ABSTRACTS

INFORMATICS
(1522-1590)

USCAP 109TH ANNUAL MEETING
2020
EYES ON YOU

FEBRUARY 29-MARCH 5, 2020

LOS ANGELES CONVENTION CENTER
LOS ANGELES, CALIFORNIA

EDUCATION COMMITTEE

Jason L. Hornick, Chair

Rhonda K. Yantiss, Chair, Abstract Review Board
and Assignment Committee

Laura W. Lamps, Chair, CME Subcommittee

Steven D. Billings, Interactive Microscopy Subcommittee

Raja R. Seethala, Short Course Coordinator

Ilan Weinreb, Subcommittee for Unique Live Course Offerings

David B. Kaminsky (Ex-Officio)

Zubair Baloch
Daniel Brat
Ashley M. Cimino-Mathews
James R. Cook
Sarah Dry

William C. Faquin
Yuri Fedoriw
Karen Fritchie
Lakshmi Priya Kunju
Anna Marie Mulligan
Rish K. Pai
David Papke, Pathologist-in-Training
Vinita Parkash
Carlos Parra-Herran
Anil V. Parwani
Rajiv M. Patel
Deepa T. Patil
Lynette M. Sholl
Nicholas A. Zoumberos, Pathologist-in-Training

ABSTRACT REVIEW BOARD

Benjamin Adam
Narasimhan Agaram
Rouba Ali-Fehmi
Ghassan Allo
Isabel Alvarado-Cabrero
Catalina Amador
Roberto Barrios
Rohit Bhargava
Jennifer Boland
Alain Borczuk
Elena Brachtel
Marilyn Bui
Eric Burks
Shelley Caltharp
Barbara Centeno
Joanna Chan
Jennifer Chapman
Hui Chen
Beth Clark
James Conner
Alejandro Contreras
Claudiu Cotta
Jennifer Cotter
Sonika Dahiya
Farbod Darvishian
Jessica Davis
Heather Dawson
Elizabeth Demicco
Katie Dennis
Anand Dighe
Suzanne Dintzis
Michelle Downes
Andrew Evans
Michael Feely
Dennis Firchau
Gregory Fishbein
Andrew Folpe
Larissa Furtado

Billie Fyfe-Kirschner
Giovanna Giannico
Anthony Gill
Paula Ginter
Tamara Giorgadze
Purva Gopal
Anuradha Gopalan
Abha Goyal
Rondell Graham
Alejandro Gru
Nilesh Gupta
Mamta Gupta
Gillian Hale
Suntrea Hammer
Malini Harigopal
Douglas Hartman
John Higgins
Mai Hoang
Mojgan Hosseini
Aaron Huber
Peter Illei
Doina Ivan
Wei Jiang
Vickie Jo
Kirk Jones
Neerja Kambham
Chiah Sui Kao
Dipti Karamchandani
Darcy Kerr
Ashraf Khan
Francesca Khani
Rebecca King
Veronica Klepeis
Gregor Krings
Asangi Kumarapeli
Alvaro Laga
Steven Lagana
Keith Lai

Michael Lee
Cheng-Han Lee
Madelyn Lew
Zaibo Li
Faqian Li
Ying Li
Haiyan Liu
Xiuli Liu
Yen-Chun Liu
Lesley Lomo
Tamara Lotan
Anthony Magliocco
Kruti Maniar
Emily Mason
David McClintock
Bruce McManus
David Meredith
Anne Mills
Neda Moatamed
Sara Monaco
Atis Muehlenbachs
Bita Naini
Dianna Ng
Tony Ng
Michiya Nishino
Scott Owens
Jacqueline Parai
Yan Peng
Manju Prasad
Peter Pytel
Stephen Raab
Joseph Rabban
Stanley Radio
Emad Rakha
Preetha Ramalingam
Priya Rao
Robyn Reed
Michelle Reid

Natasha Rekhtman
Jordan Reynolds
Michael Rivera
Andres Roma
Avi Rosenberg
Esther Rossi
Peter Sadow
Steven Salvatore
Souzan Sanati
Anjali Saqi
Jeanne Shen
Jiaqi Shi
Gabriel Sica
Alexa Siddon
Deepika Sirohi
Kalliopi Siziopikou
Sara Szabo
Julie Teruya-Feldstein
Khin Thway
Rashmi Tondon
Jose Torrealba
Andrew Turk
Evi Vakiani
Christopher VandenBussche
Paul VanderLaan
Olga Weinberg
Sara Wobker
Shaofeng Yan
Anjana Yeldandi
Akihiko Yoshida
Gloria Young
Minghao Zhong
Yaolin Zhou
Hongfa Zhu
Debra Zynger

1522 Are Medical Grade Monitors a Necessity for Digital Pathology? Initial Insights between Commercial Grade and Medical Grade Displays

Jacob Abel[1], Peter Ouillette[2], Lloyd Stoolman[2], Ulysses Balis[2], David McClintock[2]
[1]Michigan Medicine, University of Michigan, Ann Arbor, MI, [2]University of Michigan, Ann Arbor, MI

Disclosures: Jacob Abel: None; Peter Ouillette: None; Lloyd Stoolman: None; Ulysses Balis: None; David McClintock: None; David McClintock: None

Background: The display (monitor) is an integral part of the digital pathology (DP) workflow, with DP vendors required to include a specific display as part of FDA validation studies. Following radiology's footsteps, DP vendors are choosing medical grade (MG) displays advertising greater quality as compared to commercial grade (CG) displays, but at a much higher cost. MG displays are reputed to achieve greater luminance (light emitted per square area) and allow more precise control of luminance relative to CG monitors. Here, we present a pilot study comparing the luminance characteristics of two FDA-cleared DP displays (Dell MR2416 and Philips PP27QHD) to a high-end CG display (ASUS PG279Q) and typical laptop CG display (HP G5 ZBook) (Table 1).

Design: Luminance and contrast measurements were taken with an X-Rite PANTONE i1Basic Pro 2 spectrophotometer using i1 Display Pro software in a photography dark room. The displays were calibrated to 250 cd/m^2, with the exception of the laptop display, which was set to its max luminance due to setting limitations. Three sets of measurements were taken daily over a period of 5 days, performed approximately three hours apart (10AM, 1PM, and 4PM). Luminance was measured across 9 "zones" of each display, each time with the lights on and then off. Statistical analysis was performed in Microsoft Excel. Additionally, delta E-2000 values measuring the difference in relative luminance between the 8 outer zones and the central zone were calculated to assess lighting uniformity across the panel.

Results: All displays deviated significantly from the set point of 250 cd/m^2 (p-value <0.01 by one-sample T test). Statistically significant differences in luminance uniformity were found between the CG and MG displays (ANOVA and two-sample T test, p-values <0.01). The Philips PP27QHD offered the greatest max luminance followed by the ASUS, while the Dell offered the least.

Table 1: Specifications given by the manufacturer for each of the studied displays.

Manufacturer	Model	Size (inches)	Native Resolution (pixels)	Luminance Specification (cd/m^2)	Contrast Ratio
Dell	MR2416	24"	1920 x 1200	300 (180 with DICOM preset)	1000:1
ASUS	PG279Q	27"	2560 x 1440	350	1000:1
Hewlett-Packard	Zbook 17 G5	17"	1920 x 1080	300	Not Given
Philips	PP27QHD	27"	2560 x 1440	500 (350 with DICOM preset)	1000:1

Conclusions: The MG displays offered statistically significant, but relatively modest, improvements in luminance uniformity as compared to the CG displays. The Dell MG display's maximum luminance was limited as compared to the three other displays, including the two CG displays assessed. Given that the MG displays offer only minor superiority in the characteristics assessed in this study but at a much greater cost, further evaluation of whether these factors will affect pathologists' performance in a clinical setting is warranted.

1523 How Much Work Does a Surgical Pathologist Do to Generate One RVU (Relative Value Unit)? A Cross-Subspecialty Study in an Academic Tertiary Care Center

Syeda Absar[1], Jeff Prichard[2]
[1]Geisinger Medical Center, Danville, PA, [2]Geisinger, Danville, PA

Disclosures: Syeda Absar: None; Jeff Prichard: None

Background: A common metric for measuring physician work is the Relative Value Unit (RVU) determined by Medicare based on billing. A previous study conducted at our center concluded that a cytopathologist generated 12,107 RVUs in one year of work (Mestchter S, et al. Meeting Abstracts USCAP 2015). Other proposed metric of pathologist work is block, slide and level unit count. We designed this study to determine the relationship between RVU and the number of blocks, slides and levels a subspecialty pathologist needs to sign-out to earn one RVU.

Design: Our laboratory information system (Cerner CoPathPlus) was queried to determine the RVU sum per case generated by a pathologist, and the associated number of blocks, slides and levels signed-out over a period of two years from September 2017 through the end of August 2019. Eleven surgical pathology subspecialties were evaluated: Breast, Skin, ENT, GI, Gyne, Liver, Neuro, Soft Tissue/Bone, Thoracic, Thyroid and Urologic. In each subspecialty, we divided the sum of RVUs by the sum of blocks, slides and levels that were signed out in 2 years. These ratios determined the numbers needed of each to generate one RVU.

Results: Our data shows that breast pathologists read similar amounts slide levels per RVU as thyroid pathologists (p = 0.24), but significantly more slides and levels to generate one RVU than most other surgical pathology subspecialties including GI (p = 0.000000007) (Table 1). Conversely, neuropathologists read the lowest number of blocks, slides and levels per RVU.

Subspecialty	Blocks/ RVU	Slides/ RVU	Levels/ RVU
Breast	1.74	4.61	4.62
Thyroid	3.34	4.55	4.56
GI	1.15	1.90	3.91
Derm	1.54	3.00	3.77
Urologic	0.98	2.18	3.57
Gyne	2.29	2.94	3.23
Liver	1.08	2.90	3.04
Thoracic	0.98	2.66	2.73
ENT	1.28	2.18	2.32
Soft Tissue/Bone	1.82	2.08	2.20
Neuro	0.63	2.04	1.97
Grand Total	**1.44**	**2.60**	**3.65**

Conclusions: Our results show that there are statistically significant differences in the number of slides levels read by different subspecialties for each RVU generated. We are proposing the slide levels can be used as an alternative and independent metric of surgical pathology workload.

1524 Detection and Classification of Organism and Organism-like Profiles in Gastric Biopsy Helicobacter Immunohistochemistry by Digital Image Analysis (DIA)

Ibrahim Abukhiran[1], Andrew Bellizzi[1], Anand Rajan KD[1]
[1]University of Iowa Hospitals and Clinics, Iowa City, IA

Disclosures: Ibrahim Abukhiran: None; Andrew Bellizzi: None; Anand Rajan KD: *Advisory Board Member*, Roche Diagnostics Corporation

Background: Helicobacter pylori (*H.pylori*) immunohistochemistry (IHC) is a commonly ordered test in gastrointestinal pathology. Screening HP-IHC slides can be time-consuming, particularly in treated cases and in target patient populations with a low incidence rate. In these scenarios, large volumes of gastric biopsy tissue needs to be examined and the clinical impact of positive organism detection is abruptly high. We sought to utilize whole slide quantitative image analysis (QIA) by HALO (Indica Labs, New Mexico, USA) to develop a screening protocol for organism detection with high sensitivity and negative predictive value.

Design: We assembled H.pylori negative (n=115) and non-random positive cases (n=65) into a cohort (n=180) with organism incidence well above baseline population positivity rate (~5%). Whole slides were scanned at 20x (0.24μm/pixel) and presence of organisms in positive cases was confirmed by glass slide review. A HALO area quantitation (AQ) protocol was calibrated to detect IHC di-aminobenzidene (DAB) brown organism signals in a 3+ bracket and nucleocytoplasmic counterstain thresholding was set to blank. Receiver operator characteristic (ROC) analysis was performed to identify the optimal area value cutoffs.

Results: HALO QIA was able to detect organisms in 100% (65/65) of positive cases. It did not detect any organisms in the negative cases. However, it identified signal positivity (artifacts) in all cases. In positive cases, organisms tend to cover more surface area than the area of artifacts. The most common encountered artifacts were: acellular debris/mucin (29%), dark stained lymphocytes (23%) stain artifact/precipitate (23%). With ROC analysis, the area under the curve (AUC) was 0.863 (CI: 0.80-0.91, *p <0.0001*).

Figure 1 - 1524
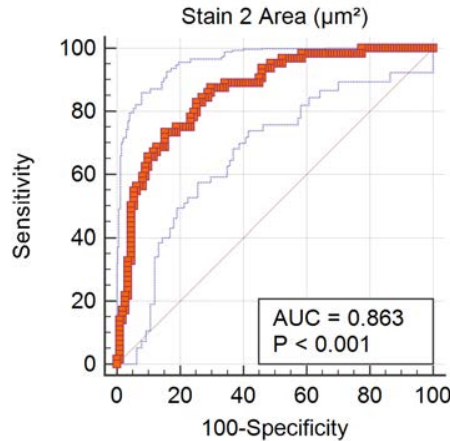
Conclusions: Whole slide QIA was 100% sensitive in detecting *H.pylori* organisms in gastric biopsies. With ROC analysis, we identified cutoffs that exhibiting high sensitivity (98.4%), and high negative predictive value (99.8%) with low cost (2.8 units relative to cost of false negative set at 100). In signal-positive organism-negative cases the detected artifacts were of a very small area (less than 0.1% of total tissue) and could be quickly reviewed and excluded by morphologic review. The high negative predictive value strongly supports the use of a simple QIA protocol in screening gastric biopsy IHC in routine practice.

1525   Medicare Trends in Pathologist Participation, Service Utilization and Payments

Simone Arvisais-Anhalt[1], Ellen Araj[1], Jason Park[1]
[1]University of Texas Southwestern Medical Center, Dallas, TX

Disclosures: Simone Arvisais-Anhalt: None; Ellen Araj: None; Jason Park: None

Background: The utilization and payment for pathology and laboratory services has not been previously reported at a national scale. Quantifying pathologist participation in Medicare services may be informative for the prediction of future utilization and payments.

Design: The Medicare Provider Utilization and Payment Data: Physician and Other Supplier Public Use File (POSPUF) from years 2012 to 2017 was aggregated. This dataset contains information on Part B services and procedures provided to Medicare beneficiaries; the data includes NPI, CPT codes, payments and place of service. Data aggregation and analysis were performed using the business intelligence software Tableau 2019.2. Physicians examined were limited to those who self-identified their specialty as "Pathology". Services and payments were examined by facility (e.g., hospital, ambulatory surgery center) and non-facility (e.g., independent laboratory, medical office) place of service.

Results: In 2017, 11,191 pathologists (female=4,533, 40.5%) provided Medicare Part B services. This is a 2.9% overall increase in pathologists from 2012; a 15.4% increase in females and a 4.2% decrease in males was observed. From 2012 to 2017, the Medicare beneficiaries receiving pathology services increased from 15,078,433 to 16,866,980. Normalized per pathologist, this is an increase from 1,386 up to 1,507. In the same period, pathology services increased from 26,634,848 to 28,791,595. Normalized per pathologist, this is an increase from 2,448 up to 2,572 services. Medicare Part B payments for all medical specialties was $68,889,358,763 in 2017; a 7.8 percent increase from 2012. In 2017, payment for pathology services was $936,201,035; a 3% percent decrease from 2012. The top ten pathology services performed in a facility were all surgical pathology. In all years surveyed, the 88305 CPT code was the most common service provided. From 2012 to 2017, the facility payment for 88305 increased from $28.70 to $30.60; in contrast the 88305 non-facility payment decreased from $55.00 to $39.00. Declines in non-facility payments were observed for other surgical pathology services.

Conclusions: From 2012 to 2017, the number of pathologists providing Medicare services increased by 2.9%. However, both the beneficiaries as well as total services increased per pathologist. Medicare payments decreased for the highest volume non-facility surgical pathology services.

1526 Automated Analysis of Plasma Cell Percentage and Light Chain Clonality in Bone Marrow Biopsy Whole Slide Images

Vahid Azimi[1], Ngoc Tran[2], Young Hwan Chang[2], Guillaume Thibualt[2], Eva Medvedova[2], Kevin Turner[2], Robert Christian[2], Phil Raess[2]
[1]Oregon Health & Science University, Portland, Multnomah, [2]Oregon Health & Science University, Portland, OR

Disclosures: Vahid Azimi: None; Ngoc Tran: None; Young Hwan Chang: None; Eva Medvedova: None; Robert Christian: None; Phil Raess: None

Background: Plasma cell myeloma (PCM) is diagnosed by visual estimation of plasma cell percentage (PC%) using CD138 immunohistochemistry (IHC) and kappa/lambda (K/L) light chain ratio using K/L in-situ hybridization (ISH) in bone marrow biopsies (BMBx) by pathologists; the cutoff for differentiating PCM from monoclonal gammopathy of undetermined significance (MGUS) is a PC% estimation of 10%. These estimations may be prone to intra- and interobserver variability; thus, development of an automated WSI-based method to standardize quantitation of PC% and clonality from BMBx could lead to more accurate prognostication and diagnosis. To our knowledge, this is the first study implementing a completely automated computer-assisted imaging analysis (CIA) method of quantifying PC% as well as K/L ISH in BMBx.

Design: H&E, CD138, and K/L ISH of BMBx were retrospectively collected and scanned from 42 patients at day +100 following autologous stem cell transplantation. Pathologist-estimated PC% and light chain clonality was extracted from pathology reports. A previously described nucleus segmentation algorithm was modified to identify PCs on CD138 IHC and K/L ISH stains. K/L ratio was calculated by # K+ cells/# L+ cells on K/L ISH, respectively. Cutoffs of >3.5 and <0.5 were used to define monoclonal kappa and lambda, respectively.

Two methods to calculate PC% were used: 1) # CD138+ cells on IHC/# nucleated cells on H&E, and 2) total area of CD138+ cells/total area of CD138+ and CD138 negative cells on IHC. (1) is the most direct measure of PC%, while (2) reduces the variation caused by using adjacent tissue sections. Area was used instead of cells for (2) due to challenges in CD138- cell segmentation.

Results: CIA PC% using method 1 (Figure 1) and method 2 (Figure 2) shows 86% (36/42) and 90% (38/42) concordance with pathologists in distinguishing between MGUS and PCM (using a 10% cutoff), respectively. Additionally, the K/L ISH algorithm demonstrates 91% (39/43) and 79% (34/43) concordance with pathologists in identifying K and L monoclonality, respectively (Table 1).

Table 1.

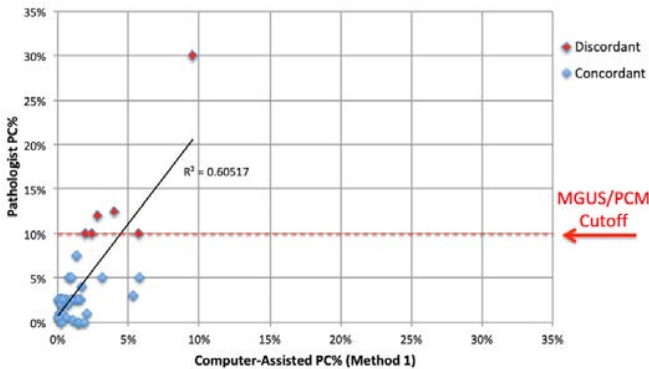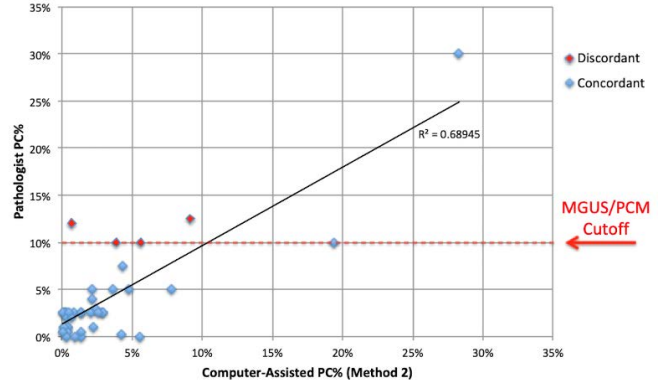| Monoclonal KAPPA by ISH	Visual estimation by pathologists	
Computer-assisted image analysis	K/L > 3.5	0.5< K/L < 3.5
K/L > 3.5	4	4
0.5< K/L < 3.5	0	35
Monoclonal LAMBDA by ISH	Visual estimation by pathologists	
Computer-assisted image analysis	K/L < 0.5	0.5< K/L < 3.5
K/L < 0.5	4	8
0.5< K/L < 3.5	1	30

Figure 1 - 1526

Figure 2 - 1526

Conclusions: PC% and K/L clonality CIA algorithms show good concordance with pathologists. Differences in concordance between the two PC% methods could be due to H&E cell oversegmentation, CD138+ cell undersegmentation, or overestimation of PC% by

pathologists. Future work will further investigate reasons for CIA and pathologist discordance and assess the relative prognostic values of CIA and pathologist PC% estimations.

1527 Detecting Specimen Contamination in Whole Slide Imaging Using Artificial Intelligence

Morteza Babaie[1], Abtin Riasatian[2], Sobhan Hemati[2], Mahjabin Sajadi[2], Adrian Batten[3], Soma Sikdar[3], Liron Pantanowitz[4], Hamid Tizhoosh[1]

[1]University of Waterloo, Waterloo, ON, [2]KimiaLab, University of Waterloo, Waterloo, ON, [3]Grand River Hospital, Kitchener, ON, [4]University of Pittsburgh, Wexford, PA

Disclosures: Morteza Babaie: None; Abtin Riasatian: None; Sobhan Hemati: None; Mahjabin Sajadi: None; Adrian Batten: None; Soma Sikdar: None; Liron Pantanowitz: *Consultant*, Leica; *Consultant*, Hamamatsu; *Consultant*, Ibex; *Grant or Research Support*, Lunit; *Grant or Research Support*, Huron; Hamid Tizhoosh: *Consultant*, Huron Digital Pathology; *Primary Investigator*, Huron Digital Pathology

Background: Recent developments in digital pathology have opened new horizons for research and diagnostic purposes. Besides obvious advantages of digitization (e.g., telepathology), many components of pathology workflow could be improved by applying artificial intelligence (AI) techniques. The presence of artifacts from various sources during slide preparation could potentially render a slide useless or difficult to read for diagnosis. Among others, specimen contamination is considered a serious artifact. The finding of suspected contaminating tissues could be a challenging obstacle in diagnosis; detecting a dissimilar (foreign) tissue fragment could take a considerable time for the pathologist to resolve. We demonstrate the use of an AI-based system to detect tissue contaminations ("floaters") in digitized slides.

Design: We collected 12 clinically confirmed tissue contamination cases (6 cases confirmed by DNA testing) from the Grand River Hospital pathology archive. After scanning slides at 40x magnification (by Huron's TissueScope LE), each fragment was split into several 400x400 pixel patches. All patches were fed into the DenseNet 121 (a pre-trained deep network) that provided a descriptor vector with a length of 1024. The most similar retrieved patches were selected based on the descriptors' distances (pattern dissimilarity) between all other patches of each fragment. By taking the average among distances of each fragment, we constructed a complete bidirectional graph for each slide. Distances from fragment A to B is not equal to the distance from B to A because of a different number of patches in each fragment.

Results: We applied a linear SVM algorithm to classify the scaled average distances from and to each fragment (two features for each fragment). All experiments have been conducted using the leave-one-patient-out validation. An accuracy of 87% was achieved for contamination detection. The sensitivity of the proposed approach was 100% whereas the specificity was observed at 83%. The false-positive cases were sometimes blurry fragments. This means our approach may also detect other artifacts.
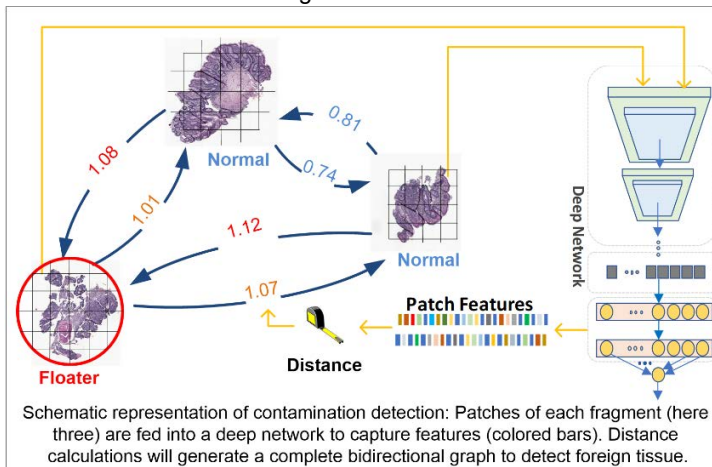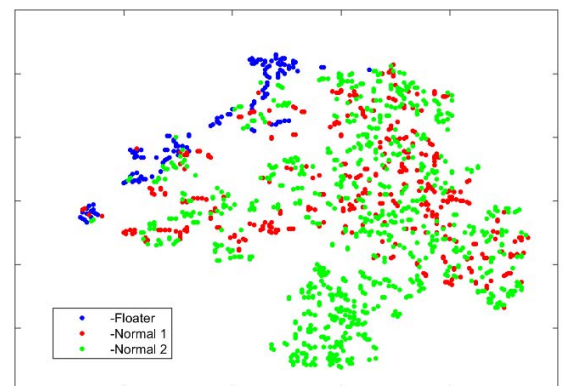
Figure 1 - 1527

Schematic representation of contamination detection: Patches of each fragment (here three) are fed into a deep network to capture features (colored bars). Distance calculations will generate a complete bidirectional graph to detect foreign tissue.

Figure 2 - 1527

Visualization of a sample slide patches similarity(with three fragments) in a 2-D spaces. t-SNE has applied to all descriptor vectors to reduce the dimension from 1024 to two.

Conclusions: Manual detection and clarification of contaminations such as floaters present in pathology slides can be a laborious task which is generally accompanied by ambiguity. Automated contamination detection for whole slide imaging in digital pathology can assist pathologists and lab technicians with this task, thereby facilitating high-quality slide processing as well as reliable diagnoses.

1528   Subpar Scanning Fidelity of Mastectomy Specimens Poses Challenges for the Adoption of Digital Pathology for Primary Diagnosis

Sarah Bowman[1], G. Zoltan Laszik[1]
[1]University of California San Francisco, San Francisco, CA

Background: High fidelity scanning of histology glass slides is one of the prerequisites for successful adoption of digital pathology for primary diagnosis. Fat-rich and faintly-stained tissues are known to pose potential problems for whole slide scanning. To assess the scope of the problem we evaluated scanning fidelity on a select set of mastectomy cases.

Design: Six randomly selected mastectomy cases from the routine surgical pathology files at UCSF from 2018 with a total of 457 slides were enrolled into the study. All slides were scanned in on Philips Ulta Fast Scanners at UCSF before and post-calibration to factory specifications. The images were reviewed by a trained image specialist (SB) and also by a pathologist (ZL) to determine the incidence of suboptimal scanning. Of the 163 slides with suboptimal scanning fidelity at UCSF, 161 slides were rescanned at Philips, Best, NL and again evaluated for the incidence of suboptimal scanning. Images with less than 100% scanning fidelity were further classified based on how much tissue was missing from the scans as follows: 1) <10mm (small area), 2) 10-30mm (medium area), and 3) >30mm (large area).

Results: The incidence of suboptimal scanning fidelity was comparable at UCSF before (152/429 [35%]) and post-calibration (163/429 [38%]) of the scanners. Of the161 slides with subpar scanning fidelity at UCSF, 81% were also suboptimal upon rescanning at Best, NL. The incidence of various classes of subpar scans at UCSF and at Best, NL is shown in Table 1. The tissues missed by the scanner were exclusively those of fat-rich portions of the sections.

Table 1. The incidence of various classes of subpar scans at UCSF and at Philips, Best, NL		
	UCSF (n=163)	**Philips, Best, NL (n=161)**
100% of Tissue Scanned	0/163 (0%)	31/161 (19%)
Category 1	62/163 (38%)	57/161 (35%)
Category 2	48/163 (29%)	24/161 (15%)
Category 3	53/163 (33%)	49/161 (30%)

Conclusions: High incidence of suboptimal scanning fidelity might hamper adoption of digital pathology for primary diagnosis for mastectomy specimens. Improvements in scanning fidelity for fat-rich tissues is warranted.

1529   Establishing a Radiology-Pathology Division – A Novel Approach Integrating Three Diagnostic Digital Imaging Modalities to Resolve Uncertain Diagnosis

Andrey Bychkov[1], Takashi Hori[2], Akira Yoshikawa[3], Yoko Masuzawa[2], Akiko Shimauchi[2], Wataru Yamashita[2], Youichi Machida[2], Junya Fukuoka[4]
[1]Kameda Medical Center, Kamogawa, Japan, [2]Kameda Medical Center, Kamogawa, Chiba, Japan, [3]Kamogawa, Japan, [4]Nagasaki University, Nagasaki, Japan

Background: Pathology and radiology are core disciplines of so-called medical imaging. There is a growing voice from the leading authorities that medical imaging can be a single specialty soon, combining expertise and advances of both disciplines. Digital technology is a medium greatly facilitating this potential switch. Radiology was universally digitized in the early 2000s. Now pathology follows the same path. While pathological diagnosis remains the gold standard, radiological findings can be essential clues to reach a correct diagnosis. Direct communication between both specialties in the frame of the radiological-pathological conference (RPC) provides an invaluable solution for challenging cases in clinical practice.

Design: Our department serves as a model facility for the primary diagnosis through digital pathology. In 2018 we achieved 100% digital workflow for biopsies, surgicals, and frozen service. In August 2018 we established a Radiology-Pathology Division, which integrated three digital imaging tools (radiology, pathology, and cytology) in a single group discussion via secure WebEx channel. The Philips IntelliSite Pathology Solution and Panoptiq platform were used for digital pathology and digital cytology, respectively.

Results: A total of 130 cases were discussed at the RPC by 09/2019, including 47 cases recorded in the standard protocol (introduced in April 2019). Baseline numbers were as follows: 2 cases/week, 20–40 min./case. Regular participants from 2–3 locations included radiologists, pathologists, senior cytotechnologist, and residents (average 5–6 people). Two-thirds of cases were referred by pathologists, and one third by radiologists. Clinical diagnosis concerned cancer in 94%. Histology and radiology digital images were provided for all cases, added by cytology in 72% cases. The vast majority of tissue samples were biopsies (85%). Main locations were thoracic (49%), abdominal (26%), and GU (15%). The final diagnosis of malignancy was rendered in 81% and excluded in 19% of cases. Histology and

cytology were initially discordant in 15% of cases. The most common decisions made at RPC were arriving at the definite diagnosis, and expanding/narrowing the differential diagnosis. Additional IHC workup was a common solution to render the final diagnosis.

Conclusions: RPC proved useful in diagnostically challenging cases. Its clinical utility was added by the educational impact since both pathologists and radiologists gained knowledge about basic diagnostic clues of each specialty.

1530   Loss of Fidelity in Whole-Slide Images Compared to Glass Slides of Brain Tumors Resected Using Cavitron Ultrasonic Surgical Aspirator
Cathryn Cadwell[1], Sarah Bowman[1], G. Zoltan Laszik[1], Melike Pekmezci[1]
[1]University of California San Francisco, San Francisco, CA

Disclosures: Cathryn Cadwell: None; Sarah Bowman: None; G. Zoltan Laszik: None; Melike Pekmezci: None

Background: Whole-slide images (WSI) hold tremendous potential for improving clinical care, research, and medical education. However, due to restrictions in the scannable area and selective scanning of regions of interest, highly fragmented and/or faintly stained tissue may not be scanned at high resolution leading to loss of fidelity in WSI compared to glass slides. Cavitron Ultrasonic Surgical Aspirator (CUSA) is often used in brain tumor resections, resulting in highly fragmented specimens. This study evaluated the extent and significance of loss of fidelity in WSI from CUSA-resected brain tumor specimens.

Design: We reviewed 296 slides from 40 CUSA-resected brain tumor cases scanned using a Philips Ultra Fast Scanner. Twenty each were selected from 2016 and 2018, between which our institution made extensive tissue processing modifications to optimize for WSI. The WSI and glass slides were reviewed by two pathologists and classified into one of three categories: 1) no loss of fidelity, 2) unscanned fragments only beyond the coverslipped area, or 3) unscanned fragments within the coverslipped area. The size of the missing fragments and potential diagnostic impact were also examined.

