochemically stained formalin-fixed paraffin-embedded tissue sections. We also develop a methodology for the creation of a combined marker by using multimodal image registration to align serial sections, enabling quantification of MenaCalc signals within TMEM rich regions.

**Design:** 60 primary breast cancer cases were stained for Mena expressing tumor cells, macrophages and endothelial cells, scanned on a digital whole slide scanner, and manually scored for TMEM by 5 pathologists. Using single fields of view, an identification/quantification algorithm was developed and its accuracy and reproducibility evaluated. Pathologists’ manual scores were compared to each other and to the algorithm. To establish an algorithm for evaluation of a combined marker, serial sections are stained for TMEM and MenaCalc, scanned, and aligned using whole-slide image registration. **Results:** Pathologist TMEM counts in individual fields of view correlated highly with those of the algorithm. Inter-pathologist and pathologist-algorithm correlations were equivalent. Day to day reproducibility of the algorithm was also high. Finally, we established that structures identified in immunohistochemically stained sections can be used to define zones for quantification of immunofluorescence signals. **Conclusions:** Automated digital pathology based analysis is a robust algorithm that performs as well as a pathologist in scoring TMEM, and reduces pathologist time from 50 to 5 minutes per case. Whole slide image registration provides a suitable platform for combining and comparing the analysis of TMEM and MenaCalc.

**Automatic Prostate Cancer Detection and Contouring On Digital Histopathology Imaging**

Wenchao Han1, Eli Gibson2, Mena Gaed4, Jose Gomez-Lemus3, Madeleine Moussa3, Joseph Chin1,5, Stephen Pautler1,5, Glenn Bauman1,4, Aaron Ward1,4,6,7

Departments of 1Medical Biophysics, 2Pathology and Laboratory Medicine, 3Surgery and 4Oncology, Western University, 5Centre for Medical Image Computing, University College London, Western University, 6Baines Imaging Research Laboratory, London Regional Cancer Program, 7Lawson Health Research Institute, London, Canada. E-mail: whan25@uwo.ca

**Content:** Quantitative pathologic assessments of prostatectomy specimens are particularly time-consuming to measure and report, and are subject to inter-observer variability. This can complicate the decision to undertake adjuvant therapy post-prostatectomy. Our goal is to enable quantitative pathology reporting via development of a machine learning-based software system for automatic detection and contouring of cancerous foci on whole-slide digital histology images. **Technology:** We have developed and measured the accuracy of a software system for automated detection and labeling of cancerous regions on digitized whole-slide images of hematoxylin and eosin-stained tissues from prostatectomy specimens Figure 1. **Design:** We obtained 20 whole-slide mid-gland prostatectomy tissue sections from 17 patients, scanned using an Aperio Scanscope (Leica Biosystems, Wetzlar, Germany) at 0.5 µm/pixel. To compensate for staining variability, we used normalized color deconvolution and our novel adaptive thresholding approach to label nuclei, luminal area, and stroma and other tissue components. We computed tissue component ratios and first- and second-order gray-level co-occurrence and gray-level run length matrix-based texture features from the tissue component maps for 4406 cancerous and 4406 non-cancerous 480 480 sub-images. Supervised machine learning (fisher classifier, logistic linear classifier, nearest mean classifier, uncorrelated quadratic classifier) classified each sub-image. We measured performance using 5-split randomized cross-validation with 20 repetitions. **Results:** Our system demonstrated robustness to staining variability in generating tissue component maps. For cancer vs. non-cancer classification, the logistic classifier gave the best results, with an error rate of 11.1% ± 3%, false negative rate of 15.0% ± 6.4%, false positive rate of 7.26% ± 2.4%, and area under the receiver operating characteristic curve of 0.95 ± 0.02. Processing time using our unoptimized single-threaded Matlab R2015a (Mathworks Inc, Natick, MA) implementation is approximately 24 hours per whole-slide image, with a typical image size of approximately 80,000 x 60,000 pixels. This suggests the potential for useful processing times using a graphics processing unit implementation of our highly-parallelizable algorithm. **Conclusion:** We developed a software system to completely map cancerous and non-cancerous regions throughout high-resolution whole-slide images of hematoxylin and eosin-stained prostatectomy tissues. Once fully validated, this system will enable quantitative pathologic assessments, supporting clinical decision making and retrospective research studies. Ongoing work includes automatic Gleason grading.
Whole Slide Imaging and Microsoft HoloLens: A Comparison of Whole Slide Imaging Viewers

Matthew G. Hanna¹, Liron Pantanowitz¹
University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: hannamg2@upmc.edu

Content: Virtual and augmented reality devices used in healthcare offer a novel means for physicians to interact with their surrounding environment. Newer headsets allow users to blend virtual reality with a real life background. The use of a mixed reality device such as the Microsoft HoloLens has great potential in pathology, especially for telepathology. The aim of this study was to investigate the ability of the HoloLens to navigate whole slide images using different WSI viewers. Technology: Microsoft HoloLens (Development Edition), a wearable holographic computer with sensors and cameras to enable mixed reality human computer interaction. User input included gaze, voice and gestures to interact with holographic content. Microsoft Edge browser was used for viewing WSI with 9 different WSI viewers (AJAX, DBViewer, DigitalScope, Google Maps API, ImageScope, ImageZoomer, Nanozoomer, OpenSlide, Zoomify). Design: Using the HoloLens 9 different WSI viewers were compared employing gestures to perform navigation tasks including scroll (tapping), panning (drag) and zoom (finger directed up for zooming in and down for zooming out). Compatibility and usability of each WSI viewer was documented. Results: Three viewers were compatible (DigitalScope, ImageZoomer, OpenSlide) and permitted digital slides to be viewed and navigated using gestures [Figure 1]. Three viewers (AJAX, Google Maps API, Nanozoomer) permitted viewing, but gestures in all navigation modes did not function as intended. Three WSI viewers (DBViewer, ImageScope, Zoomify) were incompatible with the HoloLens, mostly due to JAVA.

Figure 1: Whole Slide Imaging navigation in a mixed reality environment with the Microsoft Hololens

Conclusion: Whole slide images can successfully be viewed on the HoloLens device. However, not all WSI viewers are yet configured for use in such a mixed reality environment. Future directions should include developing WSI viewers for easy navigation with virtual and/or augmented reality devices.

Whole Slide Image File Integrity: A 10-year Look Back at Archival Images

Matthew Hanna¹, Jon Doby¹ Liron Pantanowitz¹
¹University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: hannamg2@upmc.edu

Content: Whole slide images (WSI) can help overcome storage problems with pathology glass slides such as breakage, faded stains, and loss/misplacement. However, we are not aware of literature that has investigated the integrity of long-term archived WSI. The aim of this study was to compare WSI file integrity over a 10-year period. Technology: Aperio Scanscope T2 (2005-2008), CS (2008-2010), and XT (2010-2016) whole slide scanners were used for scanning. A Microsoft Windows Server (R2 Enterprise, 64 bit, 8GB RAM, Intel Xeon E7-4880 2.4 GHz) was utilized for WSI storage, and included one data migration episode (2012). Adobe Photoshop CC 2017 was employed for image quality measurements. Design: 30 glass slides (24 H&E, 6 immunohistochemistry stains) were retrieved that had been previously scanned on Aperio scanners over a 10-year period (2005-2015). The 30 matching glass slides were re-scanned with an XT scanner for comparison. WSI were categorized as aged (2005-2010 scans), mid-age (2010-2015 scans), and recent (newly re-scanned in 2016). Qualitative and quantitative image quality measurements were performed using Adobe Photoshop to compare image files. All WSI files were also subjected to 100 open/close and compression/decompression cycles. Results: Subjective human evaluation of WSI showed no difference between archived and newly re-scanned slides. Mean compression ratios for aged, mid-age, and recent archived WSI were 23.2, 27, and 13.6, respectively (41.4% decrease compared to aged, 50% decrease compared to mid-age WSI). Average WSI file sizes for aged, mid-age, and recent WSI were 182, 199, and 214 bytes/pixel, respectively (14% difference from aged, 7% difference from mid-age WSI). Histogram color (RGB) and luminance comparative data are shown in Table 1. Aged and mid-age WSI files averaged 7.2% and 3.8% RGB color intensity variation, respectively. Recent images had 10% less noise (via luminance) than aged and mid-age WSI. WSI files showed no degradation after successive open-view-close and compression/decompression cycles. Conclusions: These data show that the integrity of WSI files archived over a decade satisfactorily represent the glass slides from which they were
scanned. Over the years WSI file size has increased, possibly due to improved sensors and/or decreased file compression ratios. Objective analysis showed wider variation in luminance for older WSI, but no notable difference in color. Repetitive opening, viewing, closing and compression-decompression cycles did not appear to impact WSI file integrity. Enhanced whole slide scanner image capture (i.e. sensors) and post-processing (i.e. software) likely play a role in varying WSI file properties.

### Value of Sharing Pathology Educational Digital Slides on Social Media

Matthew G. Hanna, Ishtiaque Ahmed, Fang He, Herman Lo, Liron Pantanowitz

1University of Pittsburgh Medical Center, Pittsburgh, PA, USA. E-mail: hannamg2@upmc.edu

**Content:** The influence of social media is emerging in pathology. Pathology images shared online can be undertaken for educational, consultation and/or marketing interests. Usage statistics and evidence of impact related to sharing digital pathology images via social media are limited in the literature. The aim of this study was to determine the benefit of posting digital images of anatomical and clinical pathology as a “case of the week” on Facebook. **Technology:** A Panoptiq imaging system (Point Grey digital camera with Sony Pregius IMX249 sensor c-mounted to an Olympus BX45 microscope using Panoptiq 3 software v3.9.3) was used to acquire panoramic images with 4x, 10x, 20x, 40x or 100x oil objectives. All images incorporated embedded z-stack video frames for multiplane focusing of key diagnostic regions. **Design:** Glass slides of interesting surgical pathology and hematology cases were used to generate panoramic digital slides with z-stacks (e.g. https://goo.gl/68PXRR). Image files were uploaded to the Panoptiq ViewsIQ portal. ViewsIQ used Amazon’s cloud storage to host files and designed an HTML5 viewer to be cross compatible with different mobile devices. Digital image navigation was tested on an iPhone 6 (iOS 8), iPad air 2 (iOS 9), and Microsoft Surface Pro 3 (Windows 10). Clinical history with quiz questions was compiled for each case using Google Forms. Weekly cases were publicized on Facebook and Twitter (@DigPathologists). Web service analytical tools (Facebook Insights, Twitter Analytics, Google Analytics) were used to capture traffic data. **Results:** After 20 weekly posted digital slides, total followers on Facebook increased by 135 and on Twitter by 186. These posts triggered 18,410 social media interactions (average 921/week) from 32 different countries, including likes and/or shares on Facebook (total, 5684; weekly average, 284), and impressions or engagements (e.g. retweets, replies, follows, likes) on Twitter (total, 12726; weekly average, 636). Social media engagement on Twitter was more actively shared by users compared to Facebook, 51% vs 14%, respectively. 64% of users correctly answered quiz questions. **Conclusions:** Prior to posting pathology images on social media websites it is important to employ applications that users are not slowed down by having to download and that are guaranteed to run on any mobile device. Amazon’s cloud infrastructure provided high uptime, ample speed, and enforced encryption of all communications and storage. Sharing interesting pathology cases on social media that are easy to view is an effective educational tool that can reach a large cohort of users.

### Pathpresenter.Com: An Innovative Platform for Teaching, Learning and Sharing of Pathology Images

Rajendra Singh, Matthew Hanna, Brandon Veremis, Anil Parwani

1University of Pittsburgh Medical Center, Pittsburgh, PA. 2Pathology, Icahn School of Medicine at Mt. Sinai, OSUMC, Columbus, Ohio, USA. E-mail: skinpathology@gmail.com

**Content:** The microscope has been around for more than a 100 years and pathologists love their microscopes. Pathology is image-rich and very visual and pathologists use many types of images and yet the workflow is very manual. Digital pathology has been around for more than a decade and despite the countless advantages it brings, digital pathology has only seen limited adoption. **Technology:** PathPresenter.com is a completely web-based platform, without the necessity for plugins or other downloads. The images are stored on a cloud-based platform with easy accessibility and a robust user interface. **Design:** PathPresenter.com is an innovative platform that bridges the gap of exposing pathologists to digital pathology for daily use. The platform has three major areas of use: presentations, high yield cases, and an extensive searchable slide library. The platform provides a streamlined workflow for teaching and learning pathology. It converts conventional PowerPoint presentations with static images to live, interactive presentations with whole slide images as digital slides. Annotation and presentation tools are available in presentation mode. **Results:** Since launching in January 2017, the platform has had more than 20,000 page visits,
Implementation of Whole Slide Imaging as a Pathology Teaching Tool and for Institutional Tumor Boards: A Resident’s Experience

Ashish Mishra1, J. Mark Tuthill1

1Department of Pathology and Laboratory Medicine, Henry Ford Health System, Detroit, MI, USA. E-mail: amishra2@hfhs.org

Content: This presentation will describe our experience implementing and utilization whole slide imaging (WSI) as a teaching tool for the pathology residents in Henry Ford Hospital, Detroit as well as our initial efforts to use WSI at institutional tumor boards. Technology: 1. Health system network. 2. Roche Ventana iScan HT Whole Slide Scanner (Roche Diagnostics Corporation, Indianapolis, IN) capable of scanning up to 360 slides at one time. 3. Internet Explorer 11 access to Roche Ventana Virtuoso WSI viewing software, 4. Basic Windows workstations as deployed throughout HFHS, 5. CoPathPlus (Sunquest Information Systems). Design: Glass slides were scanned for practice over several weeks to determine basic operation, system performance and workflow processes. Experience quickly showed that the iScan HT could be used to improve quality and efficiency of weekly unknown slide conference. A proposal was made and accepted to pilot this process. Username’s and passwords were given to all the residents allowing them to access the digital slides from any terminal in the hospital campus. Initially there was a lot of reluctance and resistance from the group to use WSI. To increase interest and enthusiasm concordance studies were presented in the monthly journal club which compared WSI and glass slides. To further interest, slides from a recently autopsy were scanned and presented during the journal. In addition, informatics lecture and luncheon meeting topics as well as a grand rounds, presentation on novel ways to use WSI were shared with the residents and other members of the department. This resulted in marked increased interest. Soon interest grew from attending physicians to use WSI for a subset of tumor boards. The same processes and procedures used for scanning slides for unknown conference were applied. Results: In Oct 2016 unknown slide conference was presented using WSI. The reaction to the quality of the histopathology system usage was excellent: nuclear contours and nucleoli were clear; navigation easy; response time was excellent with no screen lag. The conference was well received. The residents and attending loved the new format. Since then unknown conference has been presented monthly using WSI. In November 2016, we started presenting cases on WSI in GYN tumor board. GYN tumor board is unique as we typically need to present entire slides, not just the static picture of a relevant area. To do so we would project the slides via a microscope connected to the TV in the tumor board conference room a stressful and inefficient task. WSI, eliminated many problems went away in an instant rapid navigation to area(s) of interest, ease of switching between slides and ease of switching cases, were ecstatic after the tumor board. The resident was super confident and was able to explain everything in detail using WSI without fumbling. The Attending, clinicians and residents were enthused at the new format; Some had no idea that this was even technically possible. All GYN weekly tumor boards are presented using WSI. Conclusion: Whole slide imaging is a useful tool for teaching and presentation purposes. It can be easily implemented and integrated into our day to day pathology practice and resident training. The reluctance to use WSI is initially high among pathologists, but enthusiasm increases once implemented into regular practice. WSI provides for efficiencies and ease of collaboration in both educational and clinical case review settings such as institutional tumor boards.