Results: Overall, 38% of the examined WSI showed no loss of fidelity, 8% were missing tissue only beyond the coverslipped area, and 54% were missing tissue within coverslipped area (Table 1). The largest size of unscanned tissue fragments in the coverslipped area was <0.5 mm in 36% of slides, 0.5–5 mm in 17%, and >5 mm in 1%. Of the slides with unscanned tissue in the coverslipped area, 19% showed no indication of loss of fidelity in the WSI and could only be identified on review of the glass slides. There was no difference in fidelity between the 2016 and 2018 cases. All cases had at least one slide with unscanned tissue, but none of the missing fragments would have altered the final diagnosis.

	Total (n=296 slides)					
	No missing tissue	Tissue beyond cover slip	Small fragment missed (<0.5 mm)	Intermediate fragment missed (0.5 - 5 mm)	Large fragment missed (≥ 5 mm)	All missed fragments
All missing fragments detected on WSIS	N/A	N/A	40	25	2	67
Some missing fragments detected on WSIS, additional on glass slide only	N/A	N/A	41	21	0	62
Missing fragments detected only on glass slide	N/A	N/A	27	4	0	31
Total	**113** (38%)	**23** (8%)	**108** (36%)	**50** (17%)	**2** (1%)	**160** (54%)

Conclusions: Our results highlight a potential limitation of WSI for the evaluation of fragmented brain specimens obtained using CUSA, with the majority of WSI showing loss of fidelity compared to glass slides. The lack of improvement from 2016 and 2018, despite extensive tissue processing modifications, suggests that changes to the current scanning restrictions may be necessary to overcome this loss of fidelity. Although none of the unscanned fragments would have altered the final diagnosis in these 40 cases, this loss of fidelity may represent a potential liability in the clinical application of WSI to CUSA-resected specimens.

1531 Validation & Multiple Use Case Applications of Whole Slide Imaging in a High-Volume Anatomic Pathology Global Reference Laboratory

Kirti Chadha[1], Kunjal Lila[1], Ashwini Patkar[2], Shaikhali Barodawala[1], Mukul Vij[3], Anuradha Murthy[2], Vikas Kavishwar[1], Pranav Desai[4], Krishna Prasad[1], Shital Munde[5], Tejal Shah[6], Metropolis Healthcare LTD at SSRM Pathology Clinic LTD[7], Krishna Detroja[8], Ameya Khadilkar[1], Ramrao Nilkanthe[2], Dhaval Doshi[1]

[1]Global Reference Laboratory, Metropolis Healthcare Ltd, Mumbai, Maharashtra, India, [2]Integrated Oncopathology Department, Metropolis Healthcare Ltd, Mumbai, Maharashtra, India, [3]Global Reference Laboratory, Metropolis Healthcare Ltd, Chennai, Tamil Nadu, India, [4]Desai Metropolis, Surat, Gujarat, India, [5]Global Reference Laboratory, Metropolis Healthcare Ltd, Nashik, Maharashtra, India, [6]Global Reference Laboratory, Metropolis Healthcare Ltd, Vadodara, Gujarat, India, [7]Kolkata Metropolis, Kolkata, West Bengal, India, [8]Sanjeevni Metropolis, Rajkot, Gujarat, India

Background: Whole slide imaging or Digital pathology as is commonly referred to represents a substrate from which glass slides can enter the digital domain. This technology was validated by Histopathologists & received accreditation. Thereafter multiple applications were analysed & executed that enhanced the traditional workflow in anatomic pathology practice. It encouraged sub-specialty expert consultation, Intradepartmental consensus, quality assurance and medical education.

Design: Intra-observer validation for WSI on Intellisite Philips Solution was carried out for 15 Consultant Pathologists as per CAP guidelines, minimum of 60 histopathology cases with 20 slides per additional modality per Consultant with a wash out period of 2-4 weeks. A total of 1494 case cohort was included for validation including H&E (3135 slide reads), special stains (240 slide reads) & IHC (1254 slide reads). Cytology & Frozen were excluded from the scope.

The pre-existing Glass Slide Library was digitised & students had access to glass & digitized version increasing the geographical scope. The institute launched Proficiency Testing for Anatomic Pathology using WSI. Intra-departmental consensus and expert sub-specialty access was available improving accuracy and reducing TAT. Faculty were able to attend Tumour Boards and Clinical Path meetings across 7 countries.

Results: Overall concordance between WSI and conventional glass slides: 95.52%. Major discordance was 0.27% and minor discordance was 4.22%. WSI was preferable over conventional glass slides in 7.90% & conventional glass slides were preferable over WSI in 3.75% of the cases.

A total of 450 slides were digitized for Medical Education. TAT for off-site secondary expert consultation reduced by 4 - 7 days and on site by 1-2 days. Accuracy improved by 12% by consulting a sub-specialty co-faculty. 60 Clinical path meetings were attended across specialties using WSI. Proficiency testing (EQAS) Cycle time of Anatomic Pathology reduced by 20 days.

Table 1: Case cohort of 1494 cases (4629 slide reads) - Intra-observer validation of 15 pathologists

| S.No	H & E	Special stain	IHC	Total	%Major Discordance	% Minor Discordance	%Concordance	%WSI better	%Glass slide better
A	62	20	20	102	0	1.96	98.03	8.8	3.9
B	61	20	20	101	0	4.95	95.04	5.9	4.9
C	65	20	20	105	0	2.86	97.14	6.6	4.7
D	60	20	22	102	0	4.9	95.09	5.8	3.9
E	60	20	22	102	0	3.92	96.07	4.9	3.9
F	60	20	23	103	0	4.85	94.14	4.8	2.9
G	65	20	20	105	0.95	3.81	95.24	8.6	4.77
H	60	20	20	100	1	7	93	8	4
I	60	20	0	80	1.25	3.75	95	11.3	6.25
J	63	20	21	104	0	6.73	93.27	7.69	3.85
K	60	20	0	80	1.25	6.25	92.5	8.75	6.25
L	100	0	0	100	0	2	98	7	5
M	69	20	21	110	0	5	95	15	3.64
N	100	0	0	100	0	5	95	9	6
O	100	0	0	100	0	0	100	0	0
	1045	**240**	**209**	**1494**	**0.27**	**4.22**	**95.52**	**7.9**	**3.75**

Figure 1 - 1531
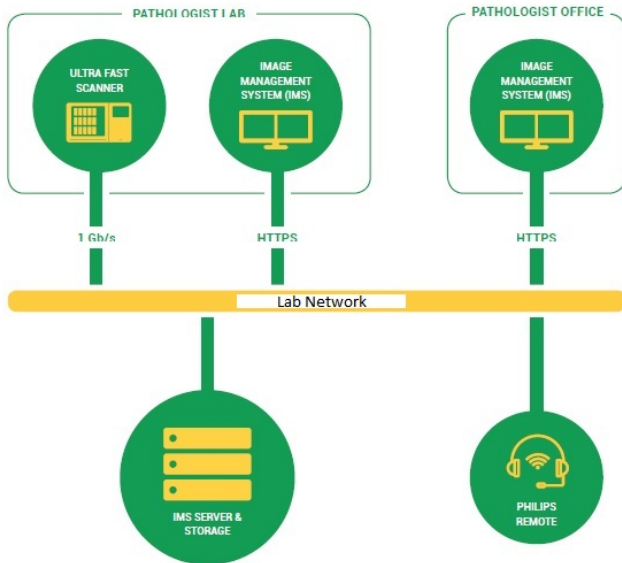
Figure 2 - 1531

Conclusions: The results of the validation showed that interpretation of whole slide images was concordant to the microscope and acceptable. WSI revolutionized histopathology education for pathologists in training with access to curated and organized cases. Intra-departmental meetings & sub-specialty expert access improved accuracy and reduced turn-around time. Proficiency testing process evolved & shortened.

1532 Automated Collection of RVU Data for Subspecialized Pathology Practice

Jennifer Chapman[1], Anton Morenko[2], Merce Jorda[3], Monica Garcia-Buitrago[4]
[1]University of Miami, Miller School of Medicine, North Miami, FL, [2]University of Miami, Miami, FL, [3]University of Miami Miller School of Medicine, Miami, FL, [4]University of Miami Miller School of Medicine/Jackson Health System, Miami, FL

Disclosures: Jennifer Chapman: None; Merce Jorda: None; Monica Garcia-Buitrago: None

Background: Due to restructuring of healthcare reimbursement, Pathologist's salary has increasingly shifted from being fixed to having a fixed component, plus added variable components provided as incentives. Variable components are used by administrators to incentivize physician behavior or productivity, the latter requiring a reliable system for measuring work. Workload has traditionally been measured by relative value units (RVU), a system that assigns points to each billable clinical activity according to reimbursement amount by current procedural terminology (CPT) code. This system is imperfect and widely criticized. While other methods for measuring workload have been developed, they are laborious to calculate in real time and have not been widely adopted.

Design: In subspecialty based practice, equal work allocation has been difficult to achieve. One reason is the lack real time visibility of number of cases and RVU assigned to each subspecialty and Pathologist. As our Department shifted from general to subspecialty practice, we sought to develop a method to electronically track the number of cases and RVUs assigned to each subspecialty and Pathologist.

Results: Within our Laboratory Information System (LIS) system database there is a subspecialty field (Sunquest CopathPlus v 6). Use of this field allowed us to create case level subspecialty assignments according to the part types submitted, and identify individual outreach consultation cases (Figure 1). We have developed a method to use the fee schedule functionality within the LIS, which would normally store pricing information, to instead define RVU values according to each CPT code. Using a query from any database browser, the criteria looking specifically at the RVU fee schedule is set along with Pathologist and subspecialty assignment. Data reports show each part type as a CPT row with a one to one association with the RVU, Subspecialty, Pathologist, and a calculation of CPT units multiplied by RVU value. Data is extracted into Excel, where data manipulation is applied (Figure 2).

Figure 1 - 1532
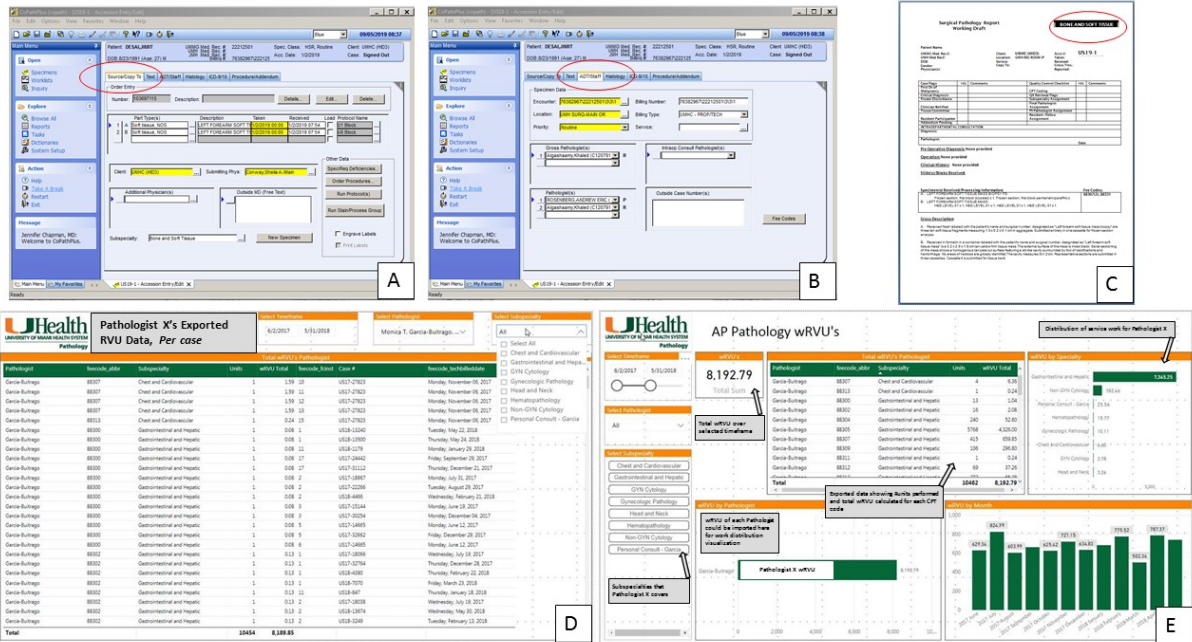
Conclusions: We have created visibility of RVU data per pathologist and per subspecialty in real time by setting frequent uploads of data from our LIS into tools such as Power BI dashboards, allowing Pathologists and Departmental leadership to understand and monitor RVUs by viewing data from individual desktops.

1533 Deep Learning-Based Pathomic Fusion for Glioma Outcome Prediction

Richard Chen[1], Ming Lu[1], Jingwen Wang[2], Faisal Mahmood[1]
[1]Brigham and Women's Hospital, Harvard Medical School, Boston, MA, [2]Brigham and Women's Hospital, Boston, MA

Disclosures: Richard Chen: None; Ming Lu: None; Jingwen Wang: None; Faisal Mahmood: None

Background: The current standard-of-care for cancer diagnosis and prognosis is the subjective analysis of histopathology slides and molecular profiles. Subjective diagnosis has shown to have a large interobserver and intraobserver variability among pathologists. Recently, there has been an increased interest in using deep learning for automated and objective grading of histology slides. However, most deep learning-based objective outcome prediction paradigms do not make use of the wealth of multimodal diagnostic data in an intuitive manner and usually rely on histology ROIs. Existing work on combining histology and molecular profiles relies on concatenation which does not make use of the most relevant features across the two modalities.

Design: In this work, we propose a unique approach for integrating histopathology and genomic data for glioma, in which we use deep learning to learn a coordinated representation of the two modalities using tensor fusion. We train two independent deep neural networks on histopathology image and genomic data respectively. Due to the lack of training samples, our network architectures for processing genomic and histopathology data are the Self Normalizing Network (SNN) and VGG16 respectively. Afterwards, we implement a multimodal deep neural network that fuses both histopathology image and genomic data. Fusion was performed by taking the outer product of the last hidden layers from the two unimodal networks. Each deep neural network was trained using the Cox partial likelihood loss, which has been demonstrated to be the same as maximizing the Concordance Index. We implemented all of our data and preprocessing code in PyTorch.

Results: Our approach is validated on glioma data from the Cancer Genome Atlas (TCGA) Pan Cancer Atlas dataset, which contains 769 samples of paired whole-slide image, genotyping and transcriptome data with ground truth survival time labels for glioma. Our results demonstrate that multimodal learning with our proposed architecture for integrating histopathology image and genomic data is able to outperform traditional approaches, with an increase in C-Index performance of 4.2%, 4.3%, and 34.2% over concatenative fusion, histopathology data, and genomic data.

Model	C-Index
Tensor Fusion	**0.7496**
Concatenative Fusion	0.7148
Histopathology	0.7177
Omic	0.4930

Figure 1 - 1533
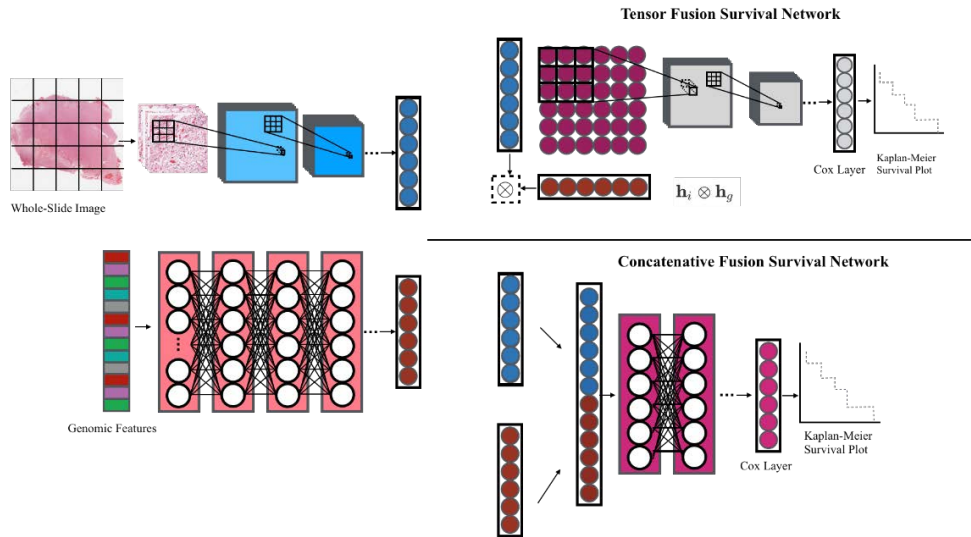
Conclusions: We propose a learning strategy for fusing genomic and histology data, which we validate on the TCGA-glioma dataset. Future work would extend this approach for other cancer types in TCGA.

1534 An Objective Deep Learning-Based Paradigm for Establishing Correspondences between Genomic and Histopathology Data

Richard Chen[1], Faisal Mahmood[1]
[1]Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Disclosures: Richard Chen: None; Faisal Mahmood: None

Background: Cancer cells exhibit enormous phenotypic and genotypic intratumoral heterogeneity in histopathology tissue and molecular profiles respectively. The spatial organization of cellular density and microvascular patterns in histopathology diagnostic slides has the potential to elucidate correspondences with molecular biomarkers, which can be targeted at the cellular level. However, such correlations between histopathological image and molecular profile features have been performed using either subjective techniques or non-learning based approaches.

Design: In this work, we design a novel deep learning approach that learns to translate molecular profiles to morphological patterns using a generative adversarial network (GAN), which we call Cycle-Aware Conditional GAN (CACGAN), illustrated in Figure 1. In this work, our generator is a mapping function that seeks to synthesize histopathology image data conditioned on the molecular profile data, and the discriminator is a convolutional neural network aims to distinguish between real and synthesized pairs of genomic and histopathology samples. We use three loss functions: an adversarial loss, a feature-matching loss and a cycle-consistency loss, which are used together to learn a mapping function that is able to synthesize realistic histopathology image data that correspond with molecular profiles.

Results: Our approach is trained and validated on glioma data from the Cancer Genome Atlas (TCGA) Pan Cancer Atlas dataset, which contains 769 samples of paired whole-slide image, genotyping and transcriptome data for glioma, with annotation for glioma histologic subtypes. Figure 2 shows synthesized histopathology tissue image results for Astrocytoma and Oligodendroglioma. The top row shows histopathology image results generated from existing methods, with the second-to-last result being CACGAN. In the bottom row, we test whether CACGAN would be able to synthesize histopathology tissue patterns corresponding to Oligodendroglioma and Astrocytoma from genomic vectors containing 1DH mutation + 1p19q codeletion and no 1DH mutation + 1p19q codeletion respectively. In an interesting result, we were able to generate varying amounts of nuclei in the histopathology image by adjusting the mutation and codeletion status of these features.
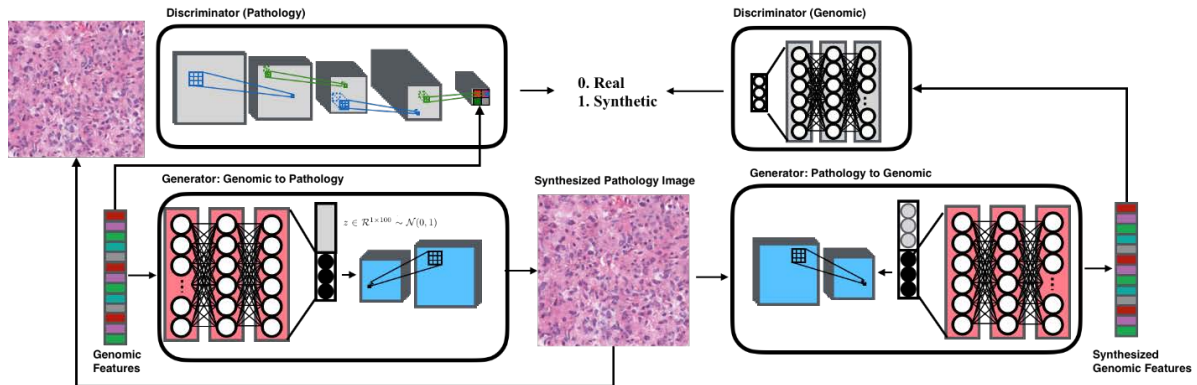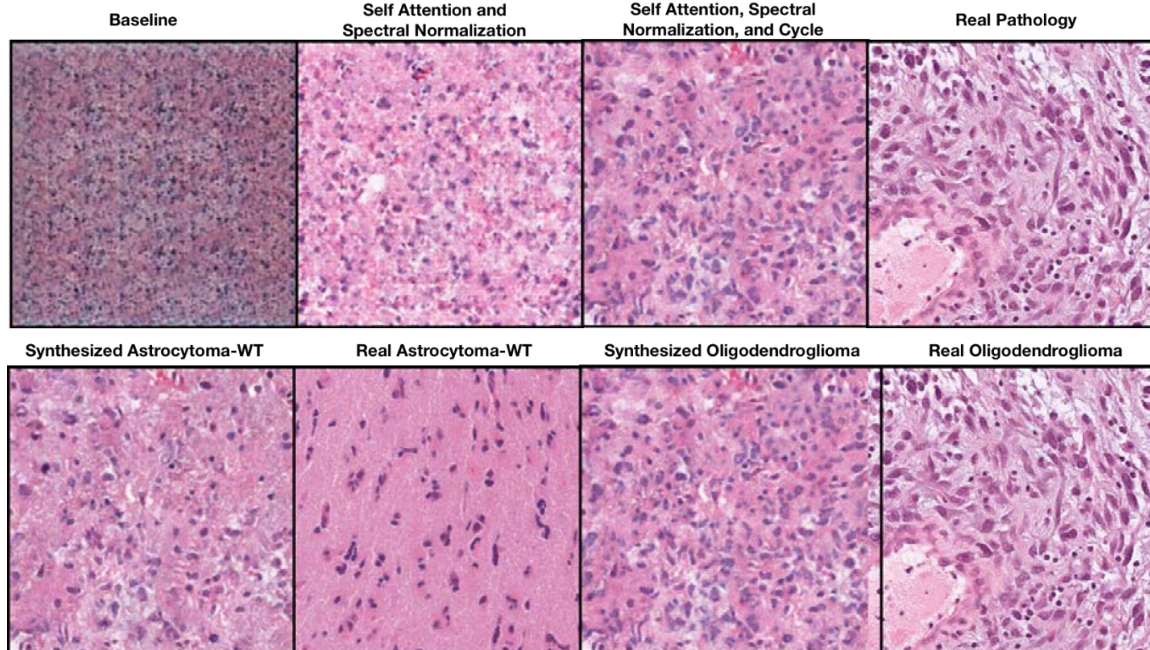
Figure 1 - 1534

Figure 2 - 1534

Conclusions: We propose a novel deep learning approach for developing correspondences between morphological patterns and molecular profiles. Future work would extend this approach to include other cancer subtypes.

1535 Unsupervised Pathology Report Classification Through Document and Word Embeddings

Jerome Cheng, University of Michigan, Ann Arbor, MI

Disclosures: Jerome Cheng: None

Background: In recent years, word embedding approaches have emerged as a versatile and powerful tool for encoding words with their contextual meaning in a numerical representation, commonly referred to as word vectors. One such method is Word2vec, developed by a team of researchers led by Tomas Mikolov at Google; it relies on a neural network to derive vectors (a series of numbers) for words in a large body of text. Heavily based on Word2vec, Doc2vec (developed by Quoc Le and Tomas Mikolov) is a similar method that encodes paragraphs as a series of numbers, in addition to the word vectors derived from each word making up the paragraphs. Applied to pathology reports, which are still commonly text based, paragraph vectors may be used to classify these reports through supervised and unsupervised methods. Supervised methods involve labelling each report as belonging to one class (e.g. benign vs. malignant), and training a machine learning model with them that can classify other reports into the appropriate category. Doc2vec may also be used in an unsupervised manner on unlabeled pathology reports to find reports that are similar to a given report, which will be the focus of this study.

Design: Diagnosis sections of 520,546 Pathology reports were extracted from our laboratory information system database using a Structured Query Language (SQL) based query. A Doc2vec model was trained with the report data for 500 epochs (iterations) using a

vector size of 300, generating unique paragraph embeddings for each report. Using the model, the top 3 most similar cases were retrieved for 100 separate cases and the predicted results were evaluated by a Pathologist for relevance.

Results: 297 out of 300 (99%) of the predicted results were contextually relevant in terms of either the lesion type or specimen location; 281 out of 300 (93.67%) of the predicted results were contextually relevant in terms of both the lesion type and specimen location. As an example, Table 1 enumerates the top 3 most similar diagnoses predicted by the trained Doc2vec model for a selected case.

Table 1 Top-3 most similar diagnoses retrieved by the Doc2vec model for a selected case
Diagnosis: Skin, left chin, shave: Squamous cell carcinoma, invasive, well to moderately differentiated, extending to the deep margin. Most similar cases: 1. Skin of right chin, shave: Invasive well differentiated squamous cell carcinoma, extending to deep margin. 2. Skin of chin, shave: Invasive, moderately to well-differentiated squamous cell carcinoma, extending to all margins. 3. Skin of left chin, shave: Invasive, well-differentiated squamous cell carcinoma, extending to deep margin.

Conclusions: Document embedding methods like Doc2vec are effective at encoding contextual information from free-text documents such as pathology reports, which has applications in research, education, and retrieval of similar cases during case sign-out. A search feature based on word and document embeddings may be built as a stand-alone program or incorporated as a feature in a laboratory information system.

1536   Improving Pathology Case Search Accuracy Using Convolutional Neural Networks
Jerome Cheng, University of Michigan, Ann Arbor, MI

Disclosures: Jerome Cheng: None

Background: In most institutions, pathology reports are stored as unstructured free-text within the laboratory information system, often leading to challenges in data retrieval. Conventional case searches for specific diagnostic entities involve querying the database for keywords associated with the diagnosis of interest, and these queries can return results less specific than intended. In this study, a convolutional neural network was applied to reports extracted with the keywords "colon" and "carcinoma", with the aim of increasing search accuracy in identifying reports containing primary colonic adenocarcinomas.

Design: The diagnosis section of 1000 pathology reports with the terms "colon" and "carcinoma" were retrieved from our laboratory information system through an SQL (Structured Query Language) Query. Each of the reports were labeled by a pathologist as either positive or negative, where cases are considered positive if the case was a primary adenocarcinoma of the colon. Negative cases comprised adenocarcinoma from other sites, metastatic adenocarcinomas, benign conditions, rectal cancers, and other cases that do not fit in the primary colonic adenocarcinoma category. The 1000 cases were separated into 2 sets – 500 each for training and testing. The training set contained 371 positive cases, and the test set had 351 positive cases. A Convolutional Neural Network model built using Keras (a neural network library) was trained to identify positive cases, and the model was applied to the test set to predict the category for each case.

Results: The CNN model classified 344 out of 351 primary colonic adenocarcinoma cases, and 126 out of 149 negative cases correctly (see confusion matrix in Figure 1), achieving an accuracy of 94% and area under the ROC Curve (AUC) of 0.98 (Figure 2).
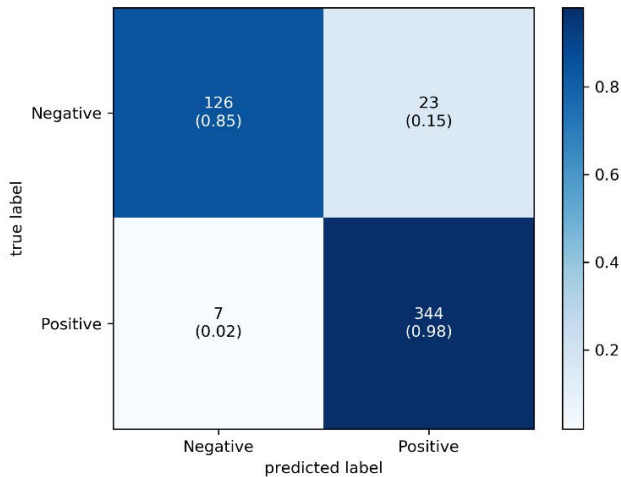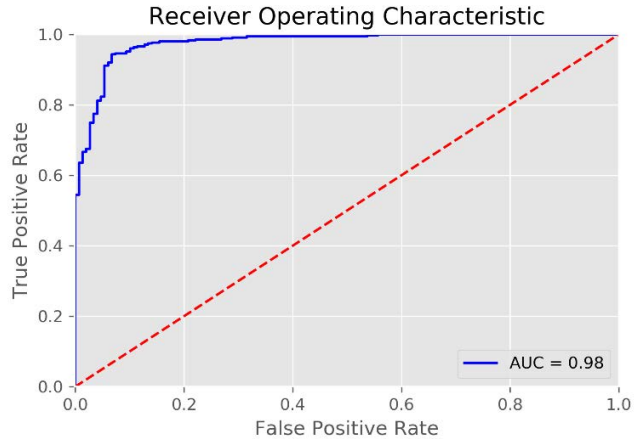
Figure 1 - 1536

Figure 2 - 1536

Conclusions: Trained CNN models by itself, or as an adjunct to keyword and pattern-based text extraction methods may be used to search for pathology cases of interest with high accuracy. Since manually labelling cases is a time-consuming endeavor, CNN assisted report identification is only recommended for large datasets, where the benefits may exceed the effort spent labelling a training set.

1537 High-Throughput WSI Scanning in the Histology Workflow of a Large Academic Lab

Thomas Chong[1], William Wallace[2]
[1]Los Angeles, CA, [2]Keck School of Medicine of University of Southern California, Los Angeles, CA

Disclosures: Thomas Chong: *Stock Ownership*, Danaher Corp.; William Wallace: *Advisory Board Member*, Leica Biosystems

Background: The advantages of whole slide image (WSI) scanning in the anatomic pathology workflow has been well-documented. Some facilities have already proceeded to all-digital workflows and have overcome the challenges with implementation of costly equipment. But for most medium to large facilities, the digital slide scanner role has been relegated to lower-volume workflows. Newer high-throughput whole slide scanners (80-100 slides/h at 40x power) is a promising step towards achieving an all-digital diagnostic workflow.

Design: Two Leica Biosystems GT450 scanners were stationed in the UCLA histology lab to receive a subset of non-time-critical cases. QA was performed on digital WSI's produced from GT450 using archival slides scanned at 40x. Three pathologists scored the quality of 180 total images from 1 unacceptable to 4 excellent, over three different sessions. The workflow was first evaluated to identify the time intervals for peak slide output and to assess the manual sort and collation times. A slide adapter was used to expedite the transfer of dried glass slides from the slide stainer/coverslipper slide racks to Leica slide racks. For each slide, we logged the scan time, region of interest size, and image file size, and manually recorded slide scan issues.

Results: Three pathologists assessed image quality of all test slides to be excellent and sufficient for diagnosis. Slide transfer time from cover-slipper racks to scanner racks using an adapter averaged 58s. Over 3 days, total scan time was 4h 40m, number of slides scanned 220, and avg scan time per slide 66s which includes rescan time. The types of errors encountered were barcode errors (16), image quality errors (18), and "no tissue" errors (3). 34 slides were rescanned, almost entirely due to spurious mounting medium from the automated coverslipper on the slide resulting in an expanded scan area.

Conclusions: In-line workflow trials of equipment is necessary to elucidate details and unforeseen issues in local implementation. In doing so, we identified pre-scan factors (e.g. mounting medium errors and possible slide label barcode quality) that can be managed, leading to more efficient scanning. Even including re-scanning due to mounting medium errors, scanning slides at 40x was faster than scan times that were previously possible at 20x. The slide throughput of the GT450 is effectively 2-3x that of the prior generation scanner (AT2) at 40x, allowing for fewer scanners for the same case volume, with no loss in image quality.

1538 Development and Online Release of "ImmunoGenius", a New Machine Learning Based Mobile Application for Immunohistochemistry Interpretation

Yosep Chong[1], Gyeongsin Park[2]
[1]Seoul, Korea, Republic of South Korea, [2]The Catholic University of Korea College of Medicine, Seoul, Korea, Republic of South Korea

Disclosures: Yosep Chong: None

Background: Immunohistochemistry (IHC) has been playing a great role in the pathologic diagnosis for the determination of the tumor origin by visualizing the protein expression. However, the IHC data is increasing exponentially and it is a huge challenge for the pathologists to correctly interpret IHC results according to the up-to-date knowledge. We designed this project to develop an expert supporting system that can quickly provide accurate IHC data and accelerate more efficient diagnosis process on the mobile platform.