Machine Learning Strategies to Optimize in Silico Genetic Variant Effect Predictions Informed by TP53 Functional and Clinical Data

Joshua F. Coleman1, Dan Jones2

1Medical Affairs, GenomOncology, LLC, Cleveland, 2Department of Pathology, The Ohio State University, Columbus, OH, USA. E-mail: josh@genomoncology.com

Content: Predicting the impact of genetic variants of uncertain significance (VUS) is a major challenge for clinical genomics. Numerous in silico variant effect prediction (VEP) algorithms purport to address this, but better validation is needed, especially to improve specificity and prevent potentially
harmful false-positive results. Using a public database of TP53 variants curated with clinical and functional data, we compared several common VEP algorithms alone and in combination using support vector machine (SVM) and deep neural network (DNN) techniques. Technology: Python language (3.5.2) with TensorFlow (0.11.0, Google, Mountain View, CA) library; R language (3.3.2); ANNOVAR (2/1/2016). Design: Tumor and germline databases of TP53 mutations were downloaded in CSV format from the UMD TP53 website (http://p53.fr, accessed 10/5/2016). Missense variants with preserved function in yeast studies selected from the tumor database were compared against inactivating germline variants associated with heritable tumor syndromes. Variant function was annotated with precompiled in silico predictions from the dbNSFP database, version 3.0, via ANNOVAR. Receiver-operator characteristics (ROC) of 21 VEP algorithms were analyzed in R. Based on the five best performing classifiers, an SVM was trained on a random 70% sampling of the variants in R, then tested on the remaining 30%. Using the same testing and training sets, a DNN with 3 layers of 30 nodes each was composed using TensorFlow. Results: The training set contained 230 inactivating variants and 352 with preserved function; the test set contained 101 and 148, respectively (831 total). Based on AUC, the five best performing individual algorithms were VEST3, FATHMM, MutationAssessor, PROVEAN and Polyphen-2 HVAR. Optimization based on ROC analyses generally improved specificity at a modest cost to sensitivity, relative to default interpretations available for 10 VEP. The accuracy of the SVM classifier was 0.933 in a 5-fold cross-validation of the training data and 0.936 with test data (AUC of 0.921). The DNN yielded an accuracy of 0.940 and an AUC of 0.976, the latter essentially similar to the AUC of VEST3. see Table 1 for sensitivity/specificity calculations. Conclusion: Optimization of VEP algorithms for individual genes is possible given sufficient functional training data. Ensemble VEP classifiers do not necessarily outperform the best individual component.

### Dynamic Models for Precision and Personalized Pathology

**Iman Tavassoly**

1 Institute for Systems Biology, Icahn School of Medicine at Mount Sinai, New York, USA. E-mail: iman.tavassoly@mssm.edu

**Content:** Dynamic models of pathological phenomena in diseases such as cancer are very useful tools to understand the pathogenesis of disease progression and development. These models study physiological and pathological phenomenon at cell and tissue level focusing not only on components within cells and tissues but also on the evolution of diseases in time and space. In this work, the implications of these models for precision and personalized pathology are discuss

| Table 1: Sensitivity/specificity calculations |
|---------------------------------------------|
| **Existing classifiers via dbNSFP** | **Individual versus ensemble classifier** | **AUC** | **Sensitivity (initial → optimized)** | **Specificity (initial → optimized)** |
|---------------------------------------------|---------------------------------------------|---------|----------------------------------------|----------------------------------------|
| VEST3                                       | Individual                                 | 0.977   | N/A → 0.946                            | N/A → 0.93                             |
| FATHMM                                      | Individual                                 | 0.934   | 1 → 0.882                              | 0 → 0.912                              |
| MutationAssessor                            | Individual                                 | 0.928   | 0.97 → 0.921                           | 0.66 → 0.786                           |
| PROVEAN                                     | Individual                                 | 0.926   | 0.95 → 0.9                              | 0.76 → 0.846                           |
| Polyphen2 HVAR                              | Individual                                 | 0.912   | 0.98 → 0.912                           | 0.62 → 0.838                           |
| Polyphen2 HDIV                              | Individual                                 | 0.902   | 1 → 0.894                              | 0.53 → 0.858                           |
| SIFT                                        | Individual                                 | 0.889   | 0.99 → 0.955                           | 0.63 → 0.772                           |
| FATHMM-MKL                                  | Individual                                 | 0.83    | 0.97 → 0.934                           | 0.43 → 0.688                           |
| SiPhy 29-way log odds                      | Individual                                 | 0.809   | N/A → 0.779                            | N/A → 0.714                            |
| LRT                                         | Individual                                 | 0.789   | 0.93 → 0.949                           | 0.71 → 0.624                           |
| GERP                                        | Individual                                 | 0.765   | N/A → 0.834                            | N/A → 0.61                             |
| phastCons 7-way vertebrate                 | Individual                                 | 0.745   | N/A → 0.752                            | N/A → 0.716                            |
| phastCons 20-way mammalian                  | Individual                                 | 0.647   | N/A → 0.855                            | N/A → 0.538                            |
| phylp 7-way vertebrate                     | Individual                                 | 0.637   | N/A → 0.707                            | N/A → 0.612                            |
| MutationTaster                              | Individual                                 | 0.602   | 0.96 → 0.97                            | 0.47 → 0.234                           |
| phyloP 20-way mammalian                     | Individual                                 | 0.591   | N/A → 0.909                            | N/A → 0.45                             |
| fitCons                                     | Individual                                 | 0.491   | N/A → 0.97                              | N/A → 0.054                            |
| MetaLR                                      | Ensemble                                   | 0.939   | 1 → 0.894                              | 0.03 → 0.92                            |
| CADD (raw)                                  | Ensemble                                   | 0.889   | N/A → 0.84                             | N/A → 0.798                            |
| DANN                                        | Ensemble                                   | 0.868   | N/A → 0.722                            | N/A → 0.858                            |
| MetaSVM                                     | Ensemble                                   | 0.65    | 1 → 0.792                              | 0.03 → 0.606                            |
| **Novel classifiers**                       | **Individual versus ensemble classifier** | **AUC** | **Sensitivity** | **Specificity** |
|---------------------------------------------|---------------------------------------------|---------|----------------|----------------|
| Novel 5-way SVM classifier                  | Ensemble                                   | 0.921   | 0.921          | 0.946          |
| Novel 5-way DNN classifier                  | Ensemble                                   | 0.977   | 0.95           | 0.932          |

AUC: Area under the curve, SVM: Support vector machine, DNN: Deep neural network, N/A: Not available
using dynamic models of cell death pathways in cancer. **Technology:** Dynamical and mathematical models were used to establish a general pipeline for precision and personalized pathology. Computational methods to detect and determine parameter spaces and finding sensitive parameters were applied. **Design:** Signaling pathways controlling cell death modalities (mainly apoptosis, necrosis and autophagy) in cancer were translated into a set of ordinary differential equations and the parameter spaces that give robustness to these models were found. These parameters were assigned to biomarkers available for quantitative measurements in patients. The predictive ability of these models was verified within current pathology literature. **Results:** Analysis of parameter spaces of cell death pathways in cancer showed some sensitive parameters which can clinically be translated to some biomarkers. These biomarkers are the main determinants of disease progression and variations in their levels can be used to make personalized diagnosis and find the prognosis for each individual patient more accurately. **Conclusions:** In the era of big data in medicine, dynamic models can be used for personalized and precision diagnosis of cancers and other diseases. Introducing them to the field of pathology requires development of basic mathematical and computational pipelines. Current example of cell death pathways and their dynamical models is a general pipeline for establishing the field of precision and personalized pathology.

**Discovery of Prognostic Factors for Gastric Cancer Based on KI67(+) Immune and Epithelial Cells Using Deep Learning**

Nathalie Harder1, Katharina Nekolla1, Nicolas Brieu1, Armin Meier1, Carolina Vanegas1, Ralf Schönmeyer1, Mehmet Yigitsoy1, Victor Matvienko1, Aleksandra Zuraw2, Ralf Huss1, Günter Schmidt1

1Research, Definiens AG, Munich, Germany. E-mail: nharder@definiens.com

**Content:** Gastric cancer is among the leading causes of cancer-related deaths worldwide. To improve treatment success, particularly for upcoming immunotherapies, it is important to characterize spatial properties of tumor-infiltrating immune cells and co-localized cancer cells. We present an automatic approach to identify prognostic factors related to KI67(+) cells based on clinical trial data comprising 248 gastric cancer patients treated at Kanagawa Cancer Centre Hospital, Japan. **Technology:** We used a convolutional neural network (CNN) to predict heatmaps of KI67(+) immune and epithelial cell densities on tissue microarray (TMA) data. Based on these heatmaps we computed features characterizing each patient, and automatically identified prognostic factors by correlating such features with clinical outcome in a Tissue Phenomics workflow. **Design:** Based on manual annotations of positive cells we generated training patches (128x128 pixels) with six patch classes: all negative (0), KI67(+) epithelial cells medium (1) and high (2) density, KI67(+) immune cells medium (3) and high (4) density, and both KI67(+) epithelial and immune cells mixed (5) [Figure 1b]. A reduced GoogLe net was trained with 100k iterations using 30k augmented patches from a patient subset reserved for training by transfer learning. By applying the network for patch-based prediction we generated KI67(+) cell density heatmaps for all TMA cores [Figure 1a]. Ratios and percentages of all classes in tumor and normal tissue cores were computed (n=43) and systematically mined as to their prognostic value. **Results:** Quantitative evaluation of the predicted heatmaps yielded 75.5% accuracy on the patient test set. Data mining identified features with high prognostic value using 100x10-fold Monte Carlo cross-validation. The percentages of class 1-patches (KI67(+) epithelial cells) in tumor [see Kaplan Meier plot Figure 1c] and class 4-patches (KI67(+) immune cells) in tissue turned out to be stable and strongly positive prognostic factors for cancer-specific death (p-value<0.05). **Conclusions:** KI67(+) epithelial and immune cell densities provide prognostic value for gastric cancer patients, which we found using a CNN with the Tissue Phenomics methodology without explicit cell segmentation. We will validate the discovered factors using data from another clinical site.

**Acknowledgment**

We thank Prof. H. Grabsch, Maastricht University, NL, for providing the data, and S. Kisiak, W. Susanto, and R. Menendez for their contribution to cell annotations.

**Generalizing Deep Learning Models for Histology Data**

Shazia Akbar1, Anne Martel1, Mohammad Peikari1, Sharon Nofech-Mozes2, Sherine Salama2
Content: Cancer cellularity is an important component of residual cancer burden in breast specimens from patients treated with neoadjuvant systemic therapy. This task is traditionally performed manually (eyeball estimation), which is not only time-consuming but also introduces inter-and intra-rater variability. Considerable progress has been made in the field of image analysis in digital pathology, increasing throughput and improving standardization. Here we propose a method for automatically determining cancer cellularity by analyzing digital slides using deep learning techniques. We develop a method for capturing complex textural appearances of cell nuclei and surrounding tissue to automatically classify image patches into one of four groups: normal tissue and, low, medium and high cellularity. Specifically, we aim to generalize our model to give high performance on a held-out test set. Technology: Deep learning has had significant impact in digital pathology resulting in state-of-the-art performance on a wide range of tasks. Specifically, convolutional neural networks, or CNNs, encompass a series of convolution and downsampling operations which enable complex textures to be captured at multiple scales without explicit modelling of tissue or cellular structures. To prevent overfitting, dropout layers can be added to reduce the size of the network thus making it more generalizable. This property is ideally suited for histology data which contains highly variable structures particularly in datasets containing cancerous tissue. Design: We designed a CNN architecture to automatically classify 2,579 image patches extracted from breast tissue whole slide images stained with H&E. The goal of the CNN is to assign a single label to image patches which represent varying degrees of cellularity. Our CNN contains three convolutional and max-pooling layers as well as two fully connected layers, the last of which outputs class predictions. We also investigated the use of dropout to improve generalizability. Our experimental setup consisted of separate train, validation and test sets; results are reported over five repeated experiments. Results: The addition of dropout in our CNN increased test accuracy performance from 51.89 ± 0.95% to 62.83 ± 1.84% (dropout rate=0.7), showing considerable performance gain by reducing network size. Confusion matrices showing prediction accuracies on the test set are shown in Figure 1. With the introduction of dropout, correct predictions increased significantly for the normal (0), low cellularity (1) and high cellularity (3) classes. Conclusion: Our results demonstrate the importance of introducing generalizability when training CNNs with histology data, due to high variability between samples, cases and cohorts. We demonstrate impressive accuracy rates on a cell classification problem which includes normal healthy tissue, which has in the past shown to be challenging to model using hand-crafted features. Our results show CNNs can be used effectively in digital pathology with some minor adjustments.

An Open Source Web Application for Real-time Display of Pending Orders

Noah G. Hoffman1

1Laboratory Medicine, University of Washington, Seattle, WA, USA.
E-mail: ngh2@uw.edu

Content: Many laboratory information systems (LIS) do not provide real-time notification of new orders, relying instead on batched, asynchronous display of information such as printed pending lists. To improve situational awareness of pending laboratory orders, we developed a web application (the “Pending Log Monitor”) that displays data continually updated from our LIS on large wall-mounted monitors or PC workstations. Users may enter comments associated with individual items. A survey was administered to evaluate usage patterns. Technology: The application is implemented in Python 2.7 using the Flask web microframework, and is hosted on a virtual machine running Ubuntu 14.04. Data is extracted from the LIS database (Sunquest Information Systems, Tucson, AZ) using custom code written in Cache (InterSystems Corporation, Cambridge, MA), and is transferred to the application server by a batched process using secure shell. User-provided comments associated with pending tests are stored in an SQLite database. Design: The application was designed for maintainability, ease of customization, stability, and rapid recovery in the result of a component failure. Logic for display and formatting of pending tests is implemented as Python functions. A simple JSON-format specification can accommodate any tabular data. Lists of pending tests defined for a given area typically correspond to one or more worksheets defined in the LIS. Results: Customized displays of pending tests have been implemented for over 35 combinations of worksheets in multiple lab areas. Pending orders for each lab area are filtered, ordered, and color coded based on elapsed time since order or receipt, priority, specimen stability, or other criteria. Data is transferred from the LIS by a batched process every four minutes. This application has replaced the use of
We studied the intraoperative consultation. Timely and properly performed intraoperative consultation reduces error and increases efficiency. A properly designed system (computer workstation or mobile app) will improve intraoperative pathology consultation significantly.