Design: We developed a mobile application for iOS and Android using a probabilistic decision tree algorithm. Using over a dozen of the major textbook including the WHO classification of tumours series and open literature data, the IHC expression database for the most neoplasm in the textbook was built. The algorithm was trained with the real IHC profile data of 639 lymphomas and 634 tumors of unknown origin (TUO) and validated with 392 lymphomas and 382 TUOs to compare the presumption accuracy.

Results: The IHC expression data of over 2009 neoplasms and 584 antibodies was built. The major features include IHC expression data search by disease and antibody name and the generation of IHC profile table and diagnosis presumption with antibody test results. The mobile application has been released on the Apple Store and Google Playstore. It can be downloaded by the searching query "ImmunoGenius" or with the QR code. The presumption accuracy using training dataset was 94.7% for lymphomas and 78.5% for TUOs. The presumption accuracy using validation dataset was 95.7% for lymphomas and 78.0% for TUOs, which is not significantly different from that of the training dataset. The major cause of inaccurate presumption was due to atypical IHC profiles of a few cases, overlapping IHC profiles between differential diagnoses, and the absence of disease-specific markers in some neoplasms.

Figure 1 - 1538
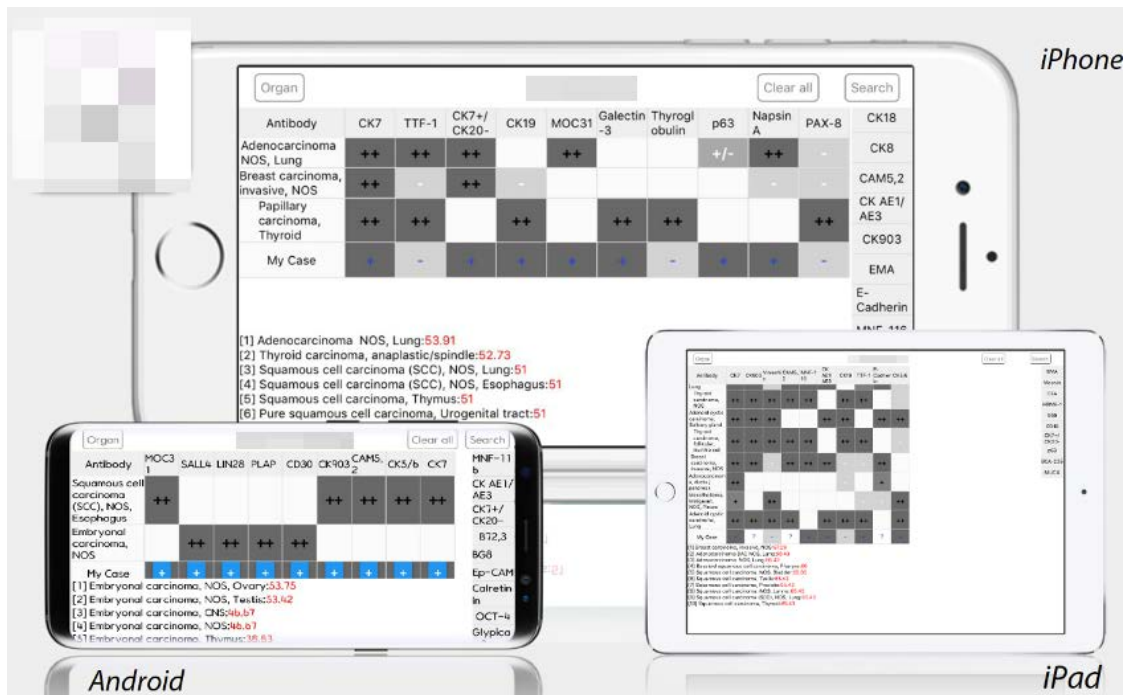
Conclusions: Better usage of IHC expression data in perspective of precision medicine can be achieved by the aids of computer expert systems. This application can be a good supportive tool for a more comprehensive and integrated interpretation of IHC results.

1539   Machine Learning as an Ancillary Tool in the Assessment of Shaved Margins for Breast Carcinoma Excision Specimens

Timothy D'Alfonso[1], David Ho[1], Matthew Hanna[1], Anne Grabenstetter[1], Dig Vijay Kumar Yarlagadda[1], Luke Geneslaw[2], Peter Ntiamoah[3], Lee Tan[1]
[1]Memorial Sloan Kettering Cancer Center, New York, NY, [2]Memorial Sloan Kettering Cancer Center, Brooklyn, NY, [3]New York, NY

Background: Breast conserving surgery for carcinoma includes excision of the primary lumpectomy specimen and separate oriented specimens taken from margins of the lumpectomy cavity. This "cavity shave" method is associated with lower rates of positive margins and fewer re-excisions. Pathologic assessment of these specimens, which are usually benign, can be time-consuming and require examination of multiple (15-40) H&E slides per case. With the increasing capabilities of digital slide scanning, computational pathology approaches could potentially improve the efficiency of this process by evaluating whole slide images (WSIs) of margins. We undertook a pilot study to determine the utility of machine learning as a screening tool for assessment of margin specimens.

Design: Lumpectomy specimens for invasive ductal carcinoma (IDC) and ductal carcinoma in situ (DCIS) for which all slides were digitally scanned were identified. Cases with positive and negative margins were randomly selected for analysis. A multi-class machine learning model was trained by WSIs to automatically segment carcinoma, benign epithelium, stroma, necrosis, adipose tissue, and background. The model utilized patches from multiple magnifications to predict tissue types. An image was classified as positive if >1 million pixels were segmented as carcinoma.

Results: 98 margin specimens were evaluated from 20 patients, consisting of 425 WSIs (mean: 4.3 slides/specimen; range: 1-10). 25 specimens (60 WSIs) contained carcinoma (18 DCIS, 7 IDC +/- DCIS). 73 specimens (365 WSIs) were benign. At the individual slide level, the model showed a sensitivity of 81% (49/60) and specificity of 85% (310/365) for identifying carcinoma. At the specimen level, sensitivity was 96% (24/25) and specificity was 60% (44/73). Figure 1 shows a margin specimen with DCIS (A, C) with cancer segmentation in red (B, D). Review of segmentation images revealed all 11 false negatives resulted from accurate segmentation, but of less than the predetermined pixel cutoff (Figure 2). The main sources of false positive classification were reactive changes (biopsy site, elastosis, cautery), fibrocystic changes, and lobular neoplasia.
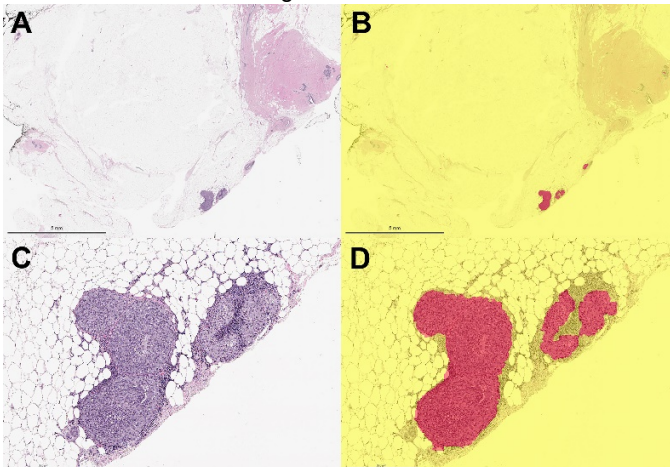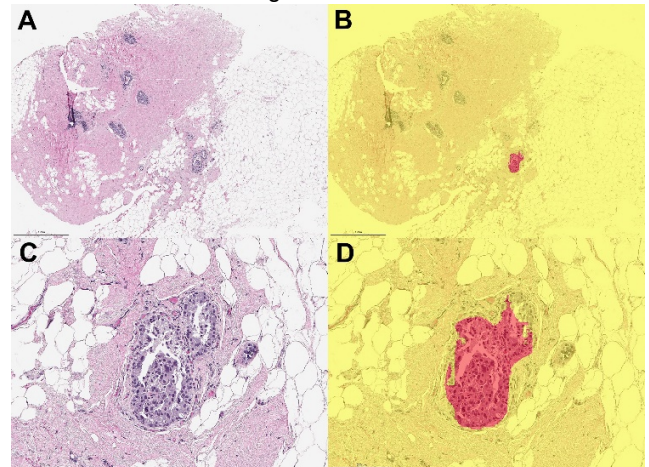
Figure 1 - 1539

Figure 2 - 1539

Conclusions: Our initial model showed good sensitivity and fair specificity for detecting carcinoma in WSIs of margins and represents a potential tool for increasing efficiency in their assessment. Further calibration of the machine learning algorithm to improve its accuracy via additional manual annotation and analysis of additional margins is ongoing.

1540   Development of Artificial Intelligence Algorithms to Quantify Nuclear Ki67 in Glandular Epithelial Cells in Barrett's Esophagus

Armando Del Portillo[1], Caitlin Hills[2], Elena Komissarova[2], Jorge Sepulveda[2], Julian Abrams[2], Antonia Sepulveda[3]
[1]New York, NY, [2]Columbia University Medical Center, New York, NY, [3]George Washington University, New York, NY

Disclosures: Armando Del Portillo: None; Caitlin Hills: None; Elena Komissarova: None; Jorge Sepulveda: None; Julian Abrams: None; Antonia Sepulveda: None

Background: Digital pathology (DP) has the potential to efficiently analyze many images with less subjectivity and bias compared to manual analysis. Barrett's esophagus (BE) is a pre-cancerous lesion, and the proliferation index of BE epithelium may serve as a biomarker of progression. To evaluate Ki67 proliferation index in the context of a BE clinical trial, we developed artificial intelligence (AI) algorithms in a DP platform.

Design: Forty-six formalin-fixed, paraffin embedded biopsy samples from 23 patients from two different institutions participating in a clinical trial for a new medical therapy for BE were studied. Sections were stained with hematoxylin, and co-immunostained with Ki67 (brown chromogen), and pan-cytokeratin (CK, red chromogen). Slides were scanned at 400x using Aperio AT-2 whole slide scanner (Leica). We manually annotated each image to include BE/gastric cardia-type epithelium, and eliminated squamous epithelium, oxyntic epithelium, and excess stroma. We developed AI algorithms using HALO-AI (v2.3.2089.30, Indica Labs) to classify tissues into CK-strong epithelium, CK-weak epithelium, stroma, and glass on the annotated images (Fig 1A-C). We then performed nuclear segmentation and determined thresholds for nuclear Ki67 positivity (nuclear counting algorithm) (Fig 1D). We used an iterative approach to evaluate false positive and false negative tissue classification and Ki67 quantification. For each image, representative glands were manually counted and compared to AI nuclear counting algorithm. We also compared manual BE area to AI BE area.

Results: We generated 11 classifier algorithms, 15 nuclear counting algorithms, and 26 unique combinations of these algorithms in order to analyze 46 stained sections. Nuclear counting algorithms were within 5% of manual count per image, and AI showed no significant difference overall compared to manual (Table 1). Classifier algorithms increased BE area by an average 55% per image (SE=19%), and significantly increased overall BE area (Table 1), requiring more precise annotations.

	Manual (mean, SE)	Optimal AI algorithm (mean, SE)	P value
BE area (mm^2)	3.92, 0.37	5.24, 0.46	0.028
Ki67 count	121.2, 9.6	120.9, 9.52	0.98

Figure 1 - 1540
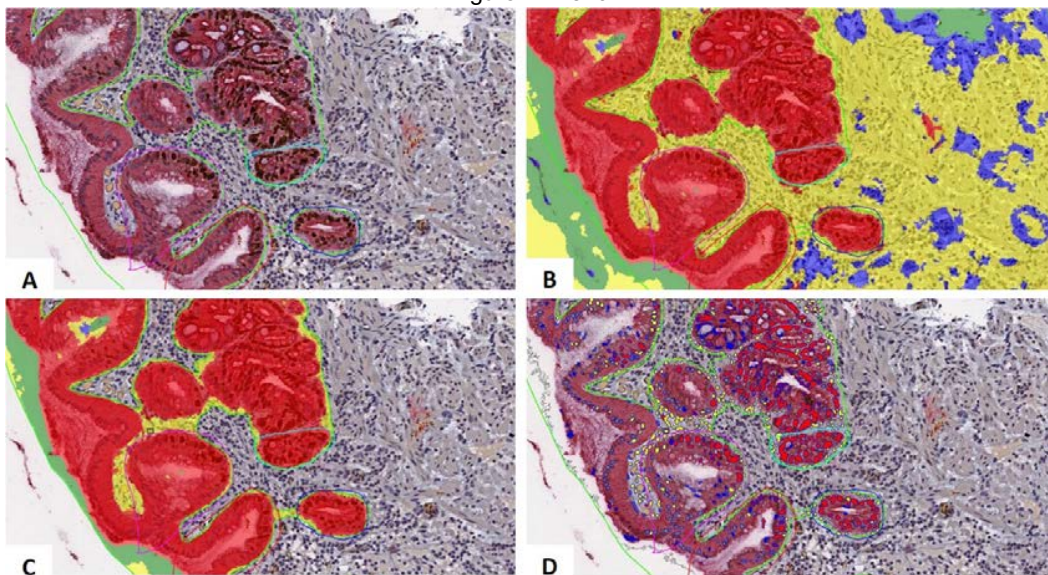
Conclusions: A combination of classification algorithms, nuclear segmentation and positivity algorithms, and precise annotations are required for accurate Ki67 quantification in BE. Tissue artifacts or variability in staining requires a labor intensive custom algorithm development per image in order to ensure accuracy using a chromogen method. Other methods with better dynamic range (e.g. fluorescence-based stains) may be less sensitive to some of these variables.

1541 The Enhancement of the Diagnostic Yield of Breast Core Needle Biopsies Using Optical Sectioning Microscopy

Sapna Desai[1], Jennifer Campbell[1], Ramapriya Vidhun[2], Richard Torres[3], Eben Olson[3], Michael Levene[4], Theresa Profeta[1], Paul Fiedler[1]

[1]Danbury Hospital, Danbury, CT, [2]Western Connecticut Health Network, Danbury Hospital, Danbury, CT, [3]Yale School of Medicine, New Haven, CT, [4]Applikate Technologies, Weston, CT

Disclosures: Sapna Desai: None; Jennifer Campbell: None; Ramapriya Vidhun: None; Richard Torres: *Stock Ownership*, Applikate Technologies, LLC; Eben Olson: *Stock Ownership*, Applikate Technologies; Michael Levene: *Stock Ownership*, Applikate Technologies; Theresa Profeta: None; Paul Fiedler: None

Background: Multiphoton microscopy is a 3-dimensional imaging technique in which two photons are used to excite fluorescence emission. This process allows for the enhanced visualization of cells, organs and tissues at a more in-depth level than single-photon microscopy, and results in improved color and nuclear clarity. Additionally, issues arising from the current process of embedding of tissues, such as tearing and folding during embedding as well as damage to the specimens, can be avoided. Published literature comparing the quality of images produced by multiphoton microscopy to single-photon microscopy is limited. Findings which determine that these images have improved diagnostic potential could be of benefit to clinicians. This study aims to demonstrate the benefits and expanded diagnostic potential of optical sectioning microscopy.

Design: Breast core needle biopsies were obtained through the Danbury Hospital Department of Pathology and the Yale University Department of Research from tissue left over after sufficient tissue for clinical diagnosis had already been collected and processed. Samples were selected by pathology tissue procurement personnel based on criteria of tissue type and cancer diagnosis, of which only 'tissue type' was recorded. Specimens were collected in cassettes labeled only with tissue type. A small portion of the sample was chemically processed in the lab to achieve optical clearing and appropriate fluorescent staining. Three dimensional scanning was performed using the multiphoton microscope in the laboratory. The image sampled was further processed using traditional embedding, sectioning, and staining. Immunohistochemical stains were performed on the cleared and traditionally processed samples for comparison.

Results: A comparison of the images of the multiphoton and traditionally processed breast core biopsies revealed that image quality is at the least comparable, if not superior using multiphoton processing. This was true for H&E stained slides as well as for the immunohistochemical stains that were performed on each core, including E-cadherin, GATA-3, CK AE1/3, HER-2, ER, PR and p63.

Figure 1 - 1541
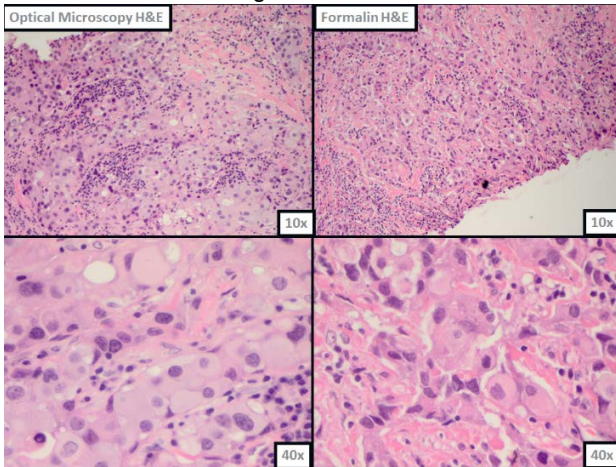
Figure 2 - 1541
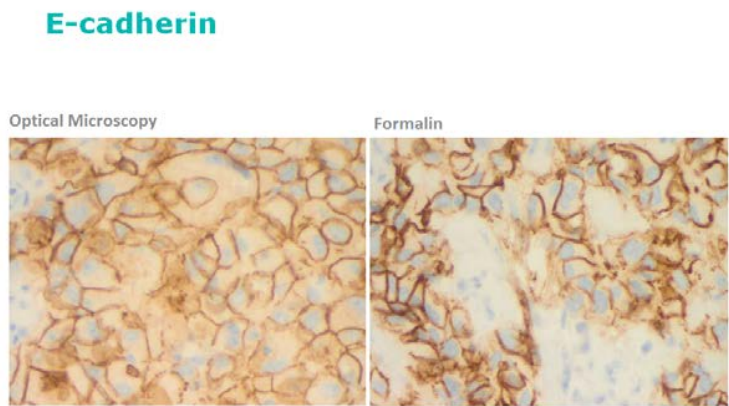
Conclusions: The results demonstrate that multiphoton microscopy is an extremely valuable advancement in the processing and diagnostic potential of breast core biopsies. While this abstract reviews the comparison of multiphoton microscopy to traditional processing, further evaluation will consist of comparing additional tissue types and immunohistochemical staining.

1542 Spectral Unmixing of Microscopic Slides with Annotations Using Multispectral Imaging and a Linear Unmixing Algorithm for Producing an Image of the Annotation and an Image of the Histologic Stain in One Scan

Thomas Flotte[1], Vivian Negron[2], Karla Kopp[1], Steven Hart[1]
[1]Mayo Clinic, Rochester, MN, [2]Mayo Clinic Rochester, Rochester, MN

Disclosures: Thomas Flotte: None; Vivian Negron: None; Karla Kopp: None; Steven Hart: None

Background: Machine learning in surgical pathology is increasing the demand for well annotated datasets for supervised learning. For decades pathologists as well as allied professionals including cytotechnologists have been annotating microscopic slides by marking the slides with ink on the coverslips. For cytology specimens, it may be necessary to retain these markings for regulatory purposes. These individually annotated slides represent substantial resources for machine learning. One approach is to scan the slides with the marks, clean the slides, and scan a second time. However, there it would be preferable to be able to scan the slides once and be able to separate the markings and the tissue stains.

Design: Formalin-fixed, paraffin-embedded sections were stained for hematoxylin, eosin, and hematoxylin and eosin (H&E). Blank slides with coverslips were marked with red, green, blue, and black markers. H&E slides were marked with all markers. Microscopic images were acquired from 440 to 720 nm with a Vectra Polaris (PerkinElmer/Akoya Biosciences) equipped with a liquid crystal tunable filter. Spectra of singly stained slides and individually marked blank slides were determined from these spectral image cubes. H&E sections with markings were spectrally unmixed using the spectra defined previously using the inForm software (PerkinElmer/Akoya Biosciences).

Results: The figure illustrates spectral unmixing. For H&E sections, the results for green marker show excellent separation of eosin, hematoxylin and the marker. The areas with blue marker show very small numbers errors in unmixing in the nuclei. The black marker showed some unmixing errors in the hematoxylin stain and the black marker. The red marker showed substantial unmixing errors in the eosin stain and red marker.

Figure 1 - 1542
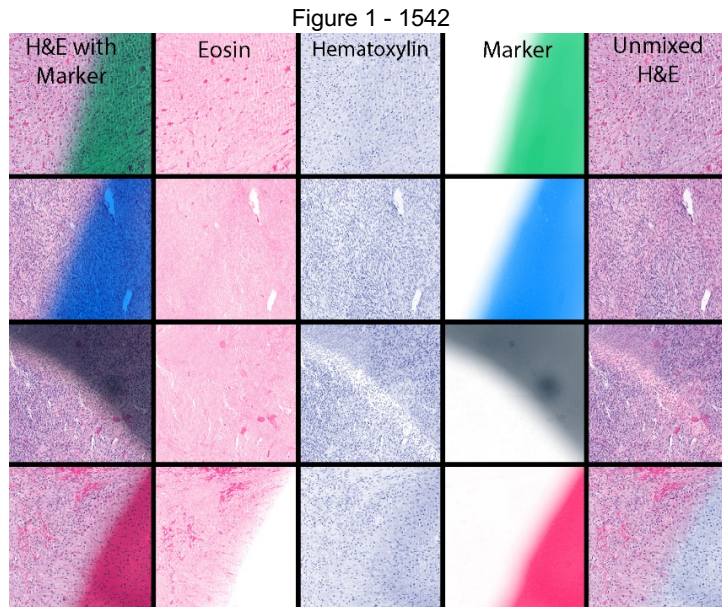
Conclusions: This study demonstrates that linear unmixing of multispectral images of microscopic slides with ink markings will produce images of both the markings and the stained sections with perfect registration from a single scan of the slides. One of the limitations of this approach relates to similarity of the spectra of the marker and the stain. For H&E sections, the technique could not separate the red marker from the eosin stain. However, the green marker unmixing was nearly perfect, the blue marker had minor errors and the black marker had a small number of errors. Thus, by using markers with colors that do not overlap substantially with the colors of the stain, microscopic slides with ink annotations can be scanned once t

1543 Validation of Digital Pathology for Secondary Diagnosis in Large Consultative Pathology Practice

Thomas Flotte[1], Vera Suman[1], William Harmsen[1], Mark Norman[1], Charlene Brown[1], Taofic Mounajjed[1]
[1]Mayo Clinic, Rochester, MN

Background: Although studies have shown whole slide imaging (WSI) non-inferiority to microscopy for primary diagnosis in surgical pathology, large studies evaluating WSI for secondary consultation are limited. We aim to compare WSI to microscopy in subspecialized secondary consultation practice.

Design: As part of a large multi-specialty validation study of digital pathology in consultation practice, 672 consultative cases were reviewed by 38 pathologists specializing in GYN, bone and soft tissue, GI/hepatobiliary, lymph node, pulmonary, and Dermatopathology (89, 98, 91, 93, 101, and 200 cases, respectively). They included 312 consecutive cases and 360 cases from targeted categories selected to capture case types of uncommon histology or poor inter-evaluator agreement. All initially received slides (n=6361) were scanned using Aperio AT Turbo scanner (×40 power). Each case was reviewed by 2 pathologists; each pathologist reviewed cases twice (glass and digital) with identical resources (stains/opinions) available for either modality. Pathologists were randomized to which modality they used first. A 2-week washout period was required between glass and digital reviews. Digital review on medical-grade monitor was performed using Aperio eSlide Manager. Diagnoses were evaluated for disagreement; disagreements that would significantly alter treatment or prognosis were considered major.

Results: The table summarizes the study results. The major disagreement rate was equivalent for the pathologists who saw the glass slides first vs. the pathologists who saw the digital images first (Path1-Glass vs. Path2-Digital) and two pathologists seeing the glass slides (disagreement on glass) and slightly better than two pathologists seeing the digital images (digital disagreements). Access to additional studies was needed in many cases; pathologists evaluating slides first were more likely to order additional slides (36%) compared to pathologists evaluating digital slides first.

Comparison between glass and digital review of consults								
Cohort	Number of slides sent ,median (IQR)	Number of additional H&Es requested by Path#1/Glass	Number of additional IHC requested by Path#1/Glass	Number of additional H&Es requested by Path#2/Digital	Number of additional IHC requested by Path#2/Digital	Number of major disagrees (Glass)	Number of major disagrees (Digital)	Number of major disagrees Path1-Glass vs. Path2-Digital
GYN	7 (3-15)	22 (24.7%)	22 (24.7%)	1 (1.1%)	8 (9.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
BST	8 (3-13)	38 (38.4%)	38 (38.4%)	12/96 (12.5%)	22/96 (22.9%)	0 (0.0%)	1 (1.0%)	1 (1.0%)
GI Liver	5 (3-9)	50 (55.0%)	36 (39.6%)	8/88 (9.1%)	12/88 (13.6%)	4 (4.4%)	7 (7.7%)	4 (4.4%)
Lymph Node	13 (7-20)	56 (60.2%)	61 (65.6%)	15 (16.1%)	30 (32.3%)	3 (3.2%)	7/92 (7.6%)	3 (3.2%)
Pulmonary	9 (5-17)	42 (41.6%)	45 (44.6%)	10/96 (10.4%)	28/96 (29.2%)	1 (1.0%)	4 (4.0%)	2 (2.0%)
DermPath	3 (2-6)	46 (23.0%)	41 (20.5%)	39 (19.5%)	55 (27.5%)	3/199 (1.5%)	4 (2.0%)	4 (2.0%)

Conclusions: This study confirms that whole slide scanning can be utilized to provide the same level of expert diagnostic interpretations as glass slides review for a variety of complex second opinion consultation cases with a diverse group of pathologists. However, the pathologists thought that they needed additional studies in a significant number of the cases to be confident that they were providing equivalent consultative opinions that may have implications for the workflow for a digital pathology consultative practice.

1544 The Impact of a Digital Solution on Tumor Board Preparation Time for Pathology Residents

Donna Fowler[1], Lincoln Sheets[2], Matthew Prime[3], Chaohui Guo[4], Athanasios Siadimas[4], Richard Hammer[1]
[1]University of Missouri, Columbia, MO, [2]University of Missouri, School of Medicine, Columbia, MO, [3]Roche, Basel, Switzerland, [4]Roche, Basel, Basel-Stadt, Switzerland

Background: A multidisciplinary tumor board (MTB) provides an interdisciplinary approach for decision-making in cancer care. Efficient preparation of MTB is critical for time-to-treatment and patient outcome. Pathological examination has been the gold standard for diagnosis in cancer and its role has been included the elucidation of etiology, pathogenesis, and prognostication. (Leong and Zhuang, 2011).

Pathologists contribute significant amount of time in MTB case preparation and play critical roles in optimal cancer treatment. It is largely unknown whether and how digital tumor board solutions facilitate pathologists in case preparation for MTBs.

Design: A prospective IRB approved cohort study was undertaken to evaluate pathology residents (PRs) time preparation for Otolaryngology (ENT) MTBs before and after the implementation of the NAVIFY® Tumor Board (NTB) solution at University of Missouri Health Care (MU). Data was collected using a digital time-tracking application. NAVIFY® integrated version is a cloud-based workflow product that integrates with the hospital EMR and displays relevant aggregated data as a holistic patient dashboard. Student's t-test was performed in cases where data met the assumption of normality, and Mann-Whitney test otherwise.

Results: Time preparation for 36 MTBs (408 cases) were evaluated: 11 (115 cases) in pre-NAVIFY® and 25 (293 cases) in post-NAVIFY®. Overall residents spent 115 hours (6905 mins) in preparing all the cases (Table 1). Data showed the average meeting preparation time per-case significantly decreased from 20.4 (SD = 9.4) mins to 15.6 (SD = 6.5) mins, representing a 24% reduction (p-value = 0.04). A reduction in variance (SD & IQR) for time preparation was also observed (Table 1).

Table 1. Pathology Resident Preparation Time of ENT MTBs Pre- & Post-NTB Implementation

ENT Tumor board	PRE-NAVIFY®	POST-NAVIFY®
Number of meetings	11	25
Number of patient cases	115	293
Total time (mins)	2,343	4,562
Mean (SD) in mins	20.4 (9.44)	15.6 (6.45)
Median (IQR) in mins	20.8 (9.85)	15.3 (7.62)
Minimum (mins)	0.5	7
Q1 (mins)	16	11
Q3 (mins)	26	19
Maximum (mins)	35	34

Conclusions: Introduction of the NTB reduced time spent by residents preparing for ENT MTBs. The observed reduction in variance for time preparation suggests that the NTB solution standardized the process for MTB preparation. Improved efficiency of case preparation and standardized workflow can enable timely discussion of patient cases, improve patient outcome and experience, and improve efficiency that may lead to cost reduction. Future investigation is needed to further examine the impact in different types of hospitals (e.g., community-based) or in a multi-site setting study.

1545   Development and Evaluation of Automated 3D Scoring System of Fluorescence In Situ Hybridization (FISH) on Formalin-Fixed Paraffin-Embedded (FFPE) Tissues Using a Confocal Whole Slide Image Scanner

Ziv Frankenstein[1], Kareem Ibrahim[1], Umut Aypar[1], Ruth Aryeequaye[1], Mamta Rao[2], Ahmet Dogan[1], Meera Hameed[1], Yanming Zhang[1], Yukako Yagi[1]
[1]Memorial Sloan Kettering Cancer Center, New York, NY, [2]Memorial Sloan Kettering Cancer Center, Hackensack, NJ

Disclosures: Ziv Frankenstein: None; Kareem Ibrahim: None; Umut Aypar: None; Umut Aypar: None; Ruth Aryeequaye: None; Mamta Rao: None; Ahmet Dogan: *Consultant*, Roche, Corvus Pharmaceuticals, Seattle Genetics, Oncology Specialty Group, Pharmacyclics, Celgene, Novartis, Takeda; *Primary Investigator*, Roche/Genentech; Meera Hameed: None; Yanming Zhang: None; Yukako Yagi: None

Background: The standard manual scoring of fluorescence in situ hybridization (FISH) analysis of formalin-fixed paraffin-embedded (FFPE) tissues is labor-intensive, time-consuming and subjective. Confocal imaging, which eliminates out-of-focus noise, offers higher resolution and more spatial information than conventional widefield imaging. The purpose of this study was to establish an automated 3D FISH scoring method.

Design: Ten de-identified archival FFPE blocks of malignant lymphoma and Ewing's sarcoma with previous FISH tests using MYC, BCL2, BCL6 and EWSR1 break-apart probes were included in the current study. For each slide, several regions of interest (ROIs) within tumor areas were selected for confocal scanning according to the corresponding H&E and IHC slides. FISH slides were digitized by a Pannoramic Confocal WSI scanner (3DHistech) with a 40× water immersion objective (0.16 um/pixel). We scanned seven layers at 0.6 μm intervals. Images were viewed, and ROIs were defined. Image analysis for 3D FISH scoring was performed with SHIMARIS PAFQ V1.0 (in-house application) that employs 3D calculations for individual cell nuclei segmentation, spot detection and distribution of break-apart probe signal patterns, including standard break-apart, and variant patterns due to truncation, and deletion, etc (Fig. 1). The accuracy of the analysis was compared with clinical manual scoring and with our previous study using a commercially available software (semi-automated).

Results: FISH slides provided high quality data for 3D analysis of spot signals in each nuclei. Using our automated 3D system, individual cell nuclei segmentation, spot detection and distribution of probes were successfully performed. The average numbers of segmented individual nuclei are 343.6, 261.1 and 100, using the automated system, the semi-automated method by the commercially available

software and the manual procedure, respectively. Using a cut-off of 10%, the percentages of abnormal signal patterns correlated well with the clinical results using manual scoring (Fig. 2).
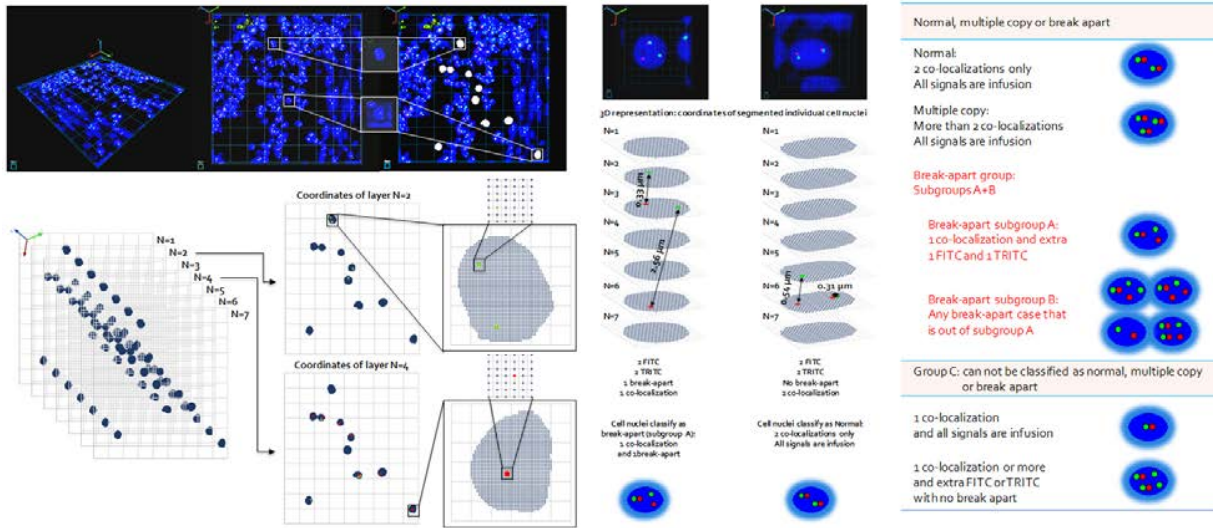
Figure 1 - 1545

Figure 2 - 1545

	Number of cell nuclei (% from total)*									
Case number:	1	2	3	4	5	6	7	8	9	10
Break-apart probe:	BCL6	BCL2	EWSR1	MYC	BCL2	BCL6	BCL6	MYC	MYC	MYC
Automated algorithm	Negative (1.3)	Negative (8.4)	Positive (81.6)	Negative (2.7)	Positive (60.2)	Negative (7.5)	Positive (74.3)	Positive (71.3)	Negative (3.6)	Negative (7.3)
Image analysis software	Negative (1.2)	Negative (5.3)	Positive (88.4)	Negative (7.5)	Positive (45.0)	Negative (3.3)	Positive (84.7)	Positive (52.7)	Negative (5.5)	Negative (3.1)
Manual procedure	Negative (-)	Negative (-)	Positive (88.0)	Negative (-)	Positive (56.0)	Negative (-)	Positive (68.0)	Positive (74.0)	Negative (-)	Negative (-)

*Percentages are of break apart signal patterns out of normal, multiple copy and break apart patterns (see Fig. 1)

Conclusions: We have established an automated method of 3D FISH scoring for FFPE tissue using confocal scanning. This method is more efficient, accurate and precise than the current standard and commercially available image analysis software. It enables the automated counting of more nuclei, precisely detecting more abnormal signal variations in nuclei and analyzing gigabyte multi-layer stacking imaging data of FFPE tissue samples.

1546 Convolutional Neural Networks Can Highly Accurately Identify Fungi in GMS Stained Sections: Computational Screening for Workflow Improvement?

Matthew Gayhart[1], Patricija Zot[2], Steven Smith[1]
[1]Virginia Commonwealth University School of Medicine, Richmond, VA, [2]Virginia Commonwealth University Health System, Richmond, VA

Disclosures: Matthew Gayhart: None; Patricija Zot: None; Steven Smith: *Consultant*, Elsevier Publishing/Amirsys

Background: At our institution, Gomori Methenamine-Silver (GMS) staining is frequently ordered on routine specimens to highlight the presence of fungal forms. However, a known issue with GMS is that artifacts may mimic fungal elements, contributing to time consuming evaluation. We propose that a computer-aided diagnostic tool that can rapidly screen for true fungi vs artifact could improve a busy surgical pathologist's workflow. To this end, we developed an image classification model to distinguish images of GMS-stained slides positive for fungal elements from images of GMS stained slides with artifact present only.

Design: Whole slide imaging at 40x magnification was performed on 101 archival cases for which GMS was ordered to asses for fungal elements (32 positive, 69 negative). The cases evaluated included 16 dermatologic, 9 gastrointestinal, 17 pulmonary, 3 cardiovascular, 26 ENT, 2 brain, 3 eye, 12 lymph node, 6 bone, 4 gynecological, 2 soft tissue, and 1 renal. From the whole slide imaging, 1013 total images (509 positive for fungal elements and 504 negative), each 154x154 pixels, were segmented and separated into a 811 training image subset and 202 testing image subset. The images were then used to create an image classification model using the ResNet-50 convolutional neural network architecture. Additionally, 28 images of GMS staining (14 positive and 14 negative for fungal elements) from outside institutions were obtained and tested using the image classifier to assess for generalizability.

Results: Of the blinded 202 image testing subset, 91 images were positive for fungal elements and 111 were negative. The GMS image classification model correctly identified 100% of the 91 positive images and 108 of 111 of the negative images (97.3%) for a sensitivity of 100% and specificity of 97.3%. The 28 outside images were then tested with GMS image classification model. Of outside positive cases, 13 of 14 (92.8%) were correctly classified, and 12 of 14 (85.7%) outside negative cases were correctly classified.

Conclusions: The image classification model is highly sensitive and specific for identifying fungi in routine surgical specimens from a variety of tissue types. Of note, no false negatives were identified. This opens the possibility of development of using a computer aided diagnostic tool to screen GMS stained images for fungal elements to improve the surgical pathologist's workflow by offloading time consuming tasks. Testing in an additional, independent, consecutive cohort is underway.

1547 "Virtual" Biomarkers? Convolutional Neural Networks Can Accurately Distinguish Chromophobe from Clear Cell Renal Cell Carcinoma

Matthew Gayhart[1], Anne Prater[1], Steven Smith[1]
[1]Virginia Commonwealth University School of Medicine, Richmond, VA

Disclosures: Matthew Gayhart: None; Anne Prater: None; Steven Smith: *Consultant*, Elsevier Publishing/Amirsys

Background: Chromophobe and clear cell renal cell carcinoma are distinct entities sharing morphologic overlap but strikingly different prognoses. Often, to distinguish between these two types of carcinoma, immunohistochemical (IHC) biomarkers are performed, adding time and cost to patient workup. We propose that an image classifier able to objectively differentiate between these two entities without IHC could decrease cost to the health system and improve turnaround time.

Design: Whole slide imaging at 40x magnification was performed on H&E slides from 56 archival renal tumor resection cases for which the final diagnosis was chromophobe (27 cases) or clear cell (29 cases) renal cell carcinoma. From the whole slide imaging, 1000 total images (500 chromophobe and 500 clear cell), each 154x154 pixels were segmented and separated into an 800 image training subset and a 200 image testing subset. The images were then used to create an image classification model using the ResNet-50 convolutional neural network architecture. Additionally, 46 images of renal cell carcinoma (23 chromophobe and 23 clear cell) were obtained from outside institutions and tested using the image classification model to assess for the generalizability of our image classification model.

Results: Of the 200 image (97 chromophobe, 103 clear cell) testing subset, all 97 of the chromophobe images (100%) and 102 of 103 (99.0%) of the clear cell images were correctly classified. The 46 outside images were then tested in the image classification model. Overall 22 of 23 outside chromophobe cases (95.7%) and 23 of 23 outside clear cell cases (100%) were correctly classified.

Conclusions: The image classification model was able to very accurately distinguish between chromophobe and clear cell renal cell carcinoma on H&E without the need to order IHC a diagnostic biomarker. While this is only a pilot study, these results highlight the potential of how machine learning based analytic techniques can improve how we practice pathology. We are presently evaluating the performance of this model in a consecutive cohort of clear cell and chromophobe carcinomas to simulate implementation of this strategy and assess its performance.

1548 Determination of Tumor Mutation Burden Using a 500+ Gene Pan-Cancer Comprehensive NGS Panel with Automated Variant Calling

Kelly Gerding[1], Kenneth Valkenburg[1], Christina Oliveras[1], James White[1], Leila Ettehadieh[1], Gustavo Cerqueira[2], Christopher Gault[1], James Hernandez[1], Eric Kong[1], Samuel Angiuoli[1], John Simmons[1], Isabell Loftin[1], Abigail McElhinny[1]
[1]Personal Genome Diagnostics (PGDx), Baltimore, MD, [2]Personal Genome Diagnostics (PGDx), Ellicott City, MD

Disclosures: Christina Oliveras: None; Leila Ettehadieh: None; Gustavo Cerqueira: None; Christopher Gault: None; Eric Kong: *Employee*, Personal Genome Diagnostics; Isabell Loftin: *Employee*, PGDx

Background: Among clinically relevant molecular biomarkers in oncology, tumor mutation burden (TMB) has recently emerged as a composite genomic metric, with potential to predict response to immune checkpoint inhibitors. However, measuring TMB is complex, technically demanding, and requires next generation sequencing (NGS) based approaches. Here, we present data from the analytical

validation of a 500+ gene, pan-cancer panel designed to report TMB, as well as single nucleotide variants (SNVs), insertion-deletions (indels), amplifications, translocations, and microsatellite instability (MSI).

Design: 118 FFPE tumor tissue samples (8 tumor types) were analyzed for TMB, utilizing our gene panel. Accuracy of TMB measured from the targeted panel was compared to whole-exome sequencing (WES) derived TMB. Additionally, >500 FFPE pan-cancer (35 tumor types) tumor tissue samples, were profiled for genomic alterations and concordance assessed via orthogonal methods (IHC, FISH, PCR, or NGS) and calculated as overall percent agreement (OPA). A proprietary, machine-learning bioinformatics pipeline was utilized for automated variant calling. Precision and repeatability of assay performance was assessed across multiple operators, instruments and days.

Results: TMB calls displayed a high level of concordance to WES-derived TMB (Pearson correlation, p=0.903) across a range of TMB scores (0.2-89.7 muts/Mbp), with an overall median of 4.5 muts/Mbp and a mean of 8.3 muts/Mbp. The lower limit of TMB measurement for the targeted panel was determined to be 1.9 muts/Mbp. Detection of sequence mutations (SNVs/indels) demonstrated an OPA of >99.9% compared to orthogonal NGS assays (n=112). Amplification detection had >93.0% OPA with FISH and IHC assays (n=176). Rearrangement detection had a >94.0 % OPA when compared to FISH and an RNA-based sequencing method (n=270). 115 pan cancer samples were analyzed for MSI and demonstrated an OPA of 100% with PCR-based approaches.

Conclusions: Our 500+ gene, pan-cancer panel provides accurate and reproducible results for the detection of clinically relevant genomic alterations pan-cancer, without the need for matched normal samples. Further verification and validation studies of this gene panel are ongoing. This targeted gene panel will employ a decentralized, kitted model to empower local labs and allow for delivery of highly accurate and timely results.

1549   Digital Image Analysis to Detect Isolated Tumor Cells Positive Lymph Nodes

Sayak Ghatak[1], Kimmie Rabe[2], Irina Stout[1], Mahmoud Khalifa[2]
[1]University of Minnesota, Minneapolis, MN, [2]University of Minnesota - Twin Cities, Minneapolis, MN

Disclosures: Sayak Ghatak: None; Kimmie Rabe: None; Irina Stout: None; Mahmoud Khalifa: None

Background: Isolated tumor cells (ITCs) within lymph node has emerged as a new pattern of lymph node metastasis in several malignancies. Endometrial carcinoma with microcystic elongated and fragmented (MELF) pattern of myoinvasion has a high metastatic potential, thus lymph node biopsy is an important tool in the correct staging and management. The current management protocols of endometrial cancer do not consider the presence of ITCs as a deciding factor for adjuvant therapy, partly because of the lack of standardized laboratory guideline for their detection. In this study, we show the proof of concept of a novel digital image analysis (DIA)-based diagnostic algorithm of sentinel and non-sentinel lymph nodes to detect ITCs in metastatic endometrial carcinoma with MELF pattern of myoinvasion.

Design: Slides of ITC-positive lymph nodes, confirmed by microscopic evaluation (hematoxylin and eosin) and cytokeratin AE1/AE3 were scanned using Aperio ImageScope™. The whole slide images (SVS file format) were accessed using QuPath and representative screenshots were acquired as Tagged Image File Formats. DIA was performed using open-source software QuPath and FIJI.

Results: Figure 1 demonstrates ITCs detection using QuPath in hematoxylin and eosin and hematoxylin-DAB images. Figure 2 demonstrates ITCs quantification using FIJI in hematoxylin-AEC images.
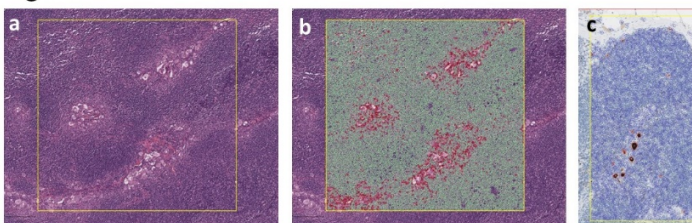
Figure 1 - 1549

Figure 1: ITC detection using QuPath. For morphologic analysis of hematoxy[...] images, representative regions of interest (ROI) of lymph nodes containing ITC are [...] training areas in QuPath. ROIs are then used to create a customized classifier (tu[...] tumor) that include stain vector deconvolution and built-in cell detection p[...] Hematoxylin Optical Density. Using the QuPath classifier, new ROIs (Fig. 1a) ar[...] tumor (red) and non-tumor (green) detection (Fig. 1b). For cytokeratin AE1/AE[...] detection, ROI-based positive cell detection function in QuPath is used to detec[...] with DAB (Fig. 1c).

Figure 2 - 1549
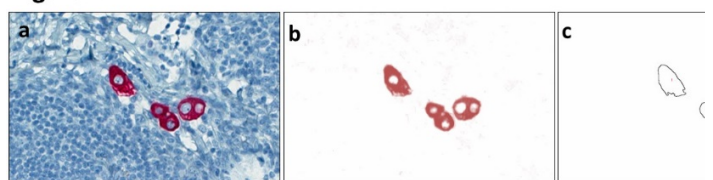
Figure 2: ITCs quantification using FIJI. Quantification of ITCs is done using FIJI. Tr[...] images (Fig. 2a) are deconvoluted into three different color channels using RC[...] vectors. The deconvoluted AEC color channel (R: 0.2743, G: 0.6796, B: 0.6803) sho[...] ITCs (Fig. 2b) is selected to impart a pixel intensity and pixel area threshold. The t[...] number of ITCs in the given ROI is calculated using particle analysis function on FIJ[...]

Conclusions: In this proof of concept study, we demonstrate two independent DIA methods of detection and quantification of ITC as lymph node metastasis in endometrial cancer with MELF pattern of myoinvasion. DIA can be used to complement conventional microscopy for detection of ITCs in lymph nodes. DIA can be used for quantification of metastatic burden in lymph nodes. DIA techniques developed in this study using open-source software can be transferable to detect ITCs in lymph node for other cancers.

1550 A Validation Study of HER2 Immunohistochemistry Digital Imaging Analysis and its Correlation with HER2 Fluorescence In Situ Hybridization Results in Breast Carcinoma

Ramon Hartage[1], Aidan Li[2], Scott Hammond[3], Anil Parwani[3]
[1]The Ohio State University Wexner Medical Center, Westerville, OH, [2]Jerome High School, Dublin, OH, [3]The Ohio State University, Columbus, OH

Disclosures: Ramon Hartage: None; Aidan Li: None; Anil Parwani: None

Background: The Visiopharm HER2 digital imaging analysis (DIA) algorithm assesses digitized HER2 immunohistochemistry (IHC) by measuring cell membrane connectivity. We aimed to validate this algorithm for clinical use by comparing with pathologists' scoring and correlating with HER2 fluorescence in situ hybridization (FISH) results.

Design: The study cohort consisted of 612 consecutive invasive breast carcinoma specimens including 395 biopsies and 217 resections. HER2 IHC slides were scanned using Philips IntelliSite Scanners and the digital images were analyzed using Visiopharm HER2-CONNECT App to obtain the connectivity values (0-1) and scores (0, 1+, 2+, and 3+). HER2 DIA scores were compared with Pathologists' manual scores and HER2 connectivity values were correlated with *HER2* FISH results. (Figure 1)

The HER2 IHC and the connectivity analyzed by Visiopharm HER2 IHC algorithm is demonstrated in the figure below as follows: A, B) one case with HER2 IHC 1+; C, D) one case with HER2 IHC 2+. E, F) one case with HER2 IHC 3+. A, C, E) HER2 IHCs; B, D, F) HER2 connectivity (green colored line) detected by Visiopharm.

Results: The concordance between HER2 DIA scores and pathologists' scores was 87.3% (534/612). All discordant cases (n=78) were only one-step discordant (negative to equivocal, equivocal to positive, or vice versa). Five cases (0.8%) showed discordant HER2 IHC DIA and *HER2* FISH results, but all these cases had relatively low *HER2* copy numbers (between 4 and 6). (Table 1) HER2 IHC connectivity showed significantly better correlation with *HER2* copy number than *HER2/CEP17* ratio (r^2: 0.851167 vs 0.818615).

Table 1. The correlation between HER2 DIA scores and pathologists' scores in 612 cases and FISH results in 442 cases.

		Visiopharm			
		Negative (0/1+)	Equivocal (2+)	Positive (3+)	Total
Pathologists	negative (0/1+)	407	25	0	432
	equivocal (2+)	41	54	6	101
	positive (3+)	0	6	73	79
	Total	448	85	79	612
HER2 FISH	Group 1 (FISH+)	3 (0.9%)	6 (9%)	53 (91.4%)	62
	Group 3 (FISH+)	0 (0)	1 (1.5%)	3 (4.5%)	4
	Group 2(FISH-)	1 (0.3%)	0 (0)	0 (0)	1
	Group 4 (FISH-)	36 (11.4%)	24 (35.8%)	2 (3.0%)	62
	Group 5 (FISH-)	277 (87.4%)	36 (53.7%)	0	313
	Total	317	67	58	442

Note: Visiopharm HER2 connectivity cut off value: 0: =0; 0< 1+ <0.12; 0.12 =<2+<0.49; 3+>0.49.

Figure 1 - 1550
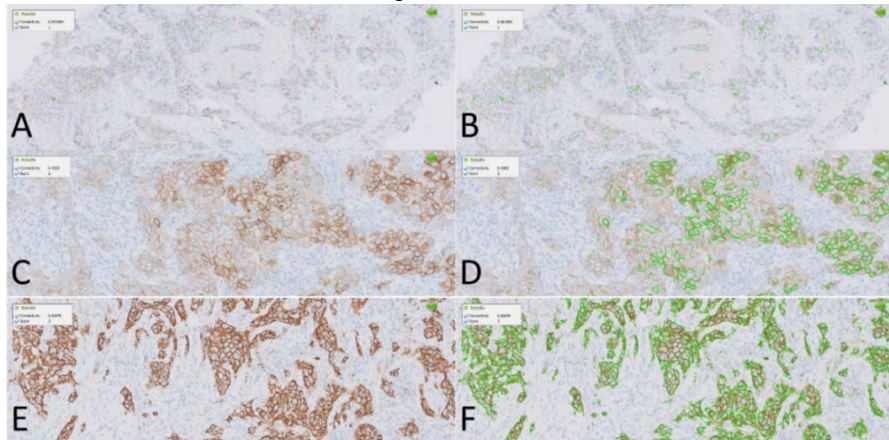
Conclusions: HER2 IHC DIA demonstrates excellent concordance with pathologists' scores and accurately discriminates between *HER2* FISH positive and negative cases. HER2 IHC connectivity has better correlation with *HER2* copy number than *HER2/CEP17* ratio, suggesting *HER2* copy number may be more important in predicting HER2 protein expression, and response to anti-HER2 targeted therapy.

1551 Artificial Intelligence Driven Neoadjuvant Chemotherapy Response Prediction in Triple Negative Breast Cancer (TNBC) Unveils Non-linear Feature Interactions

Zhi Huang[1], Zhi Han[2], Anil Parwani[3], Kun Huang[4], Zaibo Li[5]
[1]Purdue University, West Lafayette, IN, [2]Indiana University, Indianapolis, IN, [3]The Ohio State University, Columbus, OH, [4]Indiana University School of Medicine, Carmel, IN, [5]The Ohio State University Wexner Medical Center, Columbus, OH

Disclosures: Zhi Huang: None; Zhi Han: None; Anil Parwani: None; Kun Huang: None; Zaibo Li: None

Background: Pathologic complete response (pCR) to neoadjuvant chemotherapy (NAC) is a surrogate for disease-free survival in patients with triple negative breast cancer (TNBC). Identifying factors predicting pCR is crucial for clinical management and early intervention. We aimed to develop an ensemble model driven by artificial intelligence to predict the response to NAC in TNBC using clinical and image-based features.

Design: The cohort included 65 TNBCs treated with NAC and resection. Twenty-seven had pCR, while 38 had residual tumor. A multi-color immunohistochemical (IHC) assay detecting PD-L1, CD8 and CD163 was performed on biopsy sections. Tissue-wised affine registration was performed on each pair of slices. Color-based image thresholding was performed to segment CD8, CD163, and PD-L1 cells. 19 tissue-level histologic and immunostaining features together with age were collected to predict the response to NAC. Eight machine learning models were constructed, including logistic regression, neural network, decision tree, random forest classifier, AdaBoost classifier, Gradient Boosting classifier, XGBoost, and linear Support Vector Machine. All ensemble algorithms used 1000 estimators. 5-fold cross-validation on each dataset 10 times was performed. The accuracy of predictions was measured with four different metrics: AUC, F1-score, precision, and recall.

Results: Among all models, AdaBoost achieved highest performance with AUC = 0.7540, F1-score = 0.6388, precision = 0.6101, and recall = 0.6089. The decision tree also got the comparable results with AUC = 0.7506, while the linear model logistic regression only achieved AUC = 0.5856, suggesting non-linear interactions exist among features. By inspecting the feature importance as well as the decision tree plot, we found that the most pivotal feature was related to CD8. The certain ratio range of CD8 cell count to tissue area (> 1e-3 but <= 3e-3) leads to better outcome (pCR). Other variables such as the ratio of PD-L1 area to tumor area and age also played important role. A clinically applicable decision tree was then constructed to assist pathologists. The importance of the features was also calculated and ranked along with the decision tree.

Table 1. Nineteen tissue-level histologic and immunostaining features and age with their feature importance in accordance with AdaBoost algorithm. Features are sorted in descending order according to their Gini importance.

Feature name	Feature importance	Feature name	Feature importance
Ratio of CD8 cell count to tissue area	0.10154	Ratio of CD8 cell count to stroma area	0.04684
Ratio of tumor area to tissue area	0.09102	Ratio of CD8 cell count to tumor area	0.03992
Age	0.08884	Ratio of CD163 cell count to tumor area	0.03846
Ratio of CD8 area to tumor area	0.08560	Ratio of CD163 area to stroma area	0.03550
Ratio of CD8 area to stroma area	0.07710	Ratio of PDL1 cell count to stroma area	0.03042
Ratio of CD8 area to tissue area	0.07246	Ratio of PDL1 cell count to tumor area	0.02314
Ratio of PDL1 area to tumor area	0.06026	Ratio of CD163 area to tissue area	0.01378
Ratio of PDL1 area to stroma area	0.05542	Ratio of PDL1 area to tissue area	0.01314
Ratio of CD163 cell count to stroma area	0.05384	Ratio of CD163 cell count to tissue area	0.00982
Ratio of CD163 area to tumor area	0.05326	Ratio of PDL1 cell count to tissue area	0.00964

Figure 1 - 1551

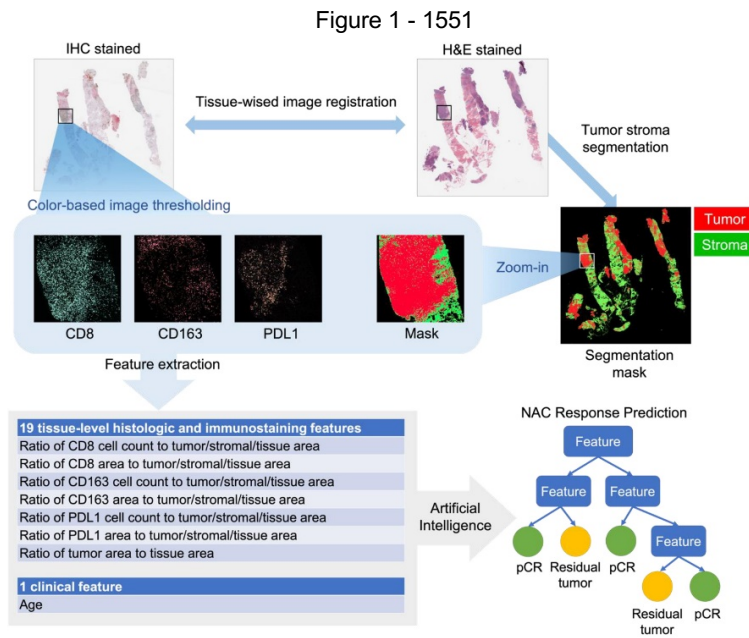Figure 1. Overview of our pipeline. H&E and IHC slides were digitally scanned into whole slide images. Tissue-wised affine registration was performed on each pair of slices. Color-based image thresholding was performed on L*a*b* color space to segment CD8, CD163, and PDL1 cells. Nineteen tissue-level histologic and immunostaining features and age (clinical feature) (Table 1) are able to predict the response to NAC.
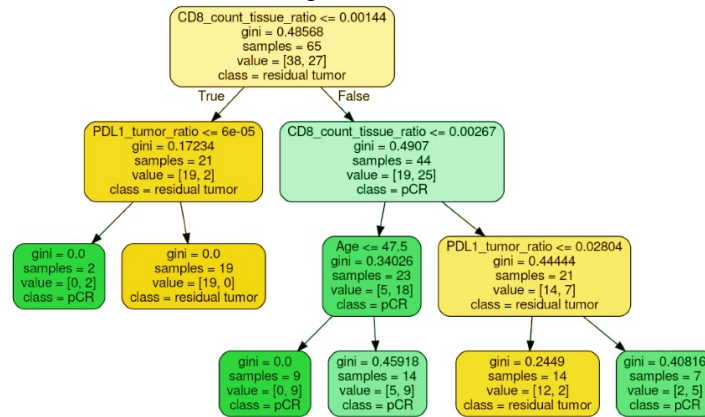
Figure 2 - 1551

Figure 2. Depth=3 decision tree constructed according to the tissue-level histologic, immunostaining, and clinical features. Rounded boxes with orange color: predict residual tumor (bad outcome); with blue color: predict pCR (good outcome).

Conclusions: Our results have demonstrated that artificial intelligence driven NAC response prediction by the ensemble method AdaBoost is able to achieve an acceptable performance. Tissue-level features derived from H&E and IHC images provide important information in a non-linear pattern.

1552 Mitosis Detection with Tiny-YOLO
Kenji Ikemura[1], Yukako Yagi[2]
[1]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY, [2]Memorial Sloan Kettering Cancer Center, New York, NY

Disclosures: Kenji Ikemura: None; Yukako Yagi: None

Background: A convolutional neural networks (CNN) model named "You Only Look Once (YOLO)," has gained recognition for its fast and efficient object detection tasks. Tiny-YOLO is a more compact model suited for real-time object detection. It is also used to quickly assess whether or not YOLO can learn a particular task before training it on a larger CNN model such as YOLOv3. In this study we tested the ability of YOLOv3-tiny (a version of tiny-YOLO) to detect mitoses in histological images.

Design: This study utilizes breast cancer images downloaded from a publicly available repository: https://mitos-atypia-14.grand-challenge.org/. The training set was composed of 928 histological images at 40x magnification from an Aperio scanner. For each image, we made a CSV file with following information: 1) whether or not mitoses exist in the image, 2) the coordinates of each mitotic figures, and 3) height and width of mitotic figures. Test set is composed of 20 unique H&E slides scanned at our institution on Aperio and 3D Histech (total of 40 images). The "gold standard" of mitosis is when three out of five pathologists mark the object as mitosis. Tiny-YOLO was configured to have batch size of 64 and run for 600 epochs. Evaluation of tiny-YOLO was made by calculating sensitivity, precision (positive predictive value), and F-measure = 2*sensitivity* precision/(sensitivity + precision).

Results: The trained tiny-YOLO tested on the Aperio test set had a sensitivity of 20%, precision of 19%, and F-measure of 0.20. Tiny-YOLO tested on 3D Histech test set had a sensitivity of 28%, precision of 15%, and F-measure of 0.20. Sample output images shown in Figure 1: A) Aperio image annotated by 5 pathologists (different colored circle for each pathologist). There are 2 mitoses in this image according to majority vote. B) Tiny-YOLO tested on Aperio image. C) 3D Histech image annotated by pathologists. D) Tiny-YOLO tested on 3D Histech image.

Figure 1 - 1552
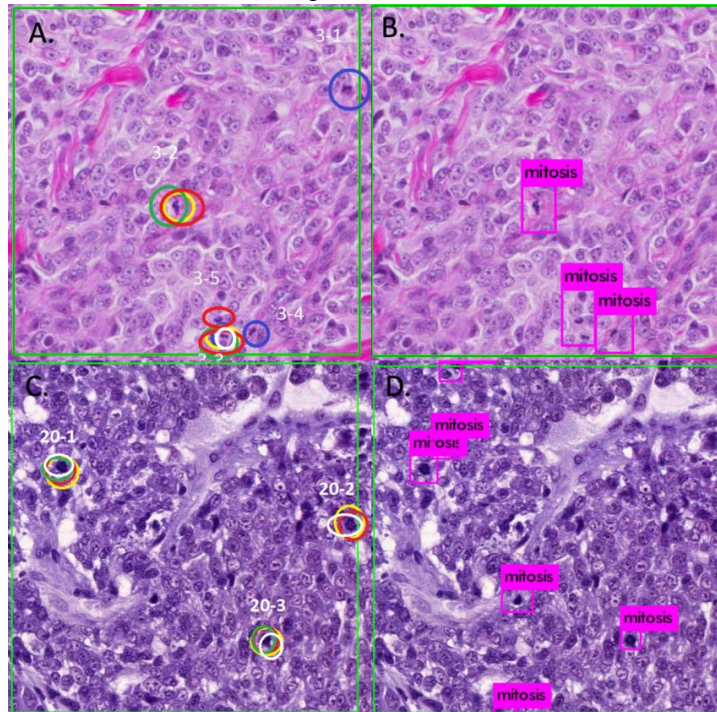
Conclusions: Although our evaluation score appears low, it clearly shows that tiny-YOLO is learning to distinguish mitosis from non-mitosis. From this result, we anticipate that full YOLO architecture can also learn to detect mitosis and with higher performance. Tiny-YOLO was also able to transfer what it learned from Aperio images to 3D Histech scanned images. These findings show the potential to generalize learning over different scanners. Moving forward, we plan to expand the data set through image augmentation and add images from various scanners to acquire generalizability.

1553 SOX10 Virtual Immunohistochemistry: The First Systematic Evaluation of a Machine Learning Algorithm with Single-Cell Resolution

Christopher Jackson[1], Keegan O'Hern[2], Meagan Chambers[2], Louis Vaickus[3], Sriharan Aravindhan[1]
[1]Dartmouth-Hitchcock Medical Center, Lebanon, NH, [2]Dartmouth-Hitchcock Medical Center and Geisel School of Medicine at Dartmouth, Lebanon, NH, [3]Etna, NH

Disclosures: Christopher Jackson: None; Keegan O'Hern: None; Meagan Chambers: None; Louis Vaickus: None; Sriharan Aravindhan: None

Background: Machine learning algorithms to predict immunohistochemistry (IHC) are being developed. To date, no algorithm capable of providing IHC prediction data with single-cell resolution has been created. We created such an algorithm using unsupervised machine learning techniques and tested its results against more traditional methods. Particular attention is placed on estimating the fraction of tumor cells in a lesion, as this is a vital quality metric in molecular diagnostics. Currently, the tumor fraction is estimated by a pathologist. Here, we employed the algorithm to predict the SOX10 immunophenotype of individual cells on H&E whole slide images (WSI) containing melanoma, allowing for the fraction of tumor cells to be automatically calculated. To our knowledge, this is the first systematic evaluation of single-cell resolution virtual IHC (vIHC).