Potential Benefits and Barriers for Implementing a Digital Communication System for Intraoperative Pathology Consultation

Buer Song¹, Nan Zhang¹, Jane Zhao¹, Edwin Anand¹, Peter L. Elkin¹, John E. Tomaszewski²

¹Biomedical Informatics, University at Buffalo, The State University of New York, New York, USA. E-mail: buersong@buffalo.edu

Content: Timely and properly performed intraoperative pathology consultation often provides critical information to ensure the success of a surgery. This is a highly stressful process for both surgical and pathological teams. Many team members with different understandings of the work flow are involved in the process. Multiple steps need to be accomplished. A complicated process like this is prone to encounter errors and delays. Common errors and delays can often be traced back to particular communication errors. These include (1) mis-labeled specimens, (2) illegible or incomprehensible requisition sheets, (3) confusing specimen orientation, and (4) suboptimal telecommunication hardware connecting the consultation lab and the operating room. In this simulation study, we attempt to systemically analyze the current work flow of intraoperative pathology consultation in a tertiary medical center, list the problems and barriers for achieving optimum efficiency and minimizing error rate, and propose a computer workstation-based or mobile device app based digital communication solution to replace the current analog environment, and compare the two systems.

Technology: We studied the intraoperative consultation process at Buffalo General Medical Center (BGMC), Buffalo NY. for the analog work flow analysis. A simulated digital communication system will be proposed and applied to the work environment to replace the current work flow. Design: The digital system attempts to completely or partially automate the steps of specimen labeling, requisition sheet fill-out, specimen tracking, specimen accessioning. Advanced communication between operating room and pathology lab with real-time video on gross and microscopic findings is also included. A detailed comparison for potential benefits and drawbacks will be performed.

Results: Multiple steps of current intraoperative consult process at BGMC are prone to errors. In the simulated work flow with an integrated digital communication system, errors in multiple analog work flow steps can be successfully eliminated or minimized. Compared to analog process, the major potential drawbacks of the digital system include system malfunctioning/break down, and workforce compliance with the new workflow. Conclusion: A digital communication system for intraoperative pathology consultation reduces error and increases efficiency. A properly designed system (computer workstation or mobile app) will improve intraoperative pathology consultation significantly.

Baikal: Streaming Data Science Platform for Laboratory Business Intelligence

Thomas Durant¹, Nathan Price², Peter Gershkovich¹, Wade Schulz¹

¹Departments of Laboratory Medicine and Pathology, Yale School of Medicine, ²Information Technology Systems, Yale-New Haven Health System, New Haven, CT 06520, USA. E-mail: thomas.durant@yale.edu

Content: In the era of healthcare system expansion and centralization of laboratory resources, monitoring quality improvement (QI) metrics becomes increasingly critical and challenging. Emerging data management technologies offer novel approaches to enhance QI practices. To this end, we present a component of a data science platform developed at our institution – Baikal – that allows for stream processing and analysis of laboratory orders and results.

Technology: Primary components include the Hortonworks Data Platform (version 2.4.2; Hortonworks, Santa Clara, CA, USA) and Hortonworks Data Flow (HDF) version 1.2 (Hortonworks, Santa Clara, CA, USA). Custom Python (version 2.7) scripts are executed within NiFi to calculate laboratory QI metrics on streaming laboratory data. QI metrics currently include turn-around time (TAT) for lab results, outstanding orders, and order volumes. Data from Health Level 7 (HL7) ORU messages and QI metrics are stored in Elasticsearch (version 2.4.2; Elastic, Mountain View, CA, USA) and can be readily visualized with Kibana (version 4.3.1; Elastic, Mountain View, CA, USA). For laboratory and EHR-related data, Cloverleaf (Infor, NY, USA) was used as the interface and integration engine.

Design: A custom emissary service was deployed to receive a stream of HL7 ORU messages from Cloverleaf. HL7 messages were validated and transformed into JSON documents for structured storage and stream processing. Custom Python (version 2.7) scripts were implemented in NiFi to denormalize messages and calculate QI metrics. Raw messages were stored in Hadoop Distributed File System and processed messages with QI metrics are routed to Elasticsearch for visualization with custom dashboards in Kibana.

Results: Baikal was deployed in July of 2016. In a representative two-month period between August 1st, 2016 and November 1st, 2016; 1.3 million tests were ordered with...
In response to a clinical need, we have developed a mobile app for calculating the minimum volume of blood required within each tube. Instrument dead volume, hematocrit correction, and other relevant parameters are accounted for. The app incorporates 168 tests and panels, encompassing 13 central laboratory instrument systems, 10 collection tube types and 5 specimen types. Results: Our house staff and nurses will integrate the app into their clinical workflow to reduce the number of tubes and volume of blood drawn. Usage, user feedback, specimen collection patterns, and patient outcomes will be followed for a longitudinal period and reported. Conclusion: In response to a clinical need, we have developed a mobile app for calculating the minimum volume of blood required for any given set of tests. The app will integrate into our hospital workflow with the primary goal to reduce iatrogenic anemia especially in the pediatric setting.

A Mobile App to Calculate Minimum Blood Volume Needed for Testing

Ray Zhang¹, Sarah M. Brown²

¹Department of Pathology and Immunology, Division of Laboratory and Genomic Medicine, Washington University School of Medicine, ²Department of Pathology, St. Louis University School of Medicine, St. Louis, MO, USA. E-mail: r.zhang32@wustl.edu

Content: Iatrogenic anemia, often caused by excessive blood drawn for laboratory testing, is a major risk factor for patient morbidity especially in the pediatric population. While reducing the volume of blood drawn for testing to a minimum has been shown to decrease incidence of iatrogenic anemia, determining the minimum volume needed for various combinations of laboratory tests is often difficult and requires laboratory assistance. Technology: We have developed a mobile app (iOS, Android) to facilitate such calculations at the patient bedside. Design: Users input tests, selected from the institutional testing menu, while the application outputs which types of, and how many, blood collection tubes are needed as well as the minimum blood volume required within each tube. Instrument dead volume, hematocrit correction, and other relevant parameters are accounted for. The app incorporates 168 tests and panels, encompassing 13 central laboratory instrument systems, 10 collection tube types and 5 specimen types. Results: Our house staff and nurses will integrate the app into their clinical workflow to reduce the number of tubes and volume of blood drawn. Usage, user feedback, specimen collection patterns, and patient outcomes will be followed for a longitudinal period and reported. Conclusion: In response to a clinical need, we have developed a mobile app for calculating the minimum volume of blood required for any given set of tests. The app will integrate into our hospital workflow with the primary goal to reduce iatrogenic anemia especially in the pediatric setting.