Design: H&E WSIs were created by scanning 12 H&E slides at 400x. The slides were de-stained, received SOX10 IHC, and were again scanned at 400x. The resulting SOX10 WSIs were digitally registered (aligned) with the H&E WSIs. The SOX10 WSI was then used to precisely annotate the H&E nuclei as either SOX10-positive or SOX10-negative. These annotated H&E images were used to train a segmentation neural network to produce SOX10 virtual Immunohistochemistry (vIHC) images.

To evaluate the vIHC network, a separate set of 1000 x 1000 pixel H&E images were produced from clinical practice. An experienced board-certified dermatopathologist manually labeled approximately 8,000 melanocytic and non-melanocytic nuclei in these images. The SOX10 vIHC analyzed the raw H&E images, and the predicted number of SOX10-positive and SOX10-negative cells was calculated.

Results: The mean error for the tumor cell fraction was 20.95%. 11/25 (44%) of the cases were within 10% of the true tumor cell fraction, while 14/25 (56%) were within 15%.

Conclusions: This marks the first SOX10 vIHC algorithm capable of single-cell resolution. It is also the first systematic evaluation of such an algorithm using tumor cell fraction. Although vIHC was highly accurate in many cases, error was significant in others. Future directions include increasing the accuracy of the algorithm by using larger data sets. The work is an example of the potential for machine learning to rapidly and cost-effectively aide pathologists in both clinical and research settings.

1554 Whole Slide Image Registration with Machine Learning

Christopher Jackson[1], Robert Hamilton[2], Louis Vaickus[3]
[1]Dartmouth-Hitchcock Medical Center, Lebanon, NH, [2]Dartmouth Hitchcock Medical Center, Lebanon, NH, [3]Etna, NH

Disclosures: Christopher Jackson: None; Robert Hamilton: None; Louis Vaickus: None

Background: Unsupervised machine learning methods have been developed that use immunohistochemistry (IHC) whole slide images (WSI) to annotate hematoxylin and eosin (H&E) WSIs. Advanced versions of these methods have used IHC performed on the same cell layer as H&E following a de-staining step. These methods have historically relied on iterative registration techniques that are slow, computationally demanding, and occasionally fail to converge. Here, we demonstrate that a neural network can be created to automatically register H&E and IHC images.

Design: 12 H&E slides from clinical practice were scanned at 400x. The slides were de-stained, and SOX10 IHC was performed on the destained slides which were again scanned at 400x. The H&E and SOX10 IHC WSIs were registered using iterative registration techniques. Sub-images were created by setting a grayscale transformation of the H&E image to the red color channel. The IHC image then underwent a geometric translation by randomly generated X and Y distances, up to 100 pixels. A grayscale transformation of the translated IHC image was assigned to the green channel. These images were then labeled using their X and Y translation distances, and their resolution was reduced by a factor of 3. 49,968 such images were assigned to a training group, and an additional 5,552 were assigned to a validation group. A network was trained for 16 epochs to predict the labeled X and Y distances for the images in the training group. The mean error of the validation group was calculated. The resulting neural network was tested on an additional 50 randomly generated test images and compared to an iterative registration technique. The mean computation times were calculated.

Results: The neural network output mean-error for the validation group was 8.18 ± 14.05 pixels, approximately half the diameter of a red blood cell (19 pixels). The mean time to register one 1,197 x 1,197-pixel image using the iterative registration method was 49.36 ± 12.06 seconds, compared to 0.18 ± 0.02 seconds for the registration neural network. Extrapolating these results to a 40,000 x 40,000 WSI, it would take 15.8 hours to register the WSI using the iterative technique, compared to 3.5 minutes to register the image using the neural network. The iterative registration method failed to converge in 4/50 (8%) of cases, something that did not occur with the neural network.

Conclusions: Rapid, high-quality WSI registration using machine learning is an ideal alternative to the classic iterative registration.

1555 Digital Image Analysis for Estimating Stromal CD8+ Tumor Infiltrating Lymphocytes in Lung Adenocarcinoma

Iny Jhun[1], Daniel Shepherd[2], Yin Hung[2], Emilio Madrigal[3], Long Le[2], Mari Mino-Kenudson[2]
[1]Stanford Health Care, Palo Alto, CA, [2]Massachusetts General Hospital, Boston, MA, [3]Boston, MA

Disclosures: Iny Jhun: None; Daniel Shepherd: None; Yin Hung: None; Emilio Madrigal: None; Long Le: *Advisory Board Member*, ArcherDx; *Stock Ownership*, ArcherDx; Mari Mino-Kenudson: None

Background: CD8+ tumor infiltrating lymphocytes (TILs) along with PD-L1 have predictive relevance in patients with non-small cell lung cancer (NSCLC) undergoing immunotherapy. While manual counting of PD-L1+ tumor cells is clinically applied, a scoring system for CD8+ TILs has not been established. There is growing evidence on stromal CD8+ TILs as a prognostic and predictive indicator, however, high inter-observer variability in CD8+ TIL density estimates limits its potential clinical application. We aimed to test the feasibility of a digital image analysis (DIA) workflow for calculating stromal CD8+ TIL density in dual PD-L1/CD8 stained slides of lung adenocarcinoma.

Design: A total of 26 lung adenocarcinoma cases were retrieved and digitized at 40x magnification (0.2214 μm pixel width and height). A DIA workflow was developed in *QuPath* (Belfast, UK) and comprised of stain estimation, cell detection, feature computation, and stain intensity classification. Two pathologists manually quantitated the percentage of stromal CD8+ TILs from digitized slides within a pre-defined specified region of interests (ROIs) with at least 200 stromal cells. A random tree classifier was built using the pathologists' annotations from 14 training cases and applied to 12 validation cases. The DIA-estimated CD8+ TIL densities were compared to manual estimates.

Results: A DIA workflow was applied to 26 dual PD-L1/CD8 stained slide images with an average of 488 (199-949) stromal cells. An average of 16.0% (0.2-54.9%) of stromal cells were CD8+ TILs. The estimated total stromal cell count and CD8+ TIL densities were in good agreement with pathologists' estimates (Table). The relative difference between DIA-estimated and pathologists' total stromal cell counts (12% in training cases, 17% in validation cases) were comparable to that in cell counts between the two pathologists (12% in training cases, 13% in validation cases). DIA-estimated stromal CD8+ TIL densities were similar to the pathologists' estimates, with a difference of 3.4% in training cases and 1.5% in validation cases. The concordance between the DIA-estimated CD8+ TIL densities and average pathologists' estimates (average difference of 2%) was higher than that between pathologists (average difference of 9%) (p=0.01).

		Total Cells						CD8+ TIL Density (%)			
		Pathologists		QuPath				Pathologists		QuPath	
	Case #	Average	% Difference[a]	Predicted	% Difference[b]			Average	% Difference[c]	Predicted	% Difference[b]
Training	1	550.5	19%	463	-19%			49%	41.1%	55%	-5.4%
	2	653.0	-21%	199	-228%			6%	2.5%	4%	2.2%
	3	748.5	-26%	704	-6%			31%	8.0%	24%	6.5%
	4	445.0	-23%	418	-6%			20%	9.6%	19%	1.1%
	5	631.5	-12%	474	-33%			38%	23.2%	35%	3.0%
	6	640.0	-29%	949	33%			5%	-0.7%	1%	3.3%
	7	573.5	-22%	539	-6%			23%	10.0%	19%	4.1%
	8	493.5	-14%	452	-9%			5%	1.3%	0%	4.3%
	9	365.0	-5%	277	-32%			9%	3.0%	8%	0.5%
	10	728.0	0%	600	-21%			27%	10.6%	24%	2.3%
	11	481.5	-7%	441	-9%			6%	3.5%	4%	2.0%
	12	553.0	-3%	554	0%			25%	19.4%	11%	13.6%
	13	455.5	-20%	364	-25%			12%	0.4%	9%	3.1%
	14	614.0	-31%	505	-22%			54%	37.3%	48%	6.5%
	Average		**-14%**		**-12%[d]**				**12.1%**		**3.4%**
Validation	15	405.0	-11%	364	-11%			6%	1.5%	7%	-0.9%
	16	368.0	-16%	277	-33%			30%	24.1%	34%	-3.8%
	17	631.0	-20%	535	-18%			19%	8.0%	19%	0.1%
	18	707.0	-36%	614	-15%			20%	5.9%	17%	2.9%
	19	734.5	-12%	628	-17%			8%	4.0%	2%	5.9%
	20	721.5	-9%	626	-15%			15%	4.7%	12%	2.2%
	21	691.5	0%	596	-16%			10%	3.2%	7%	3.1%
	22	284.5	-12%	274	-4%			11%	2.5%	8%	2.6%
	23	748.0	-24%	537	-39%			10%	1.8%	13%	-2.7%
	24	533.5	-9%	426	-25%			14%	3.0%	9%	5.8%
	25	515.5	-14%	437	-18%			18%	3.5%	17%	1.3%
	26	415.0	6%	445	7%			11%	3.3%	10%	0.9%
	Average		**-13%**		**-17%**				**6.0%**		**1.5%**
	[a] Relative difference = (pathologist 1 - pathologist 2)/average										
	[b] Note: QuPath differences are relative to the average pathologists scores										
	[c] Actual difference = pathologist 1 - pathologist 2										
	[d] This average excludes case 2										

Conclusions: We demonstrate the feasibility of applying a DIA workflow for estimating stromal CD8+ TILs in NSCLC. DIA can provide an efficient and standardized approach for quantifying stromal CD8+ TIL density.

1556   Abstract Withdrawn

1557   Automatic Assessment of Tumor Cellularity in Histopathology Images Using Weakly-Supervised Segmentation

Shivam Kalra[1], Amir Safarpoor Kordbacheh[1], Morteza Babaie[1], Hamid Tizhoosh[1]
[1]University of Waterloo, Waterloo, ON

Background: Tumor cellularity is the quantification of the relative ratio of tumor and normal cells. It is an important part of tumor assessment in histopathology. In current clinical practice, tumor cellularity is estimated by pathological review of sectioned specimens, usually stained with hematoxylin and eosin (H&E). However, the manual assessment affects both the quality and the reliability of an estimate due to the inter-observer variability.

Design: We have developed a machine-learning pipeline for automatically assessing the tumor cellularity. It consists of two steps. Firstly, a nuclei segmentation technique using color deconvolution. The nuclei regions correspond to the hematoxylin (H) component of the H&E staining. Secondly, a method to estimate the tumor cellularity by using a weakly-supervised training. We utilized a popular deep network for semantic-segmentation, namely U-Net, to infer the probability of each pixel in the image as: tumor cell, non-tumor cell, or none. The cellularity of a patch is calculated as the ratio of the expectation of two probability masks (tumor versus normal cells). We trained the network end-to-end with two loss functions: i) *binary cross entropy* between the nuclei masks obtained from the color deconvolution and the nuclei masks obtained from the segmentation, and ii) *regression loss* of the calculated and the given cellularity values.

Results: We evaluated the efficacy of our approach on the dataset provided for the Cancer Cellularity Challenge, 2019, organized by SPIE. The dataset consists of 2,579 patches at 20x resolution from WSIs extracted from 64 patients with residual invasive breast cancer. Each patch in the dataset is assigned a cellularity score (a number between 0 to 1) by an expert pathologist. We used the data augmentation techniques to increase our training samples to up to 50,000 patches. We scored a high rank correlation between the computed and the pathologist's estimate of cellularity (>90%) on the unseen data.
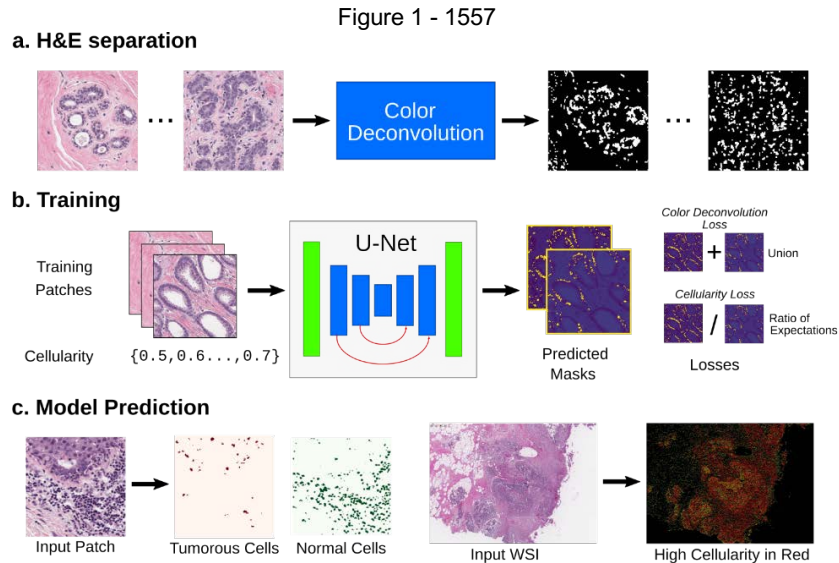

Figure 1 - 1557

a. H&E separation

b. Training

c. Model Prediction

Figure: Automatic assesment of tumor cellularity of H&E stained slides. (a) A H&E stained patch goes through a color deconvolution separating its H and E components. Only the H component representing the nuclei is selected. (b) The patch is passed to a U-Net to predict the tumorous and normal cell through the weak supervision via cellularity values and masks from (a). (c) The trained model can predict the high tumor cellularity regions in unseen H&E slides.

Conclusions: We have developed an automated approach for cellularity assessment of histopathology images. We used the data augmentation techniques to train the model with limited number of samples. The trained model is used to visualize the areas of high cellularity within WSIs. The results suggest that our method can reliably compute the tumor cellularity. Furthermore, our approach does not rely on annotations of individual nucleus since it automatically infers the high cellularity regions.

1558   A Single Section is Enough for Kidney Biopsy? Generative Adversarial Network Converts One Stain to Another While Preserving Underlying Histology (PPHM-GAN)
Masataka Kawai[1], Tetsuo Kondo[1], Kunio Mochizuki[1], Naoki Oishi[1]
[1]University of Yamanashi, Chuo, Yamanashi, Japan

Disclosures: Masataka Kawai: None; Tetsuo Kondo: None; Kunio Mochizuki: None; Naoki Oishi: None

Background: Renal biopsy is the golden standard for the diagnosis of some kidney diseases such as glomerulitis. Special stains such as H&E (Hematoxylin and Eosin), PAS, PAM, and MT (Masson's Trichrome) staining as well as various immunostains are performed routinely. These different staining techniques reflect histological structure and chemical composition of the tissue. Some structures or lesions are more readily identified with a certain stain than with others. Although close comparison between different stains on the same region of interest is often necessary, different sections are apart as near as several microns and as far as several hundred microns. This difference between sections is sometimes critical. To obviate these limitations, we developed a novel pseudo-stain converter named PPHM-GAN (PAS-PAM-H&E-MT-Generative Adversarial Network). GAN or Generative Adversarial Network is an ensemble of neural networks with generating and discriminating networks (a.k.a the generator and the discriminator respectively) that are trained competitively and (semi-)simultaneously. Recent advances have proven that GAN with convolutional neural networks can generate fake images indistinguishable from real images even by human eyes.

Design: We retrospectively collected 7 cases of renal biopsy from University of Yamanashi Hospital. All specimens were performed H&E, PAS, PAM, and MT staining before diagnosis. All slides were scanned by virtual slide scanner (VS-120, Olympus) at 40x. Several hundred patches (512x512px, 87x87 μm) per slide according to the size of the specimen were randomly cropped. PPHM-GAN has network architecture inspired by Star-GAN where the generator converts one stain to another while the discriminator classifies the stains as well as whether the image is fake or not. We adopted Wasserstein GAN with gradient penalty (WGAN-GP) as GAN loss. Converted stains or fake images were evaluated by FID (Fréchet Inception Distance) and skilled pathologists.

Results: Multi-stain to multi-stain conversion was successful at scale of 256x256px. Some difficulty in converting certain stains might come from similarity or vicinity of stains. E.g. Conversion from H&E to PAM is easier than that from PAS to PAM (FID are 1.74 and 1.817 respectively).

Figure 1 - 1558
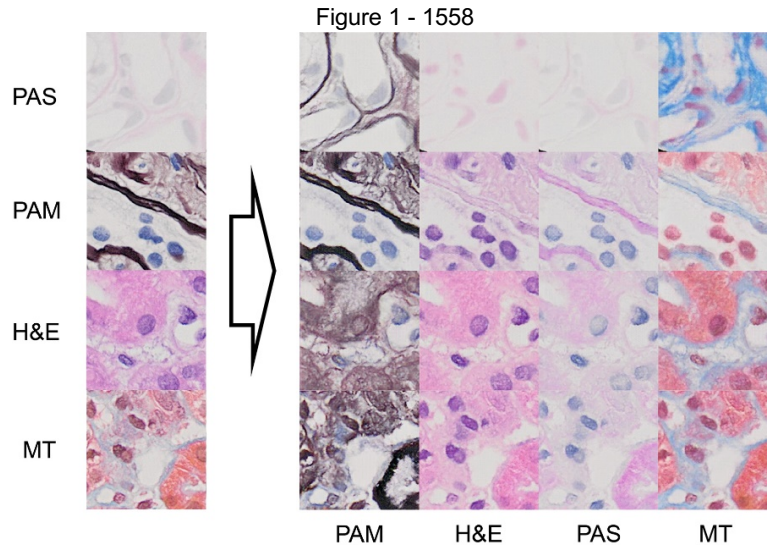
Conclusions: We proposed a GAN-based multi-stain converter named PPHM-GAN. Generated fake stains are plausible to humans at least patch level.

1559   Inter-Observer Variation in Image Quality Scoring

David Kellough[1], Trina Shanks[2], Atemnkeng Alemnji[3], Jillian Moorman[3], Anil Parwani[4], Mark Lloyd[1]
[1]Inspirata, Inc., Tampa, FL, [2]Inspirata, Inc., New Albany, OH, [3]Inspirata, Inc., Columbus, OH, [4]The Ohio State University, Columbus, OH

Disclosures: David Kellough: *Employee*, Inspirata, Inc; Trina Shanks: *Employee*, Inspirata, Inc.; Atemnkeng Alemnji: None; Jillian Moorman: None; Anil Parwani: None; Mark Lloyd: *Employee*, Inspirata, INc.

Background: QC is an important part of every endeavor and high-volume slide scanning is no exception. At the Digital Pathology Scan Center at the James Cancer Hospital at The Ohio State University, we review 1.4% of all WSIs we produce. Since go-live in May 2017, we have individually reviewed over 17,000 WSIs. To expedite this we developed an objective test to assess slide quality and assign a numeric score. This quantitative quality indicator allows us to track and compare scan quality between slide sources, across scanners, and longitudinally. Our life-of-project average score is 9.37 which comports with general pathologist satisfaction.

Most of our review has been undertaken by trained and experienced personnel. As digital adoption and slide scanning volumes skyrocket, we have needed to rely on addition reviewers and the question has arisen: Is the reliability of our QC review system based on the skill of the reviewer or the merits of the review process?

Design: Four observers at different experience levels reviewed WSIs: one highly experienced, one moderately experienced, and two novices. The same set of 10 slides (5 H&E, 2 IHC, and 2 special stains) were run on 5 different Philips UFS scanners on 5 different days (with washout days in-between) and reviewers asked to score image quality.

Slide images were scored according to our system. Whole slide images were given a 0 if significant areas were out-of-focus or cut-off by the ROI detection box. Slides not receiving a zero were then given a point for each of 5 random tissue areas in-focus at 20x and 5 random tissue areas at 40x. These scores were summed to produce a quality score of 0-10. Scores were recorded for each slide image and each reviewer and compared.

Results: The average SD for the 50 slide image views by the 4 reviewers was 0.93, indicating a fair degree of concurrence. Reviewers tended to score slides within a point more or less of one another.

The average SD for the 50 slide image views by the 2 most experienced reviewers was even more concurrent: 0.5. The two most experienced reviewers were in complete agreement on 24 of the 50 slide views. When not in agreement, the two tended to score slides within a point more or less of one another.

Conclusions: The conformity of the reviewers' scores says that our slide QC system is a valid and reliable indicator of the quality of the WSI. The greater conformity of the more experienced users' scores indicates that it is a system that works best with training and experience.

1560 Deep Learning Method for Pathology Image Compression and Reconstruction

Pratik Kubal[1], Scott Doyle[1]
[1]University at Buffalo, Buffalo, NY

Disclosures: Pratik Kubal: None; Scott Doyle: None

Background: Whole Slide Images (WSI) are large (2-5 GB per slide), necessitating image compression. Lossy compression greatly reduces disk space by removing data, while lossless compression maintains image data but cannot achieve high compression. We propose a lossy Deep Neural Network (DNN)-based approach to create an optimized image representation, which can be used to accurately reconstruct the image for viewing. We apply our algorithm to WSIs of H&E-stained colon cancer tissue from The Cancer Genome Atlas (TCGA) and compare the compression and reconstruction to lossy (JPG, downsampling) and lossless (PNG) approaches. Our method yields better compression and comparable image quality to existing algorithms, reproducing pathologically relevant image structures. This supports our method for clinical archival of pathology images.

Design: From TCGA Colon Cancer repository we collect 18 WSIs for training and 2 for validation. The DNN consists of an encoder, which creates a compressed image representation for storage, and a decoder, which reconstructs the full-size image. During training, data augmentation is performed by random flipping, rotation, and cropping to 256 by 256 image patches. Evaluation is done on lossless PNG images of 6400 by 6400 pixels, which are compressed using (a) the DNN approach outlined above, (b) bicubic resizing, and (c) JPEG compression with 75% quality. For evaluation, we calculate Peak Signal to Noise Ratio (PSNR), PSNR per compressed megabyte (MB), and compression ratio compared to the ground truth.

Results: Quantitative results are shown in Table 1. DNN-compressed images achieve higher compression versus JPG with comparable PSNR. Bicubic downsampling has higher compression, but with worse reconstruction image quality (Figure 1). In Figure 2, different patches are shown with their corresponding reconstructed images; the DNN-reconstructed images have a much greater image detail compared to bicubic downsampling, with greater compression ratio. Additionally, DNN images faithfully recreate diagnostically relevant tissue structures.

Table 1: Quantitative results of 6400 by 6400 Images

	Space on Disk per Tile in MB (Smaller is Better)				Content Metrics (Higher is Better)								
	Ground Truth	DNN Compressed	Bicubic Down-sampled	JPG	DNN-Reconstructed			Bicubic Up-sampled			JPG		
					PSNR	PSNR/MB	Compression Ratio	PSNR	PSNR/MB	Compression Ratio	PSNR	PSNR/MB	Compression Ratio
Mean	0.11 (±0.01)	0.01 (±0)	0.01 (±0)	0.01 (±0.01)	30.91 (±1.46)	4.51 (±0.32)	19.66 (±3.45)	28.68 (±2.45)	5.13 (±0.41)	23.94 (±4.06)	36.12 (±2.05)	3.86 (±0.47)	14.41 (±3.41)

Figure 1 - 1560

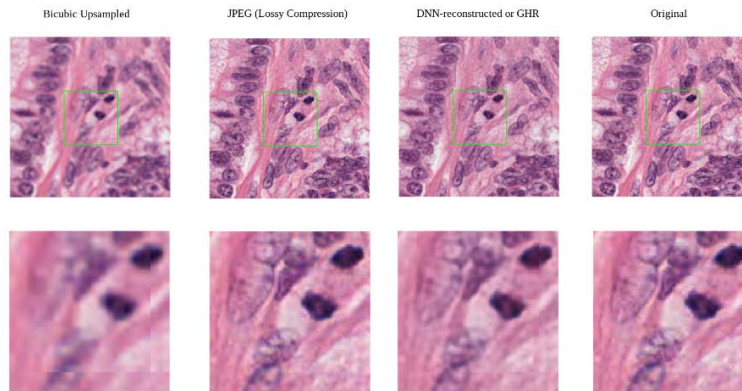Bicubic Upsampled JPEG (Lossy Compression) DNN-reconstructed or GHR Original

Figure 1: (From Left to Right) Bicubic Upsampled, JPEG compression on Original sample, DNN-Reconstructed from DNN-compressed representation, and Original sample

Figure 2 - 1560

Bicubic Upsampled JPEG (Lossy Compression) DNN-reconstructed or GHR Original
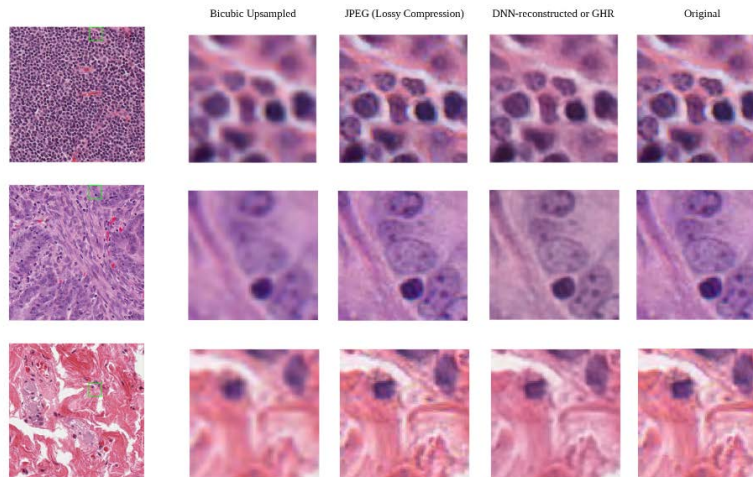
Figure 2: (Top to Bottom) Lymphocytes, Tumor Mucosa, Stroma and (Left to Right) Bicubic Upsampling, JPEG Compression on Original sample, DNN-reconstructed from DNN-compressed, and Original

Conclusions: DNN provides a better compression ratio versus JPG but with comparable image quality. Our future work will focus on developing a tile by tile compression approach, which will yield even better results by optimizing the compressed representation based on the underlying tissue type. We also plan to evaluate segmentation algorithms on the DNN-reconstructed images to demonstrate its ability to generate realistic-looking outputs.

1561   Histology Image Classification and Segmentation Labelling Platform

Chaitanya Kulkarni[1], Asmaa Aljuhani[1], Arunima Srivastava[1], Carly Vroom[1], Satoshi Hamasaki[1], Raghu Machiraju[1], Anil Parwani[1]
[1]The Ohio State University, Columbus, OH

Disclosures: Chaitanya Kulkarni: None; Asmaa Aljuhani: None; Arunima Srivastava: None; Carly Vroom: None; Satoshi Hamasaki: None; Raghu Machiraju: None; Anil Parwani: None

Background: Histology whole slides, are rich sources of information in a clinical setting. We now harness computational models that learn from these images to predict patient outcomes. Precise domain expert annotations of images help computational models focus on relevant features of the images. Most existing tools that facilitate image annotations are (a) proprietary or (b) difficult to scale with large images, or (c) not intuitive or conducive to specific annotation tasks. We present, Label-Coach, an open source, scalable, web-based tool to aid feature rich and purposeful annotations and subsequent image especially in the context of machine and deep learning.

Design: Features of Label-Coach include support for a wide range of image formats, free-hand drawing tools for robust segmentation and shape overlays (e.g., bounding boxes and circles) on the image. Collaboration between multiple annotators and rapid curation is also feasible. Additionally, large scale image data is supported with the use of Amazon Web Services and custom file management. Since histology images typically contain related compartments (tumor regions, cells, epithelium etc.), Label-Coach facilitates hierarchical annotation, with various associated labels depicting related components as explained in the results section.

Results: Label-Coach is being used to collect ground truth data for various clinical initiatives. It has been utilized to annotate wounds in photographs of burn victims' appendages, enabling the prediction of degree of burns. It has also been employed for annotations of "tall cells" in tall cell variant of thyroid cancer histology images, thus enabling robust measurement of cell morphology. Lastly, due to the unique hierarchical annotation feature, Label-Coach is being used to collect data at multiple levels within breast invasive carcinoma histology slides to annotate (1) tumor regions, (2) apoptotic and highly mitotic regions and (3) specific tumor cells with high nuclear grade, to model and predict the grade via the Nottingham Score.

Figure 1 - 1561

Label-Coach Interface

Region Of Interest Bounding Box   Free Hand Region   Cell Segmentation
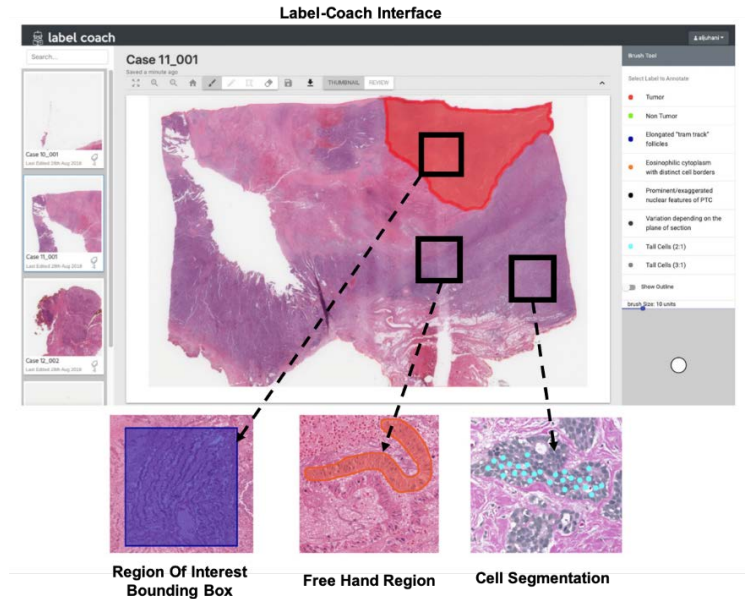
Conclusions: We have developed a robust annotation tool to infer specific understanding from diagnostic histology images. This tool enables computational models to harness useful features from these images to predict patient attributes and outcomes. Label-Coach is freely available to use through approved access from our team. We also make our code available on github.com/chaitanya2334/label-coach.

1562   Methodological Sampling Validation for Digital Image Analysis in Non-Small Cell Lung Carcinomas (NSCLC) and Tumor Heterogeneity Implications

Caddie Laberiano Fernandez[1], Ignacio Wistuba[1], Edwin Parra[1], Ruth Salazar Alejo[2], Jose Solorzano[3]
[1]The University of Texas MD Anderson Cancer Center, Houston, TX, [2]MD Anderson, Lima, Peru, [3]MD Anderson Cancer Center Madrid-España, Madrid, Spain

Disclosures: Caddie Laberiano Fernandez: None; Ruth Salazar Alejo: None; Jose Solorzano: None

Background: Digital image analysis is increasingly being used in research to accurately quantify different cell populations within the entire tissue section. Random sampling of tissue areas has been the methodology used by pathologists in order to pin down information using efficiently time and resources, being necessary a better assessment for tumor heterogeneity. The goal of this study was to validate a sampling methodology to quantify high and low density immune markers and tumor heterogeneity implications in NSCLC.
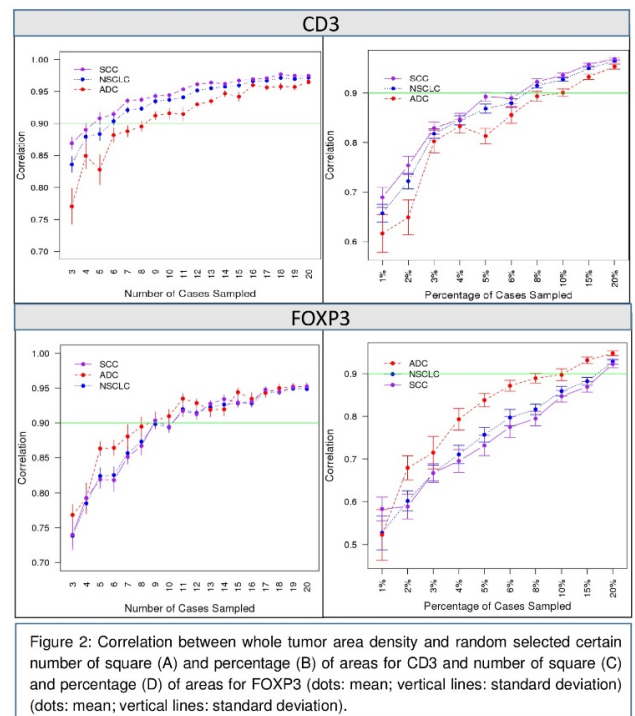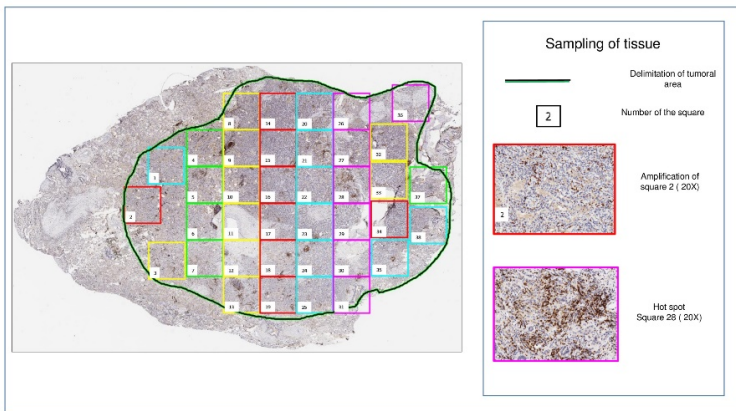
Design: The evaluation was performed on whole tumor area selected manually and on individual 1mm squares covering the same entire tumor area (Figure1).Fifty representative formalin fixed paraffin embedded (FFPE) tumor block sections from squamous cell carcinoma (SCC, N=31) and adenocarcinoma (ADC, N = 19) were used. The slides were automated stained for CD3 and FOXP3. Spearman's correlation coefficient test was used to compare the gold standard with randomly selected areas of analysis per each cohort. The slides were scanned and analyzed using an image analysis tools. Four groups were evaluated according the tumor subtype and the immune markers. There were excluded two SCC cases from further analysis that was out-layer having a small number of squares to evaluate.

Results: For immune marker CD3 and FOXP3, number of squares ranges were from 2 to 187. The increasing square's number corresponded to   a smaller variant of the mean estimation(Figure 2) . Although the correlation in both cases (ADC and SCC) started being moderated, there was an increase of to 0.9 in CD3 and FOXP3 in both kinds of tumors. SCC reached higher correlation than ADC cases across all the number of squares sampled in CD3 marker. To reach correlation of 0.9, it needed about 5 squares for SCC and 9 squares for ADC, along with 8% squares for SCC and 10% for ADC (Table1). FOXP3 in ADC achieved slightly higher correlation than SCC. To reach correlation of 0.9, it considered 10 squares for both ADC and SCC and about 10% squares for ADC and 20% for SCC. The tumor type implication was not obvious for FOXP3 as for CD3. We observed that CD3 expressed slight higher We observed that CD3 expressed slight higher in ADC than in SCC ($P = 0.062$).

			Number of squares									
			4	6	8	10	12	14	16	18	20	
	CD3	NSCLC	0.88	0.9	0.92	0.94	0.95	0.96	0.97	0.97	0.97	
		ADC	0.85	0.88	0.9	0.92	0.93	0.95	0.96	0.96	0.96	
		SCC	0.89	0.91	0.94	0.94	0.96	0.96	0.97	0.98	0.97	
	FOX P 3	NSCLC	0.79	0.83	0.87	0.89	0.91	0.93	0.93	0.94	0.95	
		ADC	0.79	0.86	0.89	0.91	0.93	0.92	0.93	0.95	0.95	
		SCC	0.79	0.82	0.87	0.89	0.91	0.93	0.93	0.94	0.95	
			Percentage of squares									
			1%	2%	3%	5%	6%	8%	10%	15%	20%	
	CD3	NSCLC	0.66	0.72	0.82	0.87	0.88	0.92	0.93	0.95	0.96	
		ADC	0.62	0.65	0.8	0.81	0.86	0.89	0.9	0.93	0.95	
		SCC	0.69	0.75	0.83	0.89	0.89	0.92	0.94	0.96	0.97	
	FOXP3	NSCLC	0.53	0.6	0.67	0.76	0.8	0.82	0.86	0.88	0.93	
		ADC	0.52	0.68	0.71	0.84	0.87	0.89	0.9	0.93	0.95	
		SCC	0.58	0.59	0.67	0.73	0.77	0.79	0.85	0.87	0.92	

Figure 1 - 1562

Figure 2 - 1562

Figure 2: Correlation between whole tumor area density and random selected certain number of square (A) and percentage (B) of areas for CD3 and number of square (C) and percentage (D) of areas for FOXP3 (dots: mean; vertical lines: standard deviation) (dots: mean; vertical lines: standard deviation).

Conclusions: In conclusion, our study shows the importance of standardizing the sampling methodology used in digital image analysis of IHC stains by examining a high and low density immune marker in NSCLC cases with different tumor heterogeneity.

1563 Generation of Synthetic Ovarian Carcinoma Microscopic Images for Proficiency Testing

Adrian Levine[1], Chien Chen Peng[2], David Farnell[3], Basile Tessier-Cloutier[1], Mitchell Nursey[4], Davood Karimi[5], Steven Jones[6], Septimiu Salcudean[1], David Huntsman[7], Stephen Yip[1], C. Blake Gilks[8], Ali Bashashati[9]
[1]University of British Columbia, Vancouver, BC, [2]University of British Columbia, Burnaby, BC, [3]Vancouver, BC, [4]BC Cancer Agency, Vancouver, BC, [5]The University of British Columbia, Vancouver, BC, [6]Michael Smith Genome Sciences Centre, Vancouver, BC, [7]British Columbia Cancer Research Institute, Vancouver, BC, [8]Vancouver General Hospital, Vancouver, BC, [9]University of British Columbia/Vancouver General Hospital, Vancouver, BC

Disclosures: Adrian Levine: None; Chien Chen Peng: None; David Farnell: None; Basile Tessier-Cloutier: None; Mitchell Nursey: None; Davood Karimi: None; Steven Jones: None; Septimiu Salcudean: None; David Huntsman: None; Stephen Yip: *Advisory Board Member*, Bayer; *Advisory Board Member*, Pfizer; *Advisory Board Member*, Roche; C. Blake Gilks: None; Ali Bashashati: None

Background: Deep learning methods used in digital pathology have demonstrated expert level performance in cancer detection and the ability to predict genomic alterations solely from images, however less work has been applied to the synthesis of imaging data. The generation of histology images has important applications, including use in education and proficiency testing, where a limited number of material can allow trainees to memorize individual test slides, rather than the general morphologic features of a histotype. We hypothesized that generative adversarial networks (GANs), a state-of-the-art deep learning method for synthesizing realistic images, could be applied to generate microscopic images of ovarian carcinoma.

Design: Our dataset consisted of 80 histopathology slides of ovarian carcinoma, comprising the five main histotypes (high-grade serous, low-grade serous, clear cell, endometrioid, and mucinous). These were scanned at 20x magnification and representative areas of cancer were annotated by a board-certified pathologist. To generate smaller training images, the digital slides were tessellated with patches of size 384 x 384 pixels, resulting in a total of 10889 images. We then used an implementation of GANs referred to as Progressive GAN to train a model to generate synthetic images. Following the completion of training, the generator component of the trained model was used to synthesize images for the five ovarian carcinoma subtypes (workflow in Figure 1).

Results: A representative set of high-resolution synthetic & real images of the five ovarian cancer histotypes is shown in Figure 2. A board-certified pathologist with digital pathology experience compared the synthetic and real images. Overall the pathologist reported that the synthetic images were less sharp than the real ones, but generally captured features characteristic of their subtype, although some patches were too small for accurate classification.

Figure 1 - 1563

Figure 2 - 1563
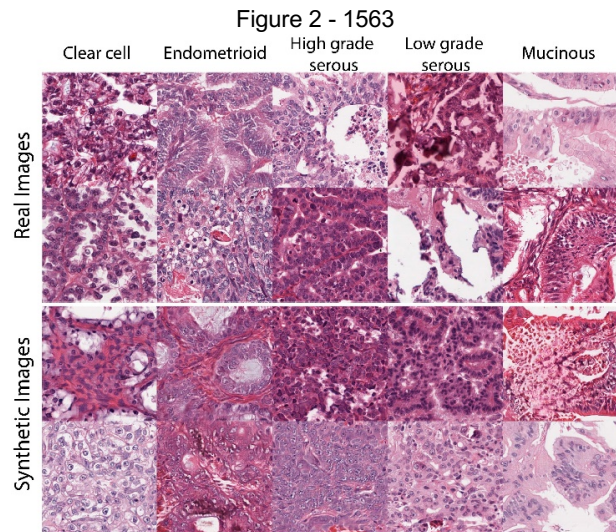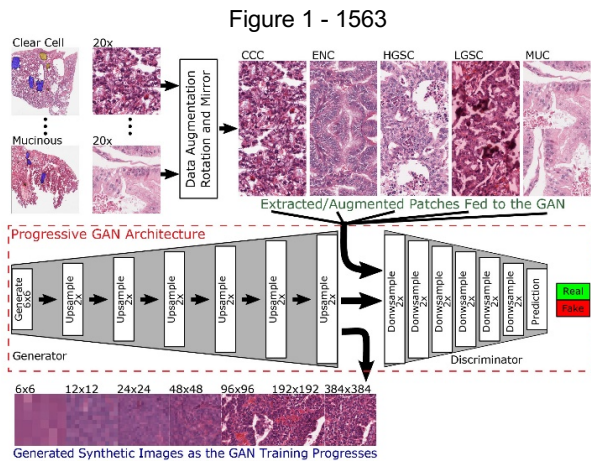
Conclusions: GANs can synthesize high-resolution histopathology images that capture salient features of different cancer types. These images are useful for augmenting small data sets in classifier training to achieve better performance in differentiating between histotypes and can provide an endless source of novel material for proficiency testing. In the future, the utility and validity of the synthesized images need to be evaluated by multiple pathologists. Furthermore, work is needed to develop methods to synthesize larger images and apply these to a wider range of cancer types.

1564 A Deep Learning Model Can Identify Rejection in Transplant Kidney Biopsies

Dmytro Lituiev[1], Sung Jik Cha[1], Jae Ho Sohn[1], Dexter Hadley[2], G. Zoltan Laszik[1]
[1]University of California San Francisco, San Francisco, CA, [2]University of Central Florida, Orlando, FL

Disclosures: Dmytro Lituiev: None; G. Zoltan Laszik: None

Background: Even though deep learning models can achieve super-human performance in many visual tasks, adoption of such models to digital pathology applications requires extensive annotations or techniques that can handle weak (per-slide) labels. Here we demonstrate application of a multiple instance learning (MIL) to classify transplant kidney needle biopsies (TxBx) with various grades of inflammation.

Design: Whole slide digital images of hematoxylin and eosin (H&E)-stained sections of TxBx with Banff normal (n=27), borderline (BL) change (n=24), and acute cellular rejection (ACR) (n=39) categories were analyzed. The cases were split into development (n=67) and test (n=23) sets in a stratified manner. A deep learning model pre-trained to recognize lymphoid aggregates (LA) on additional 44 TxBx annotated for LA was applied to extract 512 latent features and a LAs probability score from each 1024x1024 pixel patch digitized at 40x. For direct prediction, the fraction of patches per slide with predicted probability of LA presence exceeding 50% was calculated. Additionally, a MIL multilayer perceptron model was trained to predict diagnostic class using stochastic gradient descent on features extracted (512 per each slide patch). Training parameters were selected based on five-fold stratified cross-validation. Averaged predictions of the resulting models in the test set were used to calculate the area under the receiver-operator characteristic curve (AUC) and Mann-Whitney test p-value.

Results: The model pre-trained to detect presence of LA (Fig 1A) is able to differentiate biopsies with ACR from both normal (AUC=82%, p=0.031) and BL change (AUC=84%, p=0.029), but not ACR from BL change (AUC=49%, p=1.0). A MIL model (Fig 1B,C) can distinguish ACR from normal (AUC=96%, p=0.0021) and BL change (AUC = 97%, p=0.0029). The power to discriminate normal from BL change is improved but is not significantly different from random guess (AUC=69%, p=0.28).

Figure 1 - 1564
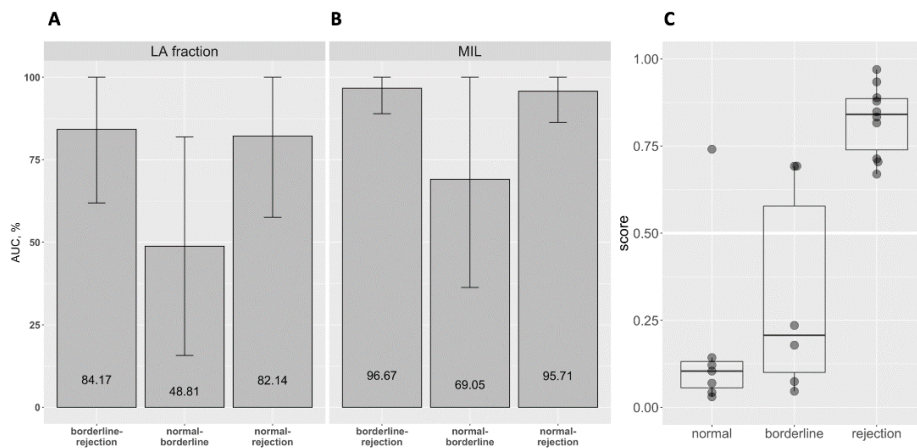
Conclusions: Biopsies with ACR can be distinguished from those with BL change and normal morphology by quantification of LA. The power to distinguish ACR, normal, and BL change is further improved by applying MIL based on automatically selected LA-specific imaging features. This demonstrates utility of multiple instance learning techniques to leverage slide-level labels to improve diagnostics of the transplant kidney rejection. MIL performance might be improved by incorporation of other features, end-to-end optimization, and by increasing the sample size.

1565 Segmentation of Breast Pathology Reports Using Deep Learning

Dmytro Lituiev[1], Maryam Panahiazar[1], Aaron Chin[1], Nolan Chen[2], Andrew Bishara[1], Dexter Hadley[3]
[1]University of California San Francisco, San Francisco, CA, [2]University of California San Francisco, Saratoga, CA, [3]University of Central Florida, Orlando, FL

Disclosures: Dmytro Lituiev: None; Nolan Chen: None; Andrew Bishara: None

Background: A significant amount of recent effort in digital histology has been dedicated to applying computer vision to learn pathological changes from digital slides. Design of such retrospective studies requires extensive curation by extracting diagnoses from electronic health records. Traditional manual curation of reports does not lend itself to scaling to large datasets, and previously used sentence-level bag-of-

words classification fails to capture hierarchical and semantic report structure, including attribution of findings to individual specimens and slides within complex reports, and may not be robust to mistyping and abbreviation. Here we apply modern deep-learning methods utilizing word-based tokenization and classification to segment pathology reports

Design: Final diagnosis section of breast pathology reports was manually segmented into six sub-sections: (1) specimen id, (2) organ or tissue source description, (3) pathological findings, (4) uninformative text, (5) slide references, and (6) punctuation. Batch-mode active learning training has been performed in three stages. Initially we segmented 4641 reports with regex and manually corrected 571 consecutively sampled reports. We trained a BiLSTM model on 400 reports, using 71 reports for validation, and 100 for testing. Resulting model was used to annotate remaining reports. Reports were ranked using maximum token entropy, and presented for manual correction. In next rounds, we added 515 and 778 reports to the pool of labelled reports and repeated the training-correction cycle.

Results: Maximum entropy per token decreased after two consecutive rounds from 1.006±0.494 and 1.080±0.304 to 0.352±0.443 bit, indicating improvement of the model. The final model achieved an area under receiver operating curve (ROC-AUC) of 99.998%.

Conclusions: In this study we demonstrate that a standard BiLSTM can successfully separate a breast pathology report into six different sections. The segmentation can be further improved by using character-level models and to extract specimen-level diagnoses from the segmented reports from large-scale datasets. These results allow for higher granularity of analysis of pathology notes by making it possible to assigning findings within a complex report to individual specimens and in some cases individual slides. This possibility will enable better curation and knowledge extraction from retrospective and potentially prospective pathology imaging data sets.

1566 Data-Efficient Histopathology Classification Using Contrastive Predictive Coding for Breast Cancer Diagnosis

Ming Lu[1], Richard Chen[1], Faisal Mahmood[1]
[1]Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Disclosures: Ming Lu: None; Richard Chen: None; Faisal Mahmood: None

Background: Convolutional Neural Network (CNN) models are touted for their ability to reduce inter-observer variability and achieve high accuracy when trained on sufficient annotated data. They have been widely applied to histology classification at the level of Regions of Interests (ROIs). However, training patch-level CNN classifiers requires the costly curation of labeled ROIs or exhaustive pixel-level annotation. The real-world need for patient stratification is often on gigapixel-sized histology slides with only slide-level labels available during training. For such tasks, given the scarcity of labeled histology data, naïve application of CNN-based Multiple Instance Learning (MIL) suffers from overfitting and poor performance. We propose to overcome this challenge using a two-stage data-efficient approach that combines self-supervised feature learning via Contrastive Predictive Coding (CPC) followed by regularized attention-based MIL.

Design: Given histology images, we first train a custom ResNet feature network on unlabeled 256 x 256 tissue patches using CPC, a self-supervised technique that uses a contrastive prediction task to learn high-level representations of the underlying data at the patch-level (bottom of Fig 1). We then apply a MIL neural network that uses attention to aggregate all patch-level features in an image and performs binary classification (top of Fig 1).

Results: We apply CPC + MIL to the binary classification of H&E stained breast cancer images in ICIAR 2018 BACH (400 labeled images of size 2048 × 1536). Using attention-MIL only as a baseline, we evaluate the performance gain by self-supervised learning via CPC. In each split, 100 images are randomly drawn for validation and the remaining 300 are used for training. We note that training CPC + MIL with hinge loss and additional regularization drastically outperforms the MIL baseline (See table), even with the CNN feature network frozen (without finetuning).

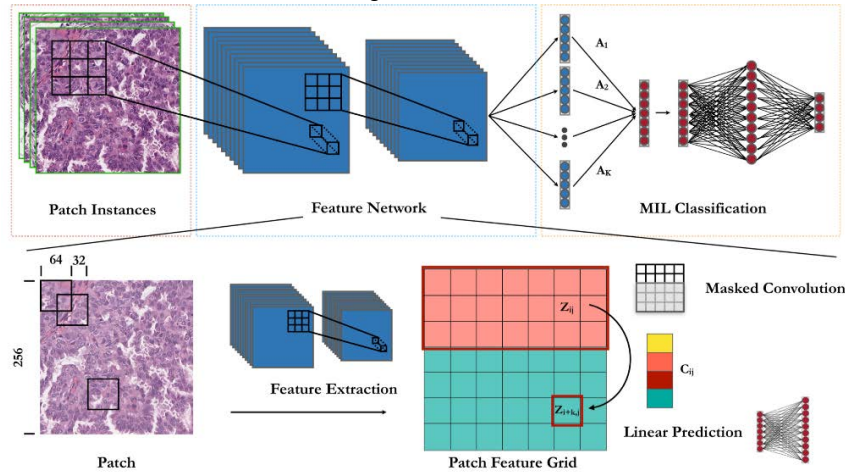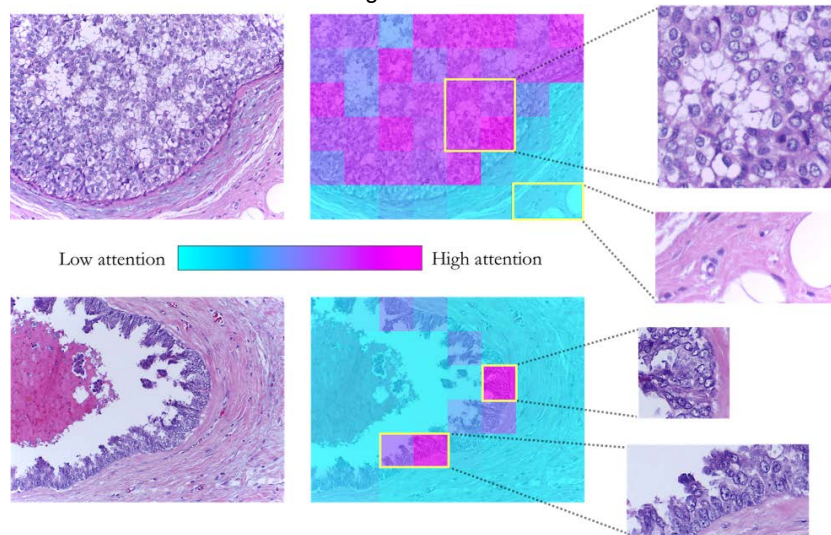| Results from random five-fold validation, ± standard deviation		
Method	Accuracy (%)	AUC ROC
MIL Only	62.6 ± 11.6	0.611 ± 0.186
MIL + CPC (Frozen Features)	90.6 ± 2.88	0.939 ± 0.024
MIL + CPC	**95.0 ± 2.65**	**0.968 ± 0.022**

Figure 1 - 1566

Figure 2 - 1566

Conclusions: We demonstrate that a deep semi-supervised approach using CPC + MIL combined with additional regularization can be effectively applied to the classification of breast cancer histology images even when MIL alone performs poorly due to overfitting on limited labeled data. Given the flexibility of our approach, we hope to scale experiments to whole slide images in the future and provide a data-efficient deep learning tool that can potentially serve as an assistive tool to reduce inter-observer variability and help pathologists improve diagnostic accuracy.

1567 Building Cytology Quality Assurance Dashboards from Daily Anatomic Pathology Laboratory Information System Database Extracts

Emilio Madrigal[1], Connor Barnhill[2], Brenda Sweeney[3], Vanda Torous[2], Martha Pitman[4], Long Le[2]
[1]Boston, MA, [2]Massachusetts General Hospital, Boston, MA, [3]Mass General Hospital, Boston, MA, [4]Massachusetts General Hospital, Harvard Medical School, Boston, MA

Disclosures: Emilio Madrigal: None; Connor Barnhill: None; Brenda Sweeney: None; Vanda Torous: None; Martha Pitman: None; Long Le: *Advisory Board Member*, ArcherDx; *Stock Ownership*, ArcherDx; Long Le: *Advisory Board Member*, ArcherDx; *Stock Ownership*, ArcherDx

Background: Cytology utilizes several quality assurance (QA) metrics that promote improved performance by cytopathologists (CP), cytotechnologists, and the lab as a whole. Well-established metrics include the ASCUS:SIL ratio, and the percentage of high-risk HPV positive ASCUS diagnoses (HRHPV+ ASCUS) in gynecologic cytology. More recently, AUS/FLUS thyroid rates have been recommended. To generate these four metrics for individual CPs and in aggregate for lab director review, several customized reports are exported from our AP-laboratory information system (LIS) as spreadsheets, followed by substantial manual data manipulation. Our current system

prohibits automation and lags by one month. We implemented an extract, transform, load (ETL) procedure to facilitate automation of these reports in real-time and to visualize them as a QA dashboard.

Design: Nightly, raw files are exported from the AP-LIS to a shared directory and later loaded into a Pandas DataFrame in Python. Rich Text Format is converted to plain text. The report for each case is split by specimen parts, and the pathologist's interpretation is parsed from the plain text (Figure 1). We use a keyword-based classification algorithm to categorize the parts based on The Bethesda System for Reporting (TBSR) Cervical Cytology and TBSR Thyroid Cytopathology diagnostic categories. ASCUS gynecologic cases are further categorized as positive, negative, or indeterminate for HRHPV+. The processed data is then loaded into a relational database for use in downstream open-source data visualization and business intelligence tools.

Results: Custom queries were written to generate anonymized and named lab administrator views, composed of a table, and 7 line and 4 bar graphs. Twelve access-controlled, personalized CP dashboards were created using 6 graphs, which are updated every 24 hours. The dashboards can be filtered by user-defined date ranges starting from January 1, 2012, to the current date.

Figure 1 - 1567
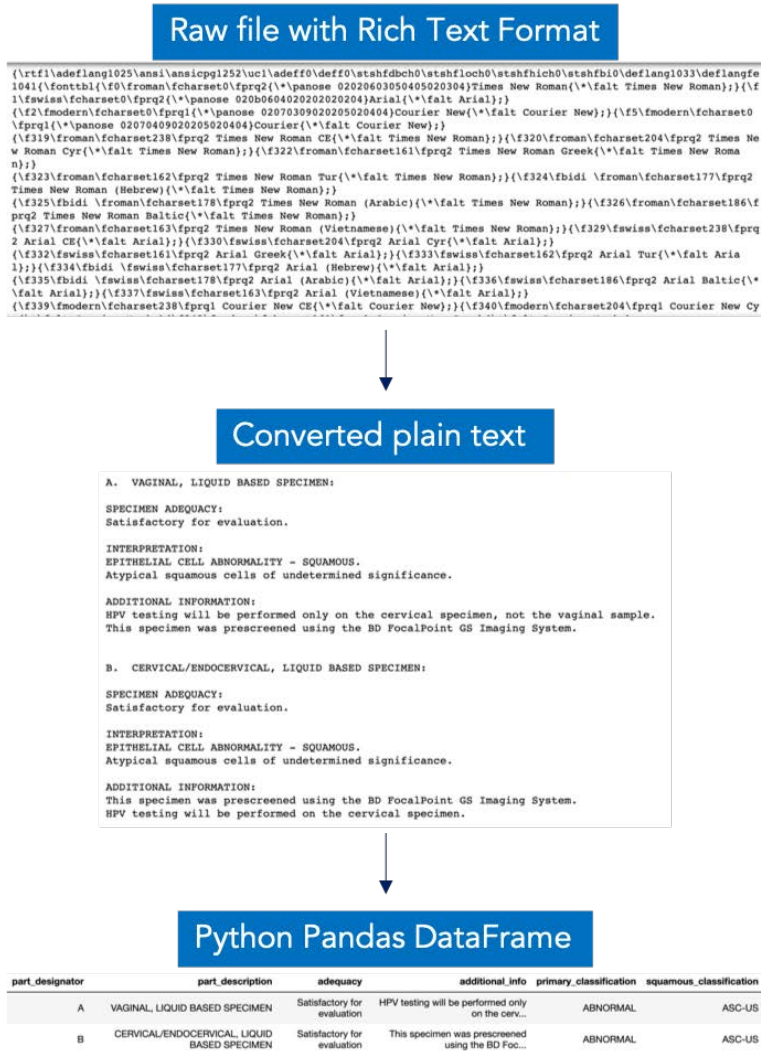
Conclusions: By using an ETL and an open-access data visualization tool, we automated the creation and reporting of personalized and aggregate cytology QA metrics for a user-friendly real-time QA dashboard. We plan to study the dashboard's impact on the performance of individual CPs and the overall lab.

1568   Holographic Visualization and Analysis of Flow Cytometry Data Using Mixed Reality Technology

Vladislav Makarenko[1], Pavel Terentiev[2], Jacob Bledsoe[3], Dmitry Korkin[2]
[1]UMass Memorial Health Care, Shrewsbury, MA, [2]Worcester Polytechnic Institute, Worcester, MA, [3]UMass Memorial Health Care, Worcester, MA

Disclosures: Vladislav Makarenko: None; Pavel Terentiev: None; Jacob Bledsoe: None; Dmitry Korkin: None

Background: Visualization of clinical flow cytometry (FC) data is becoming more challenging due to the increasing number of biomarkers used. The routine approach suggests consecutive analysis of two markers at a time using multiple 2D scatter plots. Cases with a small number of pathological cells (i.e. minimal residual disease) require particularly precise and clear visualization. Our goal is to provide proof-of-principle tool allowing simultaneous evaluation of 3 markers using a faster and more intuitive exploration of FC data to aid in diagnosis.

Design: We chose Mixed Reality (MR) as a visualization technology for our tool. The technology allows data representation as a 3D hologram immersed in a physical environment. MR data is mapped in a real 3D space and provides options to interact with the data including hand gestures, voice commands, and gaze, as well as conventional keyboard-mouse interfaces.

The developed tool consists of 3 components (Fig.1): (i) MR device to visualize and manipulate the data; (ii) graphic user interface (GUI) - gateway to manage data exchange between the storage and MR devices using custom protocol to synchronize different MR devices; (iii) database to store pre-processed FC datasets.

Publicly available raw pre-gated FC datasets (Clinical Flow Wiki) from patients with hairy cell leukemia and acute lymphoblastic leukemia were used to test the model.

Results: FC data were visualized in 3D space using Microsoft Hololens device. Cell data points were grouped based on their proximity and each group was represented as a spherical object (Fig.2), its size based on the number of related cells. The tool allows switching between different CD markers, selection of groups and representation of related data points. This approach visually improved gating of the cells and allowed distinct clustering of abnormal cell populations by using a combination of three markers for each tested entity (Fig.2).

Fig.1 Design of the developed FC data visualization system

Fig.2 Screen shot of FC data layout (CD20, CD11c and CD103

Figure 1 - 1568   Figure 2 - 1568
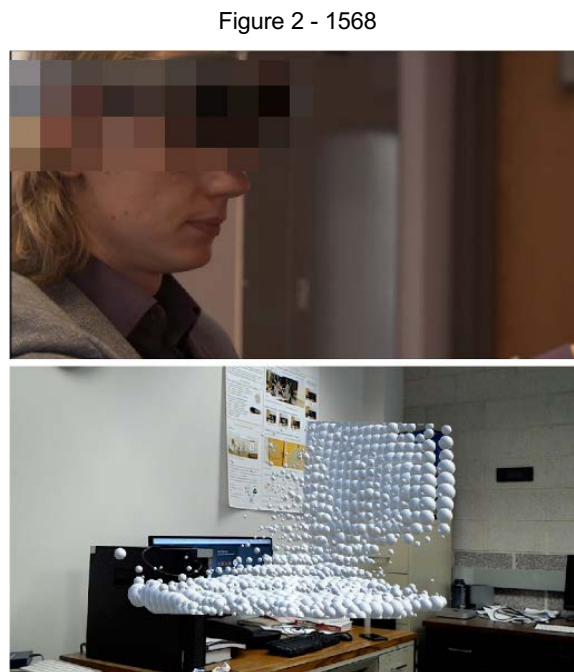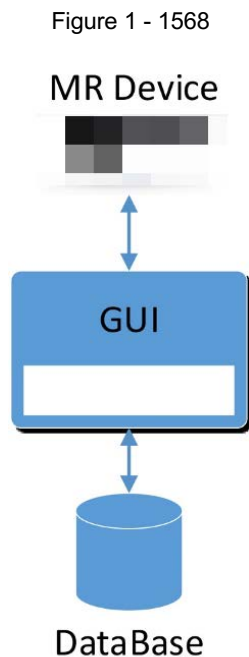
Conclusions: We developed a unique tool for holographic visualization of FC data aimed to facilitate diagnosis of complex cases. Further studies are needed to examine the performance of the tool and for comparison to current methods of FC analysis. Integration of clinical

information, principal component analysis (PCA) and dimensionality reduction algorithms as well as consideration of new ways to layout 3D data may be useful to further improve its diagnostic power.

1569   AE-Checks as a Critical Factor in Trial Design, Outcomes Research, and Utilization Studies

Hetal Marble[1], Alexander Farahani[1], Connor Barnhill[1], Jochen Lennerz[2]
[1]Massachusetts General Hospital, Boston, MA, [2]Massachusetts General Hospital, Harvard Medical School, Boston, MA

Disclosures: Hetal Marble: None; Alexander Farahani: None; Connor Barnhill: None; Jochen Lennerz: None

Background: BACKGROUND: The (A)nticipated somatic mutation frequency may differ from the clinically (E)ncountered prevalence. Publicly-available mutation databases provide reference frequencies; and their use is emphasized by current NIH/NCI data-sharing policies. In contrast, routine clinical testing is held to different standards and the encountered mutation prevalence may differ. Here, we assessed A-E-differences as a check function to explore selected tumor types.

Design: DESIGN: For E, we used all samples from our ongoing, multi-year clinical test practice and separated the top 20 most frequently mutated genes from glioblastoma (GBM), non-small cell lung carcinoma (NSCLC), acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), breast cancer, and all cancers with an anatomic site annotation of "lung". For genotyping, we target enrichment via anchored multiplex PCR in combination with next-generation sequencing on an Illumina NextSeq platform. For A we used COSMIC (mut/wt) and compared the prevalence with E by gene and cancer-type (mut/wt) using chi-square statistics (for mutational profile). For visualization, we plotted AE differences and color-coded positive (red) vs. negative (blue) deviations.