Image Standardization and Its Impact on Whole Slide Image Analysis

Mark Zarella¹, David E. Breen¹, Fernando U. Garcia¹

¹Pathology and Laboratory Medicine, Drexel University College of Medicine, Philadelphia, PA, USA. E-mail: mark.zarella@drexelmed.edu

Content: Computational analysis of H&E images relies on the staining and digitization process to produce an image that conforms to an expected range of color. Attempts to improve image analysis often begin with a color normalization step to force the data set to conform to this expectation. However, the impact of standardization on image analysis and whole slide viewing is largely unknown. Technology: We used image processing to develop and apply a novel color normalization algorithm to whole-slide images. We used eye tracking and psychovisual analysis to measure pathologist gaze patterns to evaluate the impact of color normalization on histology analysis. Design: We analyzed the color properties of publicly available H&E image sets as well as those acquired from the Drexel University cancer databank to better understand the color attributes commonly encountered in digital pathology. We used these quantities to guide the development of an algorithm that normalizes images by anchoring histologic structures to a set of target colors with the aid of a previously reported structure classification algorithm. We measured pathologist gaze patterns and diagnostic performance while they viewed unnormalized, normalized, and artificially manipulated images to characterize the impact of normalization on diagnosis. Results: We found an approximate 10-fold decrease in inter-image variability after applying color normalization. Importantly, loss associated with this transformation, as measured by the normalized mutual information, was very small, indicating its potential utility as a preprocessing step for advanced image analysis and whole-slide viewing. However, despite improvements in computational image analysis performance, we observed only modest differences in whole-slide viewing due to normalization. Conclusion: As regulatory policy continues to take shape in the arena of whole-slide imaging, a quantitative treatment of color becomes an important factor to help refine these policies. We present a novel algorithm for color normalization that substantially reduces inter-image variability while retaining the inherent information in the image, leading to greater standardization in pathology and its measurably positive impact on image analysis. More generally, we describe a color representation paradigm that enables a pathology-specific quantitative framework for color, promoting a platform by which we can compare diagnostic performance and image attributes.

An Update on the Digital Imaging and Communication in Medicine Digital Pathology Connect-a-thon

Dan Hosseinzadeh¹, Mikael Wintell², Bruce A. Beckwith³

¹Pathcore Inc., Toronto, Canada, ²Västra Götalandsregionen, Hälsos och Sjukvårdsavdelningen, Jönköping, Sweden, ³North Shore Medical Center, Miami, FL, USA. E-mail: dan.zadeh@pathcor.com
**Content:** The digital imaging and communication in medicine (DICOM) standard is the most widely used medical imaging standard around the world and has been used to archive more than 10 billion medical images. DICOM provides data standards and communication protocols for imaging that currently support more than 60 modalities including whole slide imaging and pathology metadata. Many digital pathology vendors and pathology associations, such as the College of American Pathologists, have contributed to the standard since 2005 through the DICOM Pathology Working Group (WG-26). WG-26 is currently working on a number of initiatives to expand the capabilities of DICOM for anatomic pathology including support for multispectral imaging, standards for communicating structured reports, standardized workflows for digital pathology and a multi-vendor connect-a-thon to demonstrate the benefits of a DICOM-based workflow for digital pathology. This talk focuses on the motivations for the connect-a-thon. **Technology:** The DICOM digital pathology connect-a-thon will demonstrate benefits of interoperability using DICOM to the pathology community. DICOM is a natural choice for achieving interoperability in digital pathology since it provides familiar standards for handling, storing, printing, and transmitting medical images. **Design:** The connect-a-thon will demonstrate a vendor neutral pipeline that allows imaging, archival and review [Figure 1]. All digital pathology vendors will be invited to participate with the overarching goal to show a wide range of scanners and viewing software. **Results:** The DICOM digital pathology connect-a-thon is currently in the early planning stages. The goal of the WG-26 committee is to organize the connect-a-thon for the fall of 2017. **Conclusions:** While the use of DICOM as a data and communication standard in digital pathology is still nascent, there is growing demand for DICOM in the pathology community. This demand is driven by the need to reduce disparate data formats and to future proof investments in digital pathology infrastructure. DICOM includes a file format definition and communication protocol that allow devices and software to interoperate within a networked environment. Thus it becomes possible to create a vendor neutral workflow that comprises image scanners, image archives and workstations.

**Barcode Beware: Using ISBT 128 in Positive Patient Identification for Blood Administration**

Kinjal Sunil Shah1, Cassandra D. Josephson1,2, Alexis B. Carter2

1Department of Pathology and Laboratory Medicine, Center for Transfusion and Cellular Therapies, Emory University, 2Department of Pathology and Laboratory Medicine, Children’s Healthcare of Atlanta, Atlanta, Georgia, USA. E-mail: kinjal.sunil.shah@emory.edu

**Content:** Positive patient identification (PPID) blood administration systems improve safety by ensuring correct pairing of patient to blood product. The selection of a blood administration PPID system was performed at Children’s Healthcare of Atlanta (CHOA). Regulatory, accreditation and safety issues were found during system selection which merit dissemination. **Technology:** A transfusion safety project considered two systems for PPID for blood administration in 2016. The Electronic Health Record (EHR) vendor differs from the blood bank (BB) information system (BBIS). The application names are hidden because the discovered issue is system-agnostic. **Design:** Two systems were compared for cost, integration and safety. Compliance with the following regulatory and accrediting agencies: FDA, AABB and The Joint Commission. ISBT 128 is the only barcoding and labeling system approved by the FDA. **Results:**

|                | System A | System B |
|----------------|----------|----------|
| Cost           | +        | +++      |
| Integration with EHR | +++      | ++       |
| Integration with BBIS | No integration | No       |
| FDA 510(k) cleared | Yes      | Yes      |
| Rapid infusion module | No       | No       |
| Degree of difficulty | Nursing workflow | +        | ++       |
|                | BB workflow | +++      | +        |
|                | EHR implementation | +        | +++      |
|                | BBIS implementation | +++      | +        |
| Blood product unit number and product code matching: | Dependent on interfaced data from BBIS | Yes | No |
| Acquired from ISBT 128 compliant barcode reads | Depends on setup | Depends on setup | Yes |

EHR: Electronic health record, BB: Blood bank, BBIS: BB information system, FDA: Food and Drug Administration, ISBT: International society for blood transfusion, + = modest, ++ = moderate, +++= significant

System A met all criteria for selection except that it was dependent on interfaced blood product ISBT 128 codes.

![Figure 1: Digital imaging and communication in medicine digital pathology connect-a-thon configuration](image-url)
which the BBIS could not provide discretely, introducing the potential for human error. Scanning ISBT 128 barcode data into free text fields is not an option because section 2.2 of the ISBT 128 technical specification requires software to ensure data integrity prior to acceptance. This ISBT 128 compliance issue was not known to any stakeholders prior to system comparison. A search of the medical literature did not reveal this as a potential compliance issue with these systems. **Conclusions:** System B was chosen because of its enhanced patient safety due to compliance with ISBT 128. FDA clearance and barcoding may provide a false sense of security with PPID blood administration systems. Healthcare institutions should have full knowledge of ISBT 128 requirements prior to implementing them.

**Feasibility of Converting and Viewing Whole Slide Images in Digital Imaging and Communications in Medicine Format**

Matthew G. Hanna¹, Brian Kolowitz¹, Liron Pantanowitz²

¹University of Pittsburgh Medical Center, Pittsburgh, PA, USA.  
E-mail: hannamg2@upmc.edu

**Content:** The current lack of standards with whole slide imaging (WSI) file formats has limited interoperability. In particular, the use of proprietary WSI formats has hindered incorporating WSI images into enterprise-wide imaging systems (e.g. picture archiving and communication systems) that deal with Digital Imaging and Communications in Medicine (DICOM) images. DICOM is a known standard for medical imaging, but has yet to be fully adopted in pathology.