Results: RESULTS: All encountered mutation profiles differed significantly from COSMIC (P-values<<0.01). In most cancer types, we observed lower rates of mutations on average, with notable exceptions in GBM and NSCLC. In GBM we found a higher incidence of TERT and in NSCLC, higher incidence of PIK3CA, MET, APC, RB1, STK11, respectively. Examination of the anatomic site-based mutation frequencies showed significant deviations (P-range: 0.05-0.01) that are likely due to the composite nature of the cancers tested from this site (e.g. EGFR-mutation rate is significantly lower than in the lung cancer group P<0.05).

Figure 1 - 1569

GBM		NSCLC		AML		MDS		All Lung		Breast	
APC	1%	APC	-2%	ASXL1	4%	ASXL1	2%	APC	2%	AKT1	2%
ARID1A	1%	ARID1A	1%	BCOR	1%	BCOR	-1%	ARID1A	4%	APC	2%
ATM	1%	ATM	4%	DNMT3A	13%	CUX1	-1%	ATM	4%	ARID1A	4%
ATRX	7%	ATRX	-1%	EZH2	1%	DNMT3A	1%	ATRX	4%	ATM	2%
BRAF	3%	BRAF	0%	FLT3	21%	EZH2	2%	BRAF	1%	ATRX	5%
BRCA1	2%	BRCA1	-1%	GATA2	3%	IDH1	0%	BRCA1	2%	BRCA1	2%
CIC	1%	BRCA2	-2%	IDH1	4%	IDH2	1%	BRCA2	2%	BRCA2	2%
DDX3X	0%	CDKN2A	0%	IDH2	5%	JAK2	7%	CDKN2A	3%	CDH1	7%
EGFR	12%	EGFR	16%	JAK2	1%	KRAS	-1%	EGFR	21%	CIC	0%
IDH1	9%	KEAP1	4%	KRAS	2%	NRAS	1%	KEAP1	9%	ESR1	6%
MSH6	2%	KRAS	8%	NPM1	28%	PHF6	0%	KRAS	9%	KRAS	0%
NF1	7%	MET	1%	NRAS	9%	RUNX1	0%	MET	2%	MAP3K1	7%
PIK3CA	6%	NF1	3%	RUNX1	4%	SETBP1	1%	NF1	6%	NF1	5%
PIK3R1	4%	PIK3CA	-2%	SF3B1	1%	SF3B1	14%	PIK3CA	1%	PIK3CA	8%
PTEN	17%	PTEN	0%	SRSF2	1%	SRSF2	-2%	PTEN	2%	PIK3R1	1%
RB1	6%	RB1	0%	STAG2	2%	STAG2	-1%	RB1	2%	PTEN	3%
STAG2	3%	SMARCA4	1%	TET2	6%	TET2	4%	SMAD4	2%	RB1	3%
TERT	-6%	STK11	-2%	TP53	0%	TP53	-4%	SMARCA4	5%	SMAD4	1%
TP53	20%	TERT	-1%	U2AF1	1%	U2AF1	2%	STK11	5%	SMARCA4	2%
TSC1	1%	TP53	4%	WT1	7%	ZRSR2	2%	TP53	26%	TP53	12%

Figure 1: AE differences plotted by cancer type and gene and color-coded positive (red) vs. negative (blue) deviations.

Conclusions: The mutation prevalence in COSMIC does not reflect the frequency in our clinical practice with a focus on end-stage cancer patients. While the underlying reason warrants further examination, the difference between anticipated and encountered mutation frequencies (AE check) is paramount for trial design, clinical outcomes research, and genotyping utilization studies.

1570   Education Through 280 Characters

Ammar Matloob[1], Saeed Asiry[1], Samer Khader[2]

[1]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY, [2]Montefiore Medical Center, Albert Einstein College of Medicine, Irvington, NY

Background: Medical education through social media has gained publicity in recent years. Twitter specifically, has become a major social media portal used by many pathologists around the globe to share interesting cases and papers. Our aim of this study is to measure the influence of an educational Twitter account on global cytopathology education, based on Twitter analytics data for Dr. Samer Khader account (@SamKhader). Dr. Khader is the director of the cytopathology fellowship program at Montefiore Medical Center/Albert Einstein College of Medicine.

Design: Using a Twitter analytics traffic data, we analyzed the number of tweet impressions, profile visit, tweets mentions and followers during 90 days period, starting from June, 2019 - September, 2019. Also, we analysed the engagement activity, link clicks, re-tweets and likes per tweets.

Results: Data relevance from our individual Twitter analytics account showed an increase in educational cytopathology tweets activity by 235.0% over the 90 days period leading to an increase in both tweet impressions by 953.6% (94.3K) and profile visits by 697.7% (3,680). Although, this account existed for a while it was relatively inactive until June 2019, when we started to tweet an educational material regularly reaching a rate of at least one educational tweet per day. It was also noticed that as the educational tweets activity increased by 235.0%, the number of followers increased by 138%, especially among residents and fellows globally. As Twitter activity increased with the educational cases, The engagement activity increased to 3.1% (Figure 1). The link clicks was 340 clicks (average of 4 link clicks/day). Retweets of the interesting cases were 1.2K (average of 13 retweets/day). Likes were around 2.5K (average of 28 likes/day). Replies were around 67 (average of one reply/ day) with questions or comments about the cases plus we received more questions in messages. Our tweeting activity earned 2.3K impressions/day and 203.9K impressions during the 90 days period (Figure 2).

Figure 1 - 1570

Figure 2 - 1570

Figure 1: Tweet activity, showing engagement rate, link clinks, re-tweets and likes activity during the 90 day period.
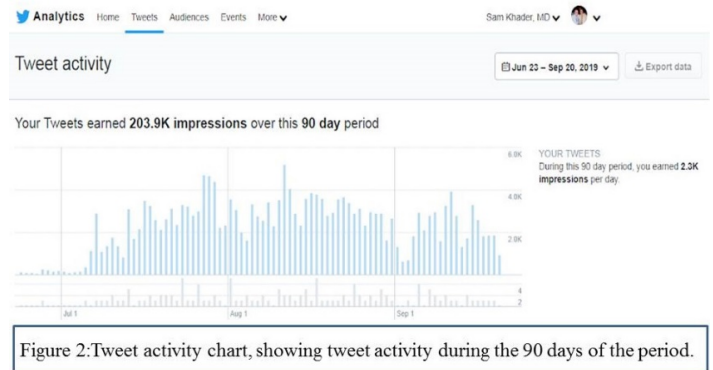
Figure 2: Tweet activity chart, showing tweet activity during the 90 days of the period.

Conclusions: The fast spread and growth of this single Twitter account in a short period of time demonstrate the usefulness of Twitter and social media in general and the role they could play in global pathology and medical education.

1571 Text Mining as a Tool to Advance Analytics in a Surgical Pathology Service

Chelsea Mehr[1], Amrom Obstfeld[2]
[1]Dallas, PA, [2]The Children's Hospital of Philadelphia, Philadelphia, PA

Background: Pathology reports contain a tremendous amount of data, however harnessing this data for secondary use in research, operational, or quality work can be challenging as the data is stored as free text which is not easily subjected to statistical analysis. While synoptic reporting and other forms of structured data capture are useful in this regard, these tools are labor intensive for users, must be maintained for long periods of time, and can only be used when the required data is known a priori. We report here the successful use of the statistical programming language R, a free application created for statistics and data analysis, as a tool for mining data from surgical pathology reports.

Design: As a proof of principle, we elected to mine cervical biopsy surgical pathology reports for the presence and type of dysplasia as well as the presence or absence of associated HPV related changes. Regular expressions, a method for analyzing text using pattern matching, which is well supported by R, was used to extract discrete data from the reports. The analysis involved first extracting final diagnosis sections from full reports, parsing individual specimens, and isolating individual diagnoses, and finally identifying the type of dysplasia present and HPV related terms.

Results: We analyzed 36,137 cervical biopsy diagnoses over an 11 year period. The report format changed twice over this period of time, resulting in three different patterns to be matched over the study period. The regular expressions were therefore constructed to be sufficiently flexible while still maintaining high levels of specificity. Likewise the precise semantics used to describe HPV related changes (e.g., "HPV", "human papilloma virus", "condyloma", etc.) had to be accommodated with highly precise patterns. Based on these data we were able to demonstrate that the use of HPV related terms vastly decreased over the study period, likely correlating with the introduction of the LAST criteria.

Conclusions: Pathology groups which require discrete data often must rely on structured data capture or highly inefficient manual reviews. Our investigation highlights the great potential of using statistical programming languages such as R to extract data from surgical pathology reports. While this specific evaluation involved cervical biopsies, the process could be applied to any data encoded in free-text.

1572 A Machine Learning Based Algorithm for Automatic Scoring of PD-L1 SP263 in Non-Small-Cell Lung Carcinoma (NSCLC)

Mohammad Saleh Miri[1], Kien Nguyen[2], Keith Earley[3], Stefanie Selck[1], Suhas Patil[4], Karel Zuiderveld[1], Guadalupe Manriquez[5], Dorothy Hayden[6], Bharathi Vennapusa[7], Christoph Guetter[2]
[1]Roche Diagnostics, Santa Clara, CA, [2]Roche Tissue Diagnostics, Santa Clara, CA, [3]Roche Diagnostics, Indianapolis, IN, [4]Santa Clara, CA, [5]Roche Tissue Diagnostics, Indianapolis, IN, [6]Roche Diagnostics, Tucson, AZ, [7]Roche Diagnostics Corporation, Indianapolis, IN

Disclosures: Mohammad Saleh Miri: None; Kien Nguyen: None; Stefanie Selck: *Employee*, Working for Roche Tissue Diagnostics; Dorothy Hayden: *Employee*, Roche Molecular Solutions; Bharathi Vennapusa: *Employee*, Roche; Christoph Guetter: *Employee*, Roche Tissue Diagnostics

Background: Currently, immunohistochemical analysis of programmed death ligand-1 (PD-L1) expression is the most widely used biomarker for determining patient suitability for anti-PD-1/PD-L1 treatment. The US Food and Drug Administration and the European Medicines Agency require a pathologist to determine the percentage of tumor cells from a tumor sample immunohistochemically stained with an approved PD-L1 antibody. Studies on PD-L1 scoring have found that there is variability between pathologists for scoring clinically significant cutoffs of both 1% and 50%. Digital pathology (DP) with image analysis (IA) has the potential to increase pathologists' precision in PD-L1 scoring. Hence, we developed a machine learning based IA algorithm for scoring whole slide non-small-cell lung carcinoma (NSCLC) at 1% and 50% cutoffs.

Design: A cohort of 88 NSCLC samples were stained for PD-L1 IHC (SP263) and used for training the classification models. The scans were segregated into 389 smaller Fields-of-View (FOVs) and annotated at the cell level by pathologists into three classes of positive (TC+), negative (TC-) tumor cells, and other cells. The cell level ground-truth (GT) collection resulted in a training set with TC+: 29975, TC-: 17433, and Other: 30437 samples. Separately, a testing cohort was utilized to assess the performance of algorithm in computing the whole slide PD-L1 SP263 score by comparing to pathologist consensus scores. The test set consists of 156 cases (adeno/squamous cell carcinoma; resections; core-needle biopsies) along with ground truth manual microscope reads (GT) determined by a consensus panel of three independent reading pathologists. Data was analyzed using pair-wise overall percent agreement rates (OPA) derived from assay threshold categorical bins (1% and 50%) and/or as continuous scaled quantitative scores.

Results: In cross-validation of the IA algorithm after training, we found that the cell-level accuracy for tumor detection, stained tumor (TC+) and non-stained tumor (TC-) cells were 88%, 93%, and 89% respectively. We found slide-level concordance across the entire range (0%-100%), with the R^2 value=0.87. OPA values for the 1% and 50% cutoff are 91% and 93%, respectively. The correlation plot between algorithm and GT score is shown in Figure 1 (each dot represents one case).

Figure 1 - 1572

Tumor Staining Percentage

$y = 0.9021x + 3.9772$
$R^2 = 0.8652$
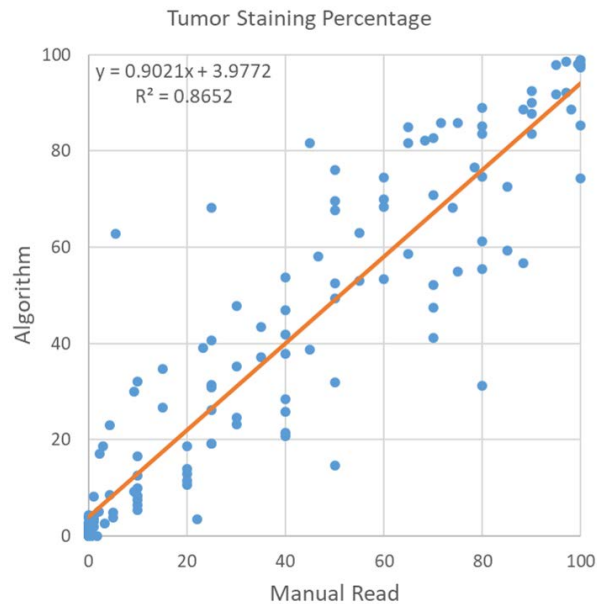
Conclusions: Overall the ability of the IA algorithm to automatically, robustly, and accurately score PD-L1 in tumor cells in NSCLC is promising, particularly as it is shown on a wide range of patient samples.

1573 Deep Learning Models for Reliable Detection of the Intraductal Lesions of the Breast

Shachi Mittal[1], Andre Balla[2], Rohit Bhargava[3]
[1]University of Illinois at Urbana-Champaign, Champaign, IL, [2]University of Illinois, Chicago, IL, [3]University of Illinois at Urbana-Champaign, Urbana, IL

Disclosures: Shachi Mittal: None

Background: Current standards for separating benign intraductal lesions from ductal carcinoma in-situ (DCIS) rely on visual inspection of morphometric features. There are no immunohistochemical markers for stratifying these patients for a precise diagnosis. This can often be time consuming and lead to diagnostic discordance. Deep learning models can enable the identification of global and local textural patterns that are indicative of an altered epithelial profile. This study focuses on utilizing deep neural networks on H&E stained slides of large surgical specimens to separate benign lesions (normal, usual hyperplasia and atypical hyperplasia) from DCIS. This will address the long-standing need for precise triaging in a quantitative manner helping the pathologist to improve diagnostic accuracy.

Design: Representative areas were chosen from about 50 surgical specimens encompassing normal breast tissue, usual hyperplasia, atypical hyperplasia with and without columnar change and ductal carcinoma in-situ. A total of 20708 images were curated for training, 8876 for calibration and 8833 for independent validation. Two pretrained neural networks, AlexNet and VGG16 were modified and retrained for feature extraction followed by classification. The learning rate for the both the models was 0.0001. A stochastic gradient loss function was used to optimize both the models on the calibration data. After inspecting the model accuracies on the calibration and independent validation (different set of patients), the classification model was then projected on a large patient area to validate the spatial performance of the model.

Results: Both the models gave high accuracy on the calibration set to distinguish the benign intraductal lesions from DCIS. However, VGG16 performed much better on the independent validation set. Therefore, we selected the later for the classification task with an overall accuracy of ~90%. A list of accuracies for both the models can be seen in Table 1. The high spatial performance of the model is evident from Fig. 1 where the benign areas are precisely separated from the adjacent DCIS.

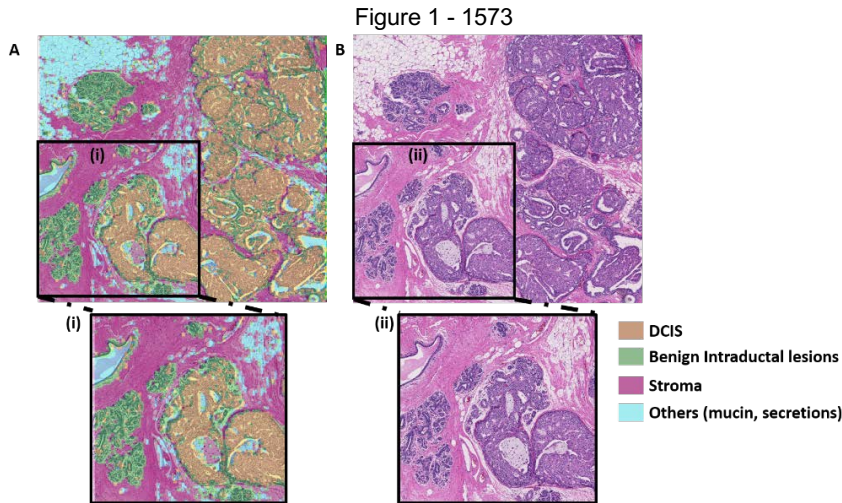| Model accuracy for the different deep learning networks used to investigate intraductal lesions		
Models	**Calibration Set Accuracy (%)**	**Validation Set Accuracy (%)**
AlexNet	93.23	76.38
VGG16*	91.28	89.57
*Both the models used the same number of training, calibration and validation examples. * Chosen model*		

Figure 1 - 1573

Figure 1. Classification of the H&E stained image using deep learning. A.H&E stained image with areas of benign and DCIS intraductal lesions identified and a zoomed in view in (i). B. The corresponding stained image with the zoomed in view in (ii).

Conclusions: Precise stratification of the intraductal lesions can help overcome the problem of under- and over diagnoses. This could also assist pathologists in borderline cases or differential diagnosis such as high-grade hyperplasia and low-grade DCIS. This pipeline, without needing extensive supplies or human input, can transform care in both high resource and low resource settings.

1574 Grading Follicular Lymphomas Using Augmented Human Intelligence

Jeffrey Mohlman[1], Jessica Kohan[2], David Ng[3]
[1]University of Utah, Salt Lake City, UT, [2]Salt Lake City, UT, [3]Seattle, WA

Disclosures: Jeffrey Mohlman: None; Jessica Kohan: None

Background: Follicular lymphoma (FL) is a common type of non-Hodgkin lymphoma that is morphologically graded to predict prognosis and direct therapy. Grading FL involves counting or estimating the number of centroblasts within follicles and 10 high power fields must be evaluated within different follicles. Reproducibility of this manual FL grading method is not high and our objective is to demonstrate the utility of applying convolutional neural networks (CNNs) to selected image fields for the purposes of grading FL.

Design: Hematoxylin and eosin stained slides from 35 cases of follicular lymphoma, including 15 high-grade (HG)(grade 3A) and 20 low-grade (LG)(grade 1-2/3), were digitally scanned using an Aperio AT2 (Leica Biosystems) slide scanner at 400X magnification. Digital files were then uploaded and viewed on QuPath (version 0.2.0-m2). Representative follicles and/or large sections from each slide were digitally circled by a hematopathologist. These areas were then extracted into 244x244 pixel tiles. Overall, 17,886 tiles were generated from 35 total cases. 14 random cases (6 HG, 8 LG) were reserved for testing, accounting for 5,573 tiles, while the remaining tiles and cases were used for training (90%) and unbiased evaluation of the model fit (validation, 10%). DenseNet-121, initialized with ImageNet weights was fit on the training set. Area under the curve (AUC) of the receiver operating characteristic (ROC) curve, precision, recall and f-1 scores were used to evaluate performance.

Results: Overall, tile classification accuracy was 94% (figure 1). 12/14 (86%) cases were correctly classified using the CNN with an AUC of the ROC curve of 0.89 (figure 2). 5/6 HG and 7/8 LG cases were correctly classified. Precision, recall and f1-scores were all 0.83 and 0.88 for HG and LG cases, respectively.
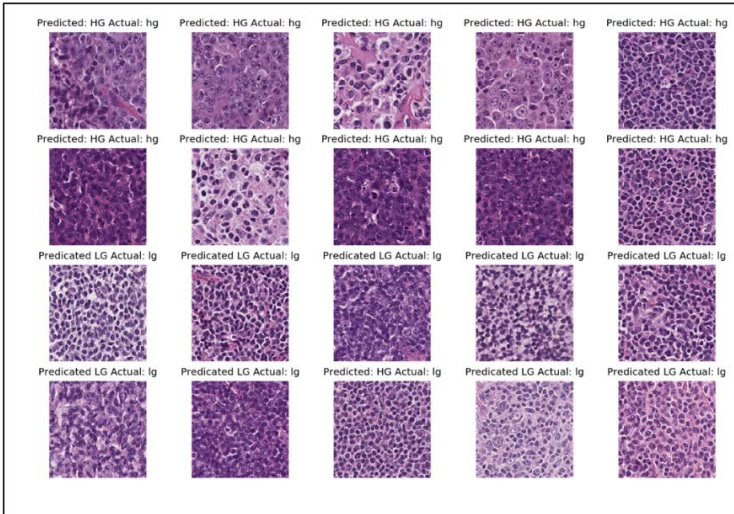
Figure 1 - 1574

Figure 2 - 1574

Figure 1: Examples of CNN performance on individual tiles randomly selected from a variety of the testing cases. LG, low grade prediction; lg, actual low grade case; HG, high grade prediction; hg, actual high grade case.

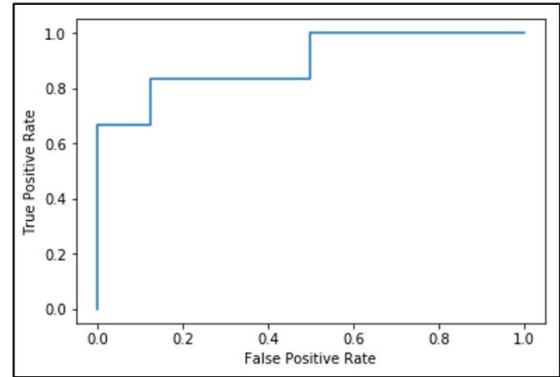Figure 2: Area under the curve (AUC) of the receiver operating characteristic (ROC) curve was 0.89.

Conclusions: CNNs hold potential for grading FLs, but more research is needed to improve the performance. The next step currently underway includes adding additional training and testing cases to evaluate if the performance characteristics can be improved. Well-validated training sets, along with clinical outcome and cost assessments, are needed prior to moving this type of assay into clinical practice.

1575 Universally Unique Identifiers (UUIDs) with URL- and Filename-Safe Base64 Encoding Provides Globally Unique Naming Conventions for Sample Identifiers

Andrew Norgan[1], Hieu La[1], Charlene Brown[1], Karla Kopp[1], Steven Hart[1], Taofic Mounajjed[1], Thomas Flotte[1]
[1]Mayo Clinic, Rochester, MN

Disclosures: Andrew Norgan: None; Hieu La: None; Charlene Brown: None; Karla Kopp: None; Steven Hart: None; Taofic Mounajjed: None; Thomas Flotte: None

Background: The development of machine learning algorithms in medicine has intensified the need for large datasets within an institution and in larger collaborative networks. In addition, the need to protect patient information may require anonymization of data. Many investigators and institutions use similar naming conventions for their specimens which results in data with the same identifiers. Universally Unique Identifiers (UUIDs; also known as Globally Unique Identifiers [GUIDs]) are an industry standard for creating random 128-bit identifiers whose specifications are maintained by the Internet Engineering Task Force (ietf.org). When UUIDs are used for sample identifiers, there is no need for a centralized registry for creating the identifiers as the probability of generation of a conflicting identifier (a collision) is negligible. Any investigator, anywhere in the world, can create an identifier using standard software that can be contributed to the dataset without concern that there will be samples with duplicate identifiers in the database.

Design: A custom software system was programmed in Python to create file names for whole slide images for our machine learning group. The program generates industry standard 128 bit UUIDs consisting of 32 characters and 4 hyphens (e.g., 00000000-0000-0000-0000-000000000000 [the nil UUID]). UUIDs are not optimally formatted as file names in some operating systems, and therefore, URL- and filename-safe Base64 encoding of the UUIDs was implemented, which reduces the total character count to 23, without loss of uniqueness (e.g., 0b45IfmXPUuQ2MvYDjg07A). Removable slide file labels are created with 2D barcodes and human readable file names. The labels are place on the slides prior to scanning to support anonymization of the slides and are used to create the filenames.

Results: To date, this system has been used to anonymize 10,000 research virtual slides without any problems or filename collisions.

Conclusions: URL-safe base64 encoded UUIDs provides a naming convention for samples that permits a decentralized approach to creating globally unique identifiers for collaborations involving multiple investigators, sites, and/or locations, as well as a general approach to anonymization of data artifacts in medicine.

1576 Development of Highly Sensitive Screening System for Metastasis Cancer Cells with Chemiluminescence Immunohistochemistry Technology

Hiroyuki Nozaka[1], Tamotsu Sugai[2]
[1]Hirosaki University, Hirosaki, Aomori, Japan, [2]Iwate Medical University, Morioka, Iwate, Japan

Disclosures: Hiroyuki Nozaka: None; Tamotsu Sugai: None

Background: Chemiluminescence method is one of highly sensitive technology for protein or gene expression assay in immunology. Chemiluminescent Enzyme Immunoassay (CLEIA) is more sensitive than fluorescent enzyme immunoassays, and it is used as a diagnostic technology for the detection of infectious diseases. On the contrast, pathological diagnosis also requires highly sensitive antigen detection technology, but chemiluminescence technology has not been applied to immunohistochemistry. The aim of this study is the development of a new highly sensitive screening system for metastasis cancer cells and the application of chemiluminescence method in histopathological diagnosis.

1. System configuration: The scanning system consisted of an automatic X/Y/Z stage, an area scan CMOS camera, and a photomultiplier tube (PMT). Stage motor control was performed by a dedicated sequencer, and image capture and photon measurement control were performed by NI LabVIEW. Image data was linked to photon count data, and total photon counts and photon counts per unit area were calculated. The overall system configuration was shown in Figure1, and the optical measurement unit was shown in Figure2.
2. System evaluation:Lymph nodes with metastasis or non-metastasis were collected from 50 patients diagnosed with primary breast cancer. IHC with Cytokeratin cocktail antibody was performed on FFPE tissue sections. Visualization was performed with DAB (Conventional IHC) / Fluorescent dye (FIHC) / Chemiluminescent dye "CDP-star" (CIHC).

Results: The FIHC and CIHC method were superior to the conventional method in both sensitivity and specificity. The CIHC method showed a significant difference in the total photon counts between non-metastasis, micro-metastasis, and metastasis. The S/N ratio of CIHC method showed higher than FIHC method.

Figure 1 - 1576

Figure 2 - 1576
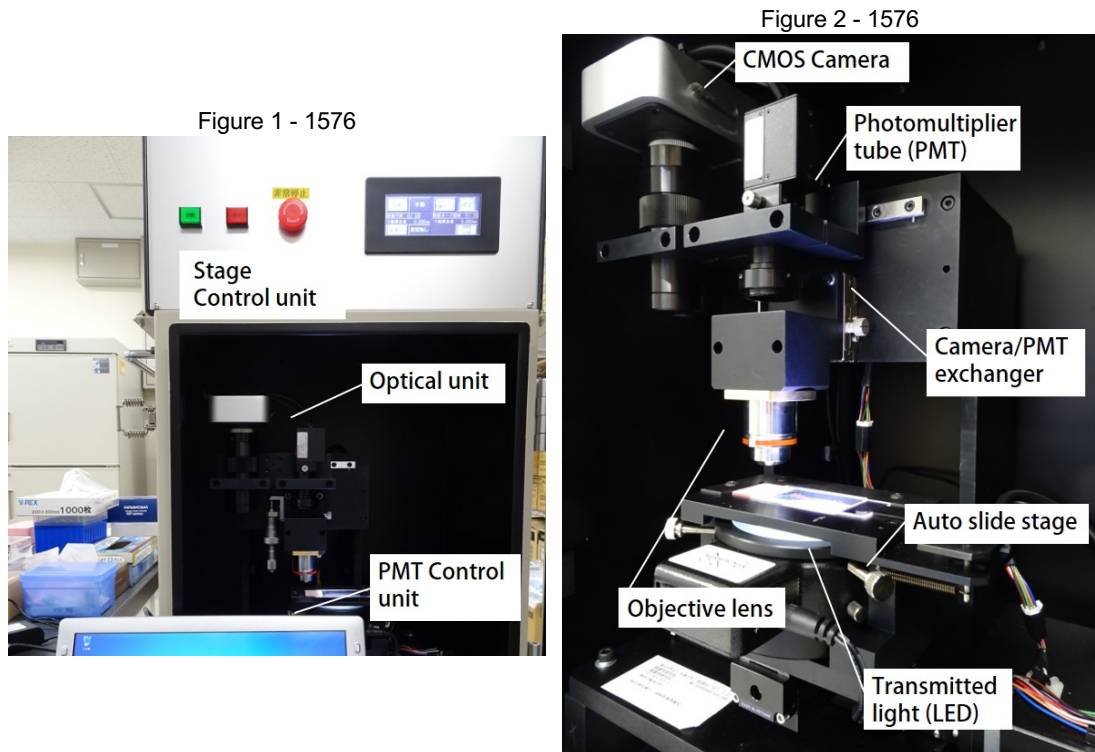
Conclusions: It was suggested that the CIHC method has excellent sensitivity and specificity, and it is useful for screening of metastasis cancer cells in lymph node. Attenuation of chemiluminescence is slow and luminescence is maintained for up to 24 hours. Furthermore, it does not require an excitation light source or bandpass filter for measurement. Therefore, the system has a simple structure and can be easily downsized. We believe that chemiluminescence technology contributes to the advancement of immunohistochemistry.

1577 Three-Dimensional Vessel Extraction in Whole Block Imaging Using Deep Neural Networks

Takashi Ohnishi[1], Alexei Teplov[1], Noboru Kawata[1], Benjamin Stueben[1], Kareem Ibrahim[1], Peter Ntiamoah[2], Canan Firat[1], Hideaki Haneishi[3], Meera Hameed[1], Jinru Shia[1], Yukako Yagi[1]
[1]Memorial Sloan Kettering Cancer Center, New York, NY, [2]New York, NY, [3]Chiba University, Chiba, Chiba, Japan

Disclosures: Takashi Ohnishi: None; Alexei Teplov: None; Noboru Kawata: None; Benjamin Stueben: None; Kareem Ibrahim: None; Peter Ntiamoah: None; Canan Firat: None; Hideaki Haneishi: None; Meera Hameed: None; Jinru Shia: None; Yukako Yagi: None

Background: In order to more clearly demonstrate the structure of cancerous tissue, components such as vessel structure, depth of invasion, etc. must be examined. Although conventional pathology images only visualize a thin cross section of tissue; micro computed tomography (microCT) allows us to fully analyze the three-dimensional (3D) characteristics of a neoplasm non-invasively. However, there is a dire need for additional analysis methods. In this study, we propose an application of analysis flow with deep neural networks and to proof of the concept, we focused to extract vessel regions from whole block images (WBIs) acquired by microCT.

Design: Figure 1 details a schematic of our analytical procedure, which uses WBI by microCT (the custom-built Nikon Metrology). 10 FFPE colorectal tissue blocks were scanned with the microCT and taken through our analysis method. In this study, we implemented a vessel extraction function using deep neural networks. The data was reconstructed into 3D volumetric images for analysis. We adopted a VNet as a basic network structure and two VNets were subsequently cascaded. A vessel region that was an output of the first VNet was enhanced by a Frangi filter and concatenated with original image, which was then inputted into the second VNet. A Frangi filter was also applied to the output of the second VNet and binarized by discriminant analysis method. Both networks trained only the vessels with 2–15 voxels diameter.

Results: All extractions were successfully conducted. Figure 2 shows an example of the cropped and enlarged images of WBI with extracted vessel region. White triangles on the WBI indicates vessel regions whose size we chose to detect. The shape of extracted vessel regions was smooth, and we could confirm vessel pathways. While it took over 8 hours to manually extract vessel regions, constructed deep neural networks could complete the process in about 10 minutes for the cropped WBI with 200x200x120 voxels using four units of GeForce GTX 1080 Ti.

Figure 1 - 1577

Figure 2 - 1577
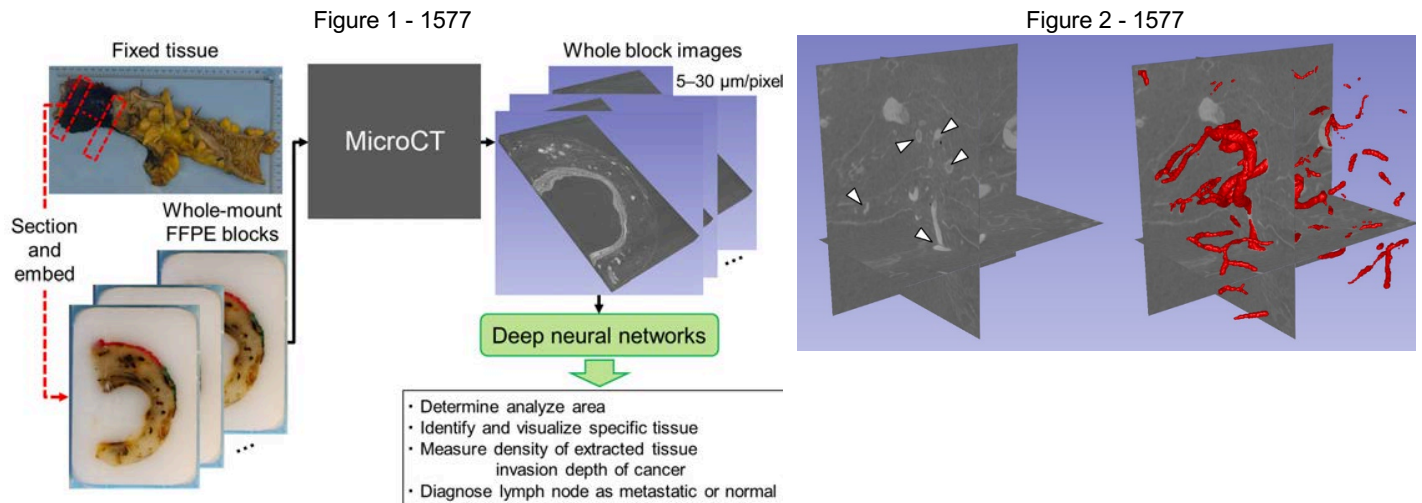
Conclusions: We have developed deep neural networks to analyze WBIs of neoplastic tissue. Our method could extract vessel regions from WBI and all extracted results were promising for future analysis. The next step is to improve the networks to measure and extract other structures such as lymph nodes and tumor lesions. In future, our analysis method might easily understand patterns of invasion and help to guide the therapeutic decisions.

1578 Large Scale Deployment of Whole Slide Imaging for Anatomical Pathology Workflow Improvement and Primary Diagnosis and Consult Reviews

Anil Parwani[1], Wendy Frankel[2], Martha Yearsley[2], Trina Shanks[3], Zaibo Li[2], David Kellough[4], Mark Lloyd[4], Xiaoyan Cui[1], Vidya Arole[1], Adrian Suarez[1], Konstantin Shilo[1], Lynn Schoenfield[2], Wei Chen[2], Gary Tozbikian[2], Ogechukwu Eze[2], Jae-Hoon Chung[5], Jose Otero[6], Abberly Lott Limbach[2], Peter Kobalka[7], Diana Thomas[6], Paul Wakely[2]

[1]The Ohio State University, Columbus, OH, [2]The Ohio State University Wexner Medical Center, Columbus, OH, [3]Inspirata, Inc., New Albany, OH, [4]Inspirata, Inc., Tampa, FL, [5]The Ohio State University Wexner Medical Center, Upper Arlington, OH, [6]The Ohio State University Wexner Medical Center, Dublin, OH, [7]The Ohio State University, Upper Arlington, OH

Disclosures: Anil Parwani: None; Wendy Frankel: None; Martha Yearsley: None; Trina Shanks: *Employee*, Inspirata; Zaibo Li: None; David Kellough: *Employee*, Inspirata, Inc; Mark Lloyd: *Employee*, Inspirata, INc.; Xiaoyan Cui: None; Vidya Arole: None; Konstantin Shilo: None; Lynn Schoenfield: None; Wei Chen: None; Gary Tozbikian: None; Ogechukwu Eze: None; Jae-Hoon Chung: None; Abberly Lott Limbach: None; Peter Kobalka: None; Diana Thomas: None

Background: Automated whole slide imaging (WSI) scanners are now rendering diagnostic quality, high-resolution images of entire glass slides and combining these images with innovative digital pathology and artificial intelligence tools that are making it possible to integrate imaging into all aspects of pathology workflow including anatomical, clinical and molecular pathology.

Design: WSI technology was implemented in the Department of Pathology at the Ohio State University Wexner Medical Center (OSUWMC) for primary diagnostics and storage of digitized slides. All glass slides were scanned by FDA-approved Philips Ultra Fast Scanner (UFS). At a time, 300 slides were loaded on the scanners and scanned at 40X magnification with 0.25 um/pixel resolution and variable bit depth. The files were stored in an iSyntax format using the Philips IntelliSite Pathology Solution (PIPS). These uncompressed files immediately became available for pathologists to access on a shared drive. Back-up compressed images were stored in a Ohio Supercomputer at a remote site. Complete laboratory information system (EPIC Beaker, Sunquest Copath) integration was accomplished as part of this project.

Results: Since May 2017, FDA-approved WSI technology has been used at the OSUWMC. To date, 1,260,919 slides from H&E, special stains and immunostains have been scanned and stored in duplicate. These are from a total of 117,038 patient surgical pathology accessions. At least 18 pathologists are using digital slides for primary diagnosis while up to 25 are utilizing digital slides for consults and/or tumor boards. The scan time per slide was 1-2 minutes dependent on the size of the tissue on the slide. The average file size ranged from 1-3 gigabytes and currently data is not compressed. Some of the challenges that were encountered included occasional failure. The successful integration with laboratory information system enabled pathologists to access images directly from their sign-out list (Figure 1)

Figure 1 - 1578
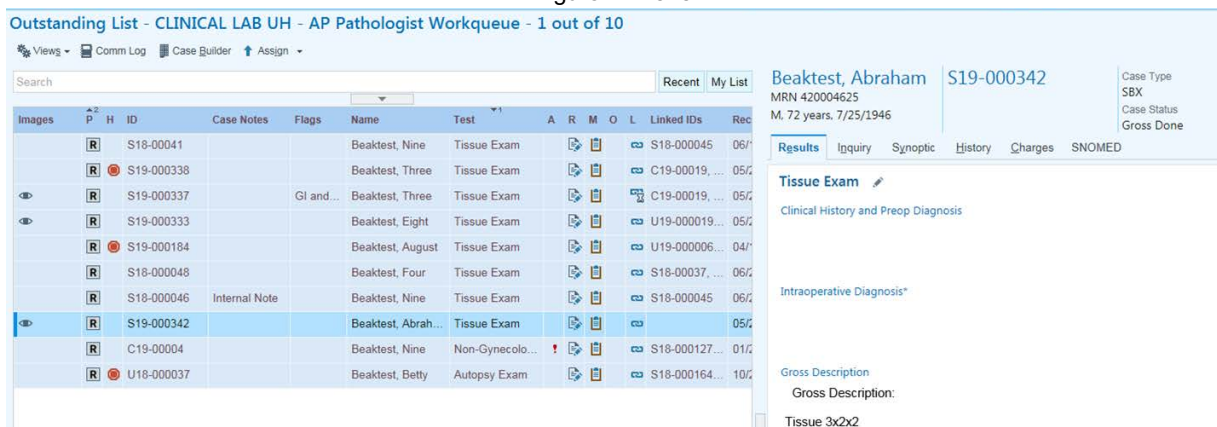
Conclusions: Our experience shows that high-throughput WSI can be successfully carried out at the level of major academic medical center such as OSUWMC and can be incorporated into the workflow of a busy academic pathology department. Future work is focused on more improvement in the workflow processes and utilizing these images for artificial intelligence algorithm development and innovate the digital sign-out process for pathologists.

1579 Oncopathologist: A Checklist Based Web-App for Generating a Cancer Report from CAP Cancer Protocols

Ashish Patil[1], Henley Jennifer[1], David Otohinoyi[2], James Cotelingam[3], Diana Veillon[4], Menchu Ong[5], Marjorie Fowler[6], Eric Wei[7], Nestor Dela Cruz[1]

[1]LSUHSC-Shreveport, Shreveport, LA, [2]All Saints University College of Medicine, Belaire, Kingston, Saint Vincent and the Grenadines, [3]Louisiana State University Shreveport, Shreveport, LA, [4]LSU Health Shreveport - Pathology, Shreveport, LA, [5]Department of Pathology at LSU Shreveport, Shreveport, LA, [6]LSU Health Sciences Center, Shreveport, LA, [7]LSU Health Sciences Center Shreveport, Shreveport, LA

Disclosures: Ashish Patil: None; Henley Jennifer: None; David Otohinoyi: None; James Cotelingam: None; Nestor Dela Cruz: None

Background: Nearly 88 cancer reporting protocol templates are available on the CAP site. Filling a cancer template and manually generating a report is time consuming and there are chances of committing errors. Use of electronic checklist based application expedites rapid and accurate generation of cancer reports from the College of American Pathologist (CAP) cancer protocol, aids in diagnostic process, standardizes reporting of data and facilitates exchanges of exchange of information between the Pathologist and the Clinician. We have designed a web-based application called "Oncopathologist" which acts as an electronic checklist for concise reports for the physicians, and this can meet the necessary requirements of accreditation.

Design: The web app can be accessed at (https://oncopathologist.com/) and is designed to enable users to easily organize cancer and biomarkers reporting protocol using the templates provided by CAP. The app is scripted using HTML and JAVA script, specifically, HTML5 and CSS 3. The script was organized on php framework Laravel for easy manipulation and is hosted on whogohost on server, wgh16.whogohost.com.

Results: The Oncopathologist automatically formats the cancer templates to generate cancer reports. The report can be exported in the Microsoft word format, from which the output can be copied and pasted manually into the laboratory information system (LIS) softwares like Copath, Novopath, EPIC, etc. Our app does not use PHI (protected health information) to generate the report output and our webserver does not store or process the PHI. The average turn around time to generate a cancer report from CAP Cancer template manually is around 35 to 45 minutes, but using our checklist based web-app one can generate the cancer report in 10 to 15 minutes.

Conclusions: The Oncopathologist is a free web-app that can generate cancer reports in concise, clear, readable format from cancer templates and can help the health care professionals to improve the efficiency of cancer reporting.

1580 Automated Identification of Glomeruli in Renal Biopsies by Machine Learning on Digital Pathology Specimens: A Multi-Institutional Study Highlights the Need for Cross-Institutional Algorithm Training to Mitigate Variability in Histologic Material

Jason Pettus[1], Maxwell Smith[2], David Wilbur[3], Lynn Cornell[4]
[1]Lebanon, NH, [2]Mayo Clinic Arizona, Scottsdale, AZ, [3]Corista, LLC, Concord, MA, [4]Rochester, MN

Disclosures: Jason Pettus: None; Maxwell Smith: None; David Wilbur: *Employee*, Corista, LLC; Lynn Cornell: None

Background: Machine learning for digital pathology specimens can improve pathologists' efficiency, accuracy, and reproducibility via prescreening with automated identification of particular features. Studies using uniform histologic material from a single institution have shown promise. Generalized application of machine learning requires validation on slides prepared by different institutions, and it is unknown how well algorithms may perform in this setting. In this study, we used machine learning to identify glomeruli on renal biopsies as a test platform to determine performance across single and multi-institutional environments.

Design: Randomly selected, adequately sampled renal core biopsy cases (71) consisting of 4 stains (H&E, trichrome, silver, PAS) from 3 institutions were digitized using either Aperio AT Turbo or AT2. Non-sclerotic glomeruli were manually annotated by 3 renal pathologists using a digital tool. Cases were divided into training (n=52) and validation (n=19) cohorts. An algorithm was trained to develop 3 convolutional neural network (CNN) models which tested case cohorts for the detection of glomeruli intra- and inter-institutionally. Raw CNN search data from each of the 4 slides per case were merged into composite fields of view (FOV) containing putative glomeruli. The sensitivity and specificity of glomerulus detection (compared to annotated truth), and FOV area, were calculated for each model/cohort.

Results: The sensitivity, specificity, and area of FOV results for each of the training/validation sets are shown in table 1. Figure 1 shows an example of algorithm output on a prescreened digital slide. Detection of glomeruli was best by algorithms trained on material from one institution and tested on similar material from the same institution (e.g. site 1a vs. 1a). Detection sensitivity degrades when training and test material originate from different sites (e.g. site 1abc vs. 2). Training using a combined set of digital slides from both institutions improves performance but not to the level of a single institution. In the best model, the FOV area decreases, meaning the precision of detection is best.

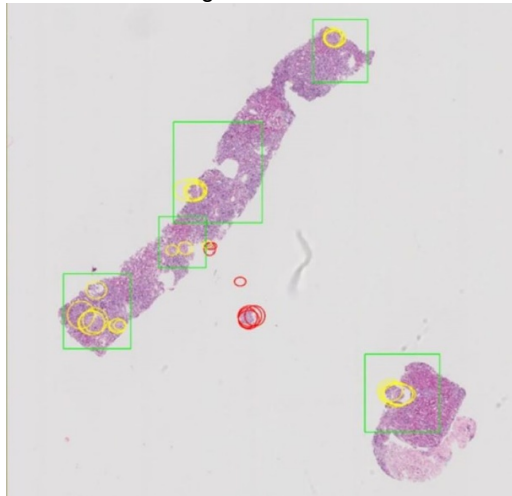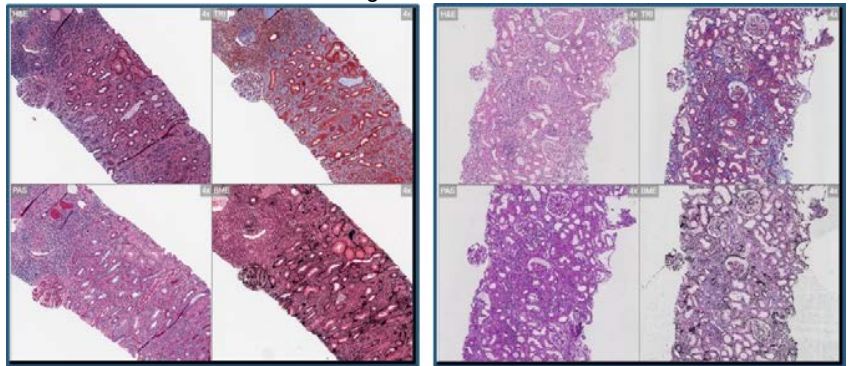| Model - testing type/cohort	Sensitivity (%)	Specificity (%)	FOV area vs. Slide area
1 - intra-institutional/1	92	89	0.008
1 - intra-institutional/1a	90	98	0.016
2 - intra-institutional/1b	93	86	0.011
2 - inter-institutional/1c to 2	77	97	0.010
3 - intra- & inter-institutional/1abc2	89	92	0.012

Figure 1 - 1580   Figure 2 - 1580

Conclusions: There is substantial inter-institutional difference in histology preparation with respect to machine learning (figure 2), which likely accounts for significant algorithm performance differences between the models. Our data highlight the need for diverse training sets for the development of generalizable machine learning histology algorithms.

1581 Data-Driven Quality Improvement in the Modern Histology Lab

Robert Seifert[1], Vektra Casler[2], Nada Al Qaysi[2], Leah Williams[3], Sherri Flax[1], Srikar Chamala[2]
[1]Department of Pathology, Immunology and Lab Medicine, University of Florida, Gainesville, FL, [2]University of Florida, Gainesville, FL, [3]UFHealth Shands Hospital, Gainesville, FL

Disclosures: Robert Seifert: None; Vektra Casler: None; Nada Al Qaysi: None; Leah Williams: None; Sherri Flax: None; Srikar Chamala: None

Background: While many laboratory information systems (LIS) can natively generate basic quality metrics (e.g., TAT or volume reports), data is often limited in detail. The discrete data elements captured by modern LIS offer an opportunity for more robust process evaluation. Thorough utilization of such data is challenging for lab directors as it requires not only intimate knowledge of one's lab operations but also an understanding of nuanced, "big data" informatics approaches. These advanced data-mining techniques reveal important elements such as specific TAT breakdowns or time-dependent spikes in volume which may be "buried" amid other LIS data points.

Design: Tissue containers, blocks, and slides are barcoded, and nearly every significant event each item experiences is tracked in the EPIC Beaker AP LIS (Epic Systems Corporation, Verona, WI). Histology staff collaborated with our pathology informatics team to generate dynamic data visualization reports using SAP BusinessObjects (SAP, Walldorf, Germany) and custom Python scripts to discover meaningful, actionable information. Challenges examined during implementation include:

1. Early detection of TAT breakdowns
2. Enumerating time-dependent volume spikes
3. Combining volume and financial data for pathology trainees

Results: A breakdown in TAT occurred for prostate biopsies from a specific contributor (Figure 1). This breakdown would have been missed by conventional TAT reports (Figure 1, dark line) which consider total volume. We implemented a dynamic TAT report with the capability to filter multiple variables (e.g. surgical specialty, specimen source, contributing clinic). This led to process changes which prevented delay of care.

Our immunohistochemistry section offers same-day TAT for tasks ordered before 10 AM. While overall volume remained stable, a spike in ancillary test ordering was noted at just before 10 AM (Figure 2) for which current staffing and instrumentation were inadequate. This data was presented to hospital administration to obtain funding for additional staff and instrumentation.

A spike in block volume was noted on cases grossed only by trainees. This data was combined with reimbursement data (Table 1) to visualize the cost of submission of unnecessary blocks. This proved to be a useful tool for trainee education.

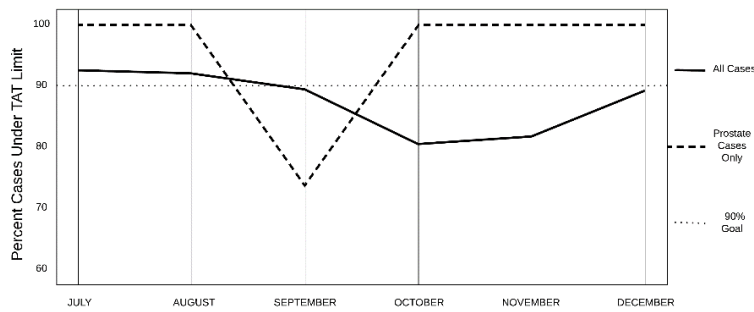| Table 1. Net collection per H&E slide produced for a given 88305 collection, including material cost, pathologist labor and technologist labor. Parenthetical dollar values indicate loss.	
# of H&E Slides per 88305	Net Collection
5	$165.82
10	$121.17
15	$82.92
20	$43.42
25	$12.82
30	($25.43)

Figure 1 - 1581

Figure 1. Percent of slides meeting TAT overall compared to percent meeting TAT for just prostate biopsy specimens submitted from a specific contributing insitution. A drop was identified in this population in September due to a process failure.
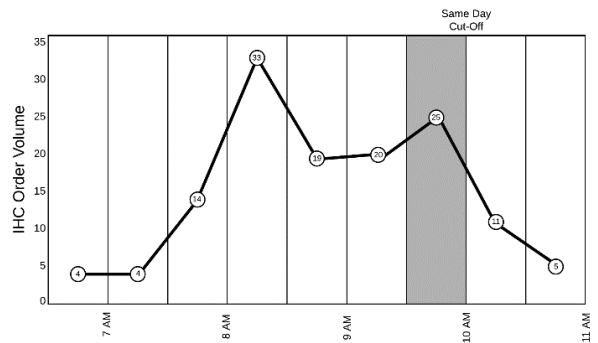
Figure 2 - 1581

Figure 2. Over 25 immunohistochemistry tasks ordered within 15 mins of same-day cutoff, 10 AM, exceeding the cumulative capacity of available staff and instrumentation.

Conclusions: Native LIS reporting functionality can be leveraged to improve multiple areas within the histology lab. We present a series of tailored approaches for comprehensive histology lab improvement.

1582   Prognostic Significance of Tumor Mutation Burden in Endometrial Endometrioid Carcinoma
Maryam Shahi[1], Alexander Baras[2]
[1]Johns Hopkins Medical Institutions, Baltimore, MD, [2]Baltimore, MD

Disclosures: Maryam Shahi: None; Alexander Baras: None

Background: Aggregate genomic measures such as tumor mutation burden (TMB), microsatellite instability (MSI), and other mutation signatures have received much attention recently as potential biomarkers predictive of specific therapeutic responses. In endometrial endometrioid carcinoma (EEC) MSI is commonly observed from germline or sporadic etiologies affecting mismatch repair genes. Response to checkpoint inhibitors (CPI) has been associated with MSI high status in a variety of human malignancies, including EEC. More recently, TMB has emerged as a putative biomarker predictive of responsiveness to immuno-oncology (IO) therapy, specifically CPI. However, the prognostic significance of TMB as well as optimal threshold(s) in ECC needs further evaluation. In this study we investigated the prognostic significance of TMB and compared that with the performance of MSI status in EEC.

Design: Genomic data of EEC cases in The Cancer Genome Atlas (TCGA) database (Nature 2013) were studied (n=307). In this cohort, MSI status was determined by 5 marker panel (MSS/MSI-L n=185 (60.3%); MSI-H n=121 (39.4%), indeterminate n=1 (0.3%)). TMB was determined by the number of non-synonymous mutations per Mbps of coding sequence (ie exome). We selected TMB cut-points by examining inflection points of the log-rank (LR) test statistics across the spectrum of TMB observed in this cohort. Overall survival (OS) available for this TCGA cohort was used for these analyses and outcomes were shown using Kaplan-Meier plots.

Results: Based on the two peaks of the LR test statistic observed across the spectrum of TMB observed in this cohort, we divided the cohort into three groups of low, intermediate and high TMB categories using cut points of 1.44 and 6.89 mutations per MBps. OS was not not significantly different based on MSI status. Surprisingly, low and high TMB strata exhibited significantly better prognosis as compared to the intermediate TMB group.

Figure 1 - 1582
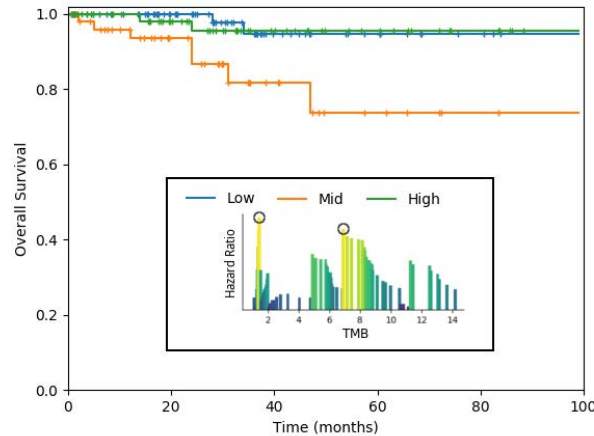
Conclusions: In the context of EEC, MSI appears well suited as a predictive biomarker for response to IO since in this non IO cohort, MSI status did significantly stratify outcomes; as such, any stratification observed in IO cohorts can be properly attributed to predicting responsiveness. In contrast, in EEC TMB exhibited a significant and non-linear ("inverse Goldilocks phenomenon") association to outcome in this non IO treated cohort. Thus, interpretation of TMB as a predictive biomarker in IO treated cohorts needs to be assessed relative to the apparent prognostic signature described above.

1583   Manual vs Digital Scoring of Ki67 in Breast Cancer; A Validation Study

Ahsan Siddiqi[1], Xochiquetzal Geiger[1], Tracy Majewicz[1], Aziza Nassar[1]
[1]Mayo Clinic, Jacksonville, FL

Disclosures: Ahsan Siddiqi: None; Xochiquetzal Geiger: None; Tracy Majewicz: None; Aziza Nassar: None

Background: Recent data suggest that ki67 level >10-14% defines a high-risk group in terms of prognosis in breast cancer. A recent meta-analysis reported a shorter OS for patients with positive Ki67 (estimated HR of 1.73; 95% CI: 1.37-2.17). Following the POETIC trial, it becomes evident that Ki67 levels at baseline and after two weeks following aromatase inhibitor therapy can predict patients who are most likely to have increased risk of recurrence and hence require additional chemotherapy treatment. Recurrence risk was lower (4.5%) for patients with a low baseline and at 2-weeks ki67 levels (<10%); on the other hand the risk was higher (19.6%) for patients with a high baseline and at two weeks Ki67 levels (≥ 10%). In the current study, we sought to validate Ki67 at our institution using Aperio digital image analysis and compare the results to the manual method in breast cancer patients.

Design: We extracted 87 patients with a breast cancer diagnosis from our pathology database (57 luminal subtype A and 30 Luminal subtype B). We performed Ki67 immunostaining (monoclonal MIB-1 clone) in paraffin-embedded whole tissue sections. Two pathologists reviewed and annotated the ki67 using manual counting and two histotechnologists performed the image analysis quantitation for hotspots using Aperio. All demographic and outcome data were collected. Descriptive statistical analysis was performed and the Bland and Altman methodology was used to evaluate the agreement between manual and digital image analysis and a paired t-test was performed to see if there were any significant differences between the reviewers.

Results: The mean age of the cohort was 66.8 years (range 47.0 – 89.0). The tumor grades using the Nottingham grading system is as follows: grade 1 (29/ 33.3%); grade 2 (49/ 56.3%) and grade 3 (9/10.4%). There were 71 (81.6%) invasive ductal carcinomas; 13 (14.9%) invasive lobular carcinomas and 3 (3.4%) mixed carcinomas. See table and figures 1 and 2.

| **Table:** Ki67 Scores	
	Total (N=87)
Digital Result	
Mean (SD)	9.4 (7.8)
Median (Range)	6.7 (0.3-33.3)
Manual Score Ki-67 Breast-Pathologist 1 (%)	
Mean (SD)	9.2 (9.3)
Median (Range)	5.0 (0.0-40.0)
Manual Score Ki-67 Breast- Pathologist 2 (%)	
Mean (SD)	6.2 (7.7)
Median (Range)	2.0 (0.0-35.0)
Difference between Manual Score Ki-67-Pathologist 1 and Manual Score Ki-67-Pathologist 2	
Mean (SD)	3.1 (6.1)
Median (Range)	1.0 (-17.0-24.0)
Difference between Digital Result and Manual Score Ki-67-Pathologist 1	
Mean (SD)	0.1 (5.6)
Median (Range)	0.4 (-21.0-16.4)
Difference between Digital Result and Manual Score Ki-67-Pathologist 2	
Mean (SD)	3.2 (4.3)
Median (Range)	2.0 (-7.1-18.3)

Figure 1 - 1583

Figure 1. Bland and Altman Plot

Digital Result vs. Manual Score Ki-67 Breast

(Pathologist 1)

Figure 2 - 1583

Figure 2. Bland and Altman Plot

Digital Result vs. Manual Score Ki-67 Breast
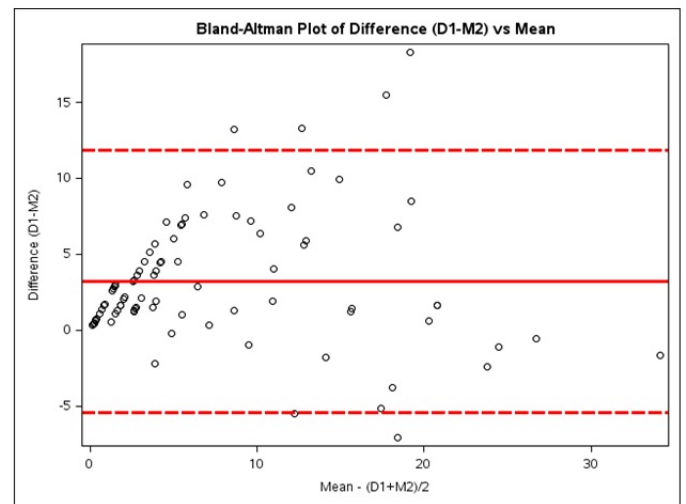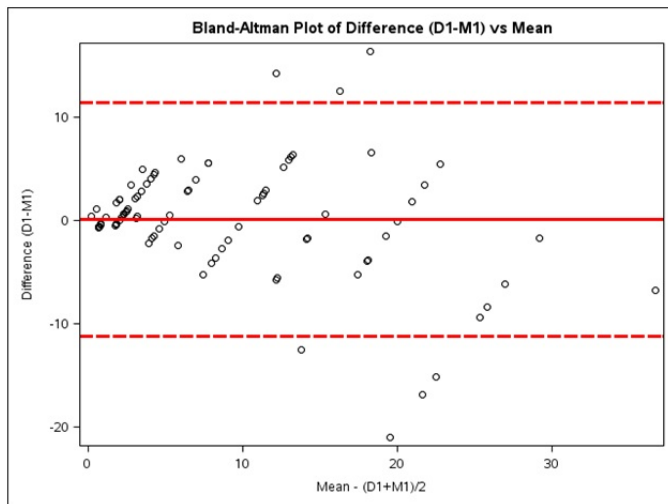
(Pathologist 2)

Paired t-test

Difference	DF	t Value	PR > |t|
D1 − M1	86	0.21	0.8351

Paired t-test

Difference	DF	t Value	PR > |t|
D1 − M2	86	6.89	<.0001

Conclusions: A significant difference is noted between pathologist 1 and 2 in regards to manual scoring. The manual and digital scores between pathologist 1 and Aperio demonstrated excellent agreement versus poor agreement between pathologist 2 and Aperio. This validates the use of digital scoring in providing more reliable ki67 proliferation labeling index to negate the effects of interobserver variability among pathologists. Furthermore, it is a better tool to assess prognosis for risk recurrence in patients who are pretreated with endocrine therapy.

1584 'PRPL_code Synopsis' is an Accurate and Efficient Synoptic Report Parsing Program for Pathology Database-Based Research

Aryeh Stock[1], Noam Harpaz[2], Huaibin Mabel Ko[1]
[1]Icahn School of Medicine at Mount Sinai, New York, NY, [2]Mount Sinai Medical Center, New York, NY

Disclosures: Aryeh Stock: None; Noam Harpaz: None; Huaibin Mabel Ko: None

Background: Performing large population studies requiring review of pathology reports is a laborious task requiring hours of manual review of thousands of pathology reports. Efforts have been made to standardize the content of pathology reports such as Systematized Nomenclature of Medicine (SNOMED) and synoptic reporting. Despite these efforts, large-scale reviews of historical data require vast amounts of time and resources to accomplish. We present here the first use of our resident-built narrative report text parsing software (PRPL_code Synopsis) to perform a retrospective review of over 1,200 pathology reports.

Design: PRPL_code Synopsis is a program written in Python 3.7.4 (pandas 0.25.1, xarray 0.12.3) and R 3.6.1 (tidyverse 1.2.1) which can parse text from narrative pathology reports and organize it as a spreadsheet to produce analysis-ready data. A Structured Query Language (SQL) search for all primary colorectal carcinomas was performed on the laboratory information system (LIS) of a major tertiary care center. Cases were divided based on the presence or absence of a synoptic report. Key regions of text were identified, isolated and assigned categorical scores representing distinct regions of the colon.

Results: Manual review of 10% of the reports revealed a concurrence of >95% between the software and manual review by a pathologist. Discrepant cases were manually reviewed. Of the discrepant cases, approximately 40% were rescored in favor of the software.

Conclusions: PRPL_code Synopsis has been demonstrated to represent a viable tool for precision data extraction with a high accuracy and low error rate. Further development will likely improve its performance when presented with a broader range of content and larger datasets.

1585 Multi-Scale Deep Learning for Small Object Detection in Whole Slide Images: Tumor Bud Detection in Colorectal Cancer without IHC Staining

Chen-Yu Sun[1], Weiguo Liu[2], Scott Doyle[3]
[1]University at Buffalo, SUNY, Buffalo, NY, [2]East Amherst, NY, [3]University at Buffalo, Buffalo, NY

Disclosures: Chen-Yu Sun: None; Weiguo Liu: None; Scott Doyle: None

Background: High tumor budding in colorectal cancer is a significant risk factor for nodal involvement in adenocarcinoma arising in polyp and it is associated with adverse prognosis. Tumor cells can be stained with immumohistochemistry, but it is not as common as hematoxylin and eosin (H&E). Manual tumor bud detection on H&E stained images is laborious and inconsistent among pathologists. This study aims to automate the process using deep learning to provide accurate and consistent detection. Tumor bud is defined as single cells or small clusters of less than five cells at the advancing front of the tumor, which are impossible to detect in a whole slide image (WSI) without high magnification. Current criteria for evaluating budding is selecting "hotspots" on a H&E section at low scale and counting the number of buds at high scale. We mimic this approach using multi-scale deep learning.

Design: Our methodology employs three classifiers, each trained on a specific task. The first tier is a patch-based convolutional neural network (CNN) which classifies tiles of a digital WSI at low magnification (4x) into three tissue regions: tumor mucosa, stroma, and normal glands. The second tier focuses on regions of stroma (where tumor buds are likely to appear) selected