The aim of this study was to determine the feasibility of converting proprietary WSI files and viewing them in DICOM format. **Technology:** Glass slides were scanned on an Aperio scanner at 20x magnification (0.5 micron/pixel). Software used for DICOM conversion included a command line tool plug-in DICOMizer (Orthanc) and FFEI Sierra OpenSlide to DICOM Converter. DICOM conversion was performed on a Hewlett-Packard Z420 PC (3.2 GHz, 64-bit, 16GB RAM). DICOM files were imported to www.orthanc-server.com and viewed using their WSI webviewer powered by OpenLayers toolkit. **Design:** WSI of glass slides scanned on an Aperio scanner and publically available WSI files (NDPI, SCN, MRXS, TIFF, BIF) from OpenSlide were converted to DICOM using Orthanc and FFEI Sierra converter software. The process of DICOM conversion and viewing was compared to using radiology DICOM images. **Results:** All WSI file types successfully converted to DICOM. Conversion times averaged 0.3 seconds/MB with Orthanc DICOMizer and 2 seconds/MB with FFEI Sierra converter. DICOM file sizes for converted images were 52% smaller than original native WSI files. DICOM converted WSI files were supported by the WSI webviewer, which allowed pan and zoom functionality. However, the default color profile (e.g. YBR) of only Orthanc converted files differed from the original WSI [Figure 1]. **Conclusions:** The conversion of WSI with different file formats into DICOM format is possible and can be performed using freely available open-source software (e.g. Orthanc). Viewing DICOM converted WSI is achievable, but depending on the photometric interpretation employed (i.e. YBR_422) may result in altered color representation. RGB color space profile converted files retained source WSI color fidelity. Converted DICOM file sizes were at least half the size of original proprietary WSI files, which thus necessitate less storage space.

**Towards Computable Cancer Synoptic Reports**

W. Scott Campbell¹, Geoff Talmon¹, Audrey Lazenby¹, Alexis Carter², Rajesh Dash³, James R. Campbell¹

¹University of Nebraska Medical Center, Pittsburgh, PA, ²Children’s Hospital of Atlanta, Atlanta, GA, ³Duke University School of Medicine, Durham, North Carolina, USA. E-mail: wcampbel@unmc.edu

**Content:** The synoptic report contains the summative pathology assessment of malignant tissue to communicate diagnostic and prognostic information to clinicians. To realize the full potential of synoptic data for use in patient care, decision support, analytics and population health, the synoptic report elements must be computable, that is the data elements must be machine readable and support computational analysis. To date, this objective has not been achieved. A collaborative terminology development effort to expand SNOMED CT content in support of computable cancer synoptic reports is presented. **Technology:** In conjunction with the LOINC-
SNOMED International collaborative agreement, the SNOMED CT concept model was expanded to support robust definitions for Observable entity concepts and was employed to encode the College of American Pathologists (CAP) cancer protocol worksheets. **Design:** Investigators at University of Nebraska Medical Center in conjunction with pathologists from the CAP and trained SNOMED CT terminologists analyzed the CAP protocols for colorectal and breast cancers. SNOMED CT definitional content was authored using the expanded SNOMED CT concept model and incorporated into the pathology information system at the Nebraska Medical Center. HL7 messages were used to transmit encoded synoptic data to an institutional tissue biobank. Biobank queries were performed to support clinical and research use cases. **Results:** A total of 194 concepts were developed for colorectal cancer and invasive breast cancer CAP worksheets including biomarker worksheets. Between October 2016 and March 2017, 81 breast and 45 colorectal tumors were encoded as part of routine rendering of anatomic pathology diagnoses and successfully transmitted to the biobank registry. Queries of the biobank database supported the use cases submitted by clinicians and researchers and demonstrated the ability to use synoptic data independently of the context of the original synoptic worksheet. Data query use cases will be presented, as well as, SNOMED CT concepts definitions created. **Conclusions:** The properties of SNOMED CT provided the necessary computable underpinnings to support a standards-based approach to capture, transmit and computationally assess syntopic data. Additional work is ongoing to encode all CAP cancer worksheets. Authored content is available from the NLM UMLS knowledge server.

### Next Generation Decision Support Tool for Variant Reporting

**Michael G. Zomnir**, **Jochen K. Lennerz**, **Maciej Pacula**, **Nishchal Nadhamuni**, **Lev Lipkin**, **Sekhar Duraisamy**, **Enrique Dominguez Meneses**, **Allison MacLeay**, **Saeed H. Al Turki**, **Zongli Zheng**, **Miguel N. Rivera**, **Valentina Nardi**, **Dora Dias-Santagata**, **Long P. Le**, **A. John Iafrate**

1. Center for Integrated Diagnostics, Massachusetts General Hospital, Boston, MA, USA. E-mail: mizomnir@andrew.cmu.edu

**Content:** Bioinformatics pipelines have markedly improved next-generation sequencing data analysis and genotyping; however, the decision to include a variant in the final report remains challenging. We aim to address this challenge by creating a next-generation decision support tool for variant reporting, using next generation sequencing data and metadata as our inputs. **Technology:** Here we employed a machine learning approach, leveraging the scikit-learn machine learning library of the Python programming language, to capture the collective clinical sign-out experience of six board-certified molecular pathologists and build a decision support tool for variant reporting. **Design:** We extracted all clinically reviewed and reported variants from our laboratory database and tested several classification models. We used ten-fold cross validation for our final variant call prediction model that derives a contiguous score from 0-1 (no-yes) for reporting. Through cross validation, we tested several different supervised machine learning approaches, including Naïve Bayes, Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines. We also performed feature selection to identify our most predictive independent variables. **Results:** For each of the 19,594 variants, our pipeline generates 507 features resulting in a matrix of roughly 9.9 million data points. From a comparison of Naïve Bayes, decision trees, random forests, and logistic regression models we selected the latter because logistic regression assigns individual coefficients for each feature, which increases interpretability. The model results in 1% false negatives and 2% false-positives. The final model’s Youden indices are 0.87 and 0.80 for screening and confirmatory cut-offs, respectively. Re-training the model on a different assay confirmed the transferability of the approach. We additionally derived individual pathologist-centric models (“virtual consensus conference function”) and a drill-down functionality allows review of the underlying features contributing to a particular score for clinical implementation. **Conclusion:** Our decision support tool for variant reporting is one approach to capture the clinical genomics sign-out experience.

### Web-based Facilitation of Communication and Coordination of Clinical Laboratory Services

**Aaron West**, **Michael Dowlin**, **Gregory Buffone**

1. Department of Pathology, Texas Children’s Hospital, 2. Department of Pathology, Baylor College of Medicine, Houston, TX, USA. E-mail: gbuffone@bcm.edu

**Content:** Clinical Laboratories, particularly those serving geographically distributed health care systems, demand an efficient and effective means of communication, coordination and documentation of various aspects of their services. Specifically, four key areas, including staffing, instrument readiness, patient and specimen handling and sharing, and shift-to-shift transitions all contribute to the level, efficiency and quality of service delivered by any given clinical laboratory. The ability to manage these dimensions of service in real-time requires a means of creating and sharing content relevant to each that can be proactively addressed at one or more operational and professional levels. **Technology:** A web-based application, Electronic Quality for Laboratories (EQL), has been created to address the aforementioned needs of the Pathology Laboratories serving the Texas Children’s health system in Houston, Texas. EQL links workbenches,
sections, shifts, laboratories and campuses across the system in near real-time (screen refresh rate 1 min) for the purposes of communication and coordination along the dimensions of service highlighted above. The facility to share information in this manner not only improves immediate outcomes but also provides documentation for ongoing quality assurance and regulatory review. Checklist for shift-to-shift hand-off and for task management is supported and can be customized to address different functions. **Design:** Real-time communication and coordination is organized along several interrelated hierarchies: operational structure, technical and administrative roles and professional roles are logically linked in the workflow and information sharing functionality. EQL also accommodates campus and discipline-specific (e.g., General Laboratory, Histology, Blood Bank) operational integration. The use of EQL reduces 15 different means of logging and communicating operation issues or potential issues to five for a 66% overall decrease. **Results:** Additionally, much of the data captured in EQL is now available in a structured form for analysis and reporting to affect operational and or quality improvement. For example, issues logged in EQL can be designated as a potential causal event. Those events can be programmatically associated with performance outcomes defined in an interfaced Quality Dashboard. Ongoing systematic association of events and outcomes provides a means of not only simplifying the quality review process for laboratories but also enables efficient and effective analysis and improvement of systems failure. **Conclusion:** EQL linked to a dashboard monitoring operations/quality provides the facility to move beyond simple discovery of poor outcomes to concurrent understanding of potential causes.

### Development of a Novel Quality Assessment Tool for Digital Microscopy

**Emily L. Clarke**1,2, Alexander Sykes1, David Brettle2, Alexander Wright3, Anna Boden1, Darren Treanor1,2,3,4

1Pathology, University of Leeds; 2Pathology, Leeds Teaching Hospitals NHS Trust, Leeds, UK; Departments of 3Clinical Pathology and 4Clinical and Experimental Medicine, Linköping University, Linköping, Sweden. E-mail: e.l.clarke@leeds.ac.uk

**Content:** Image fidelity is of utmost importance when using digital pathology for routine diagnostic work. Digital radiology has developed many Quality Assurance (QA) measures to help ensure that the display is suitable for diagnosis, but little work has been conducted within the corresponding area of digital pathology. An international, cross-site audit was conducted to evaluate the performance of displays and viewing environments within two tertiary histopathology departments using a unique QA tool. **Technology:** We designed a QA tool based on the RGB values of haematoxylin and eosin obtained from analysing the spectral data of stained biopolymer. The tool was developed using MATLAB, and involved a 5x5cm haematoxylin and eosin coloured patch with a superimposed random letter of progressively varying RGB values from the background colour. The test environment was created using the previously published web-based experiment platform, Prospector, and involved 180 test images shown individually and in random order. **Design:** Two pathology departments were involved in the audit; one where most primary diagnoses are made digitally using medical grade displays, and one that uses light microscopes only. The audit was divided into three phases across the two institutions. Participants were enrolled on an opportunistic basis. A total percentage correct for each test was obtained for each participant. **Results:** Eleven participants completed 11 tests in Phase 1. Phase 2 was carried out by 6 participants who completed 16 tests in total. Phase 3 included 26 tests conducted by 6 participants. The results will be presented, including comparison of performance between the two institutions, comparison between different displays, comparisons between participants, and variation in performance with changes in the environmental conditions. **Conclusions:** The use of the QA tool provides a standardized method of comparing displays and viewing environments. Our findings are in accordance with existing International Colour Consortium recommendations regarding the importance of ambient light consistency when using digital pathology for primary diagnosis. Further development of the QA tool will be conducted.

### A Machine-learning Model for Personalized Trial Data Exploration

**Maciej Pacula**1, Lev Lipkin1, Enrique Dominguez Meneses1, A. John Iafrate1, Long P. Le1, Jochen K. Lennerz2

1Department of Pathology, Center for Integrated Diagnostics, Massachusetts General Hospital, Boston, MA 02114, USA. E-mail: mpacula@mgh.harvard.edu

**Content:** Benefits and risks for individual patients cannot be easily extrapolated from aggregated clinical trial data. We construct and validate a tool that uses artificial intelligence to predict the individualized risks and benefits of changing blood-pressure medication from a standard to an intensive regimen using the SPRINT trial data. We then use the model to identify two clinically relevant populations with drastically different risk-benefit profiles. Finally, we make the tool available as a mobile/web application: http://mpacula.mgh.harvard.edu

**Technology:** We used machine learning to derive a weighted k-nearest neighbor model which can identify similar patients and use those patients to estimate the hazard ratios for ‘benefits’ and ‘risks’ as originally defined in the SPRINT trial [Figure 1]. **Design:** The data matrix of the SPRINT trial contains >1,000,000 data points distributed over >28 variables. We used a randomly selected 80% subset (N=7488) of the SPRINT data as a training set, then validated the predictions on
a held-out set of 20% (N=1873) patients. **Results:** The model demonstrated a statistically significant separation on the held-out set. Briefly, within the standard treatment, the high-risk group had a 2.04x higher event rates for the primary events (p<0.01) and 1.47x higher rates of adverse events (p<0.001); in the intensive group the numbers were comparable with 2.05x (p<0.001) and 1.41x (p<0.001), respectively. These unbiased validation data confirm that our model can accurately predict unseen data. We further used the model to discover that the risk ratio for adverse events from intensive therapy differs in women with/or without hypercholesterolemia whereas male patients with a history of clinical cardiovascular disease and risk factors have the greatest benefit with relatively few adverse events. **Conclusion:** In conclusion we report an artificial intelligence approach to employ shared trial data for personalized data exploration. We envision that similar tools will become a powerful approach to explore shared trial data to improve patient care.

**Figure 1:** The SPRINT trial (a) compared two regiments (standard vs. intensive) for lowering blood pressure. Using the published trial data (b) we developed and validated a patient tool (c) and a data exploration app (d) that calculates risk and benefits for individual patients.

---

**Bringing the Blood Bank to the Bedside: Multi-institutional Evaluation of Gaps in Positive Patient ID Systems for Blood Administration**

**Content:** Positive patient identification for blood administration (PPID-BA) systems have improved safety by preventing wrong blood products from being transfused. However, poor technology and/or implementation strategy may fail to improve or worsen safety and can jeopardize compliance with regulations. Guidance is needed for information technology, transfusion, and nursing staff to improve utilization and for software developers to reduce existing functionality gaps. **Technology:** Various PPID-BA systems, blood bank information systems and electronic health records served as sources of the guidance. All PPID-BA systems were 510(k) approved by the FDA. **Design:** A multi-institutional group of pathology informaticists convened to create guidance on selecting and implementing PPID-BA systems. Recommendations were categorized into groups and sequenced according to typical workflow. Only those elements which were determined to be system-agnostic by consensus were selected for inclusion. **Results:** The categories of guidance were as follows: Design, Implementation, Validation and Training. Design was broken into subcategories of why, who, what, where, when and how in which the “how” was workflow design. Workflow design was the largest subcategory of design with the following phases: pre-transfusion specimen collection, patient armbands, blood product barcode usage, PPID-BA software algorithm, transfusion documentation, post-transfusion auditing and downtime preparation. **Conclusions:** PPID-BA systems are considered by some to be primarily a nursing tool, but transfusion services feed these systems with data and have the most responsibility for ensuring regulatory compliance. Each PPID-BA system should have a set of standard basic functions including pre-built quality assurance reports, audits and safety monitors. The group unanimously felt that both the ISBT 128 donor identification number and ISBT 128 product code should be required to match between the scanned barcodes on the blood product label and the data present in the PPID-BA system that is acquired from the blood bank information system. This is critical to patient safety because a single blood product may be split into modified (e.g., washed, irradiated) and non-modified components. Third party safety reports for PPID-BA systems are warranted to help institutions pick the right software. Future work may include surveying users of PPID-BA systems to determine what other guidance may be necessary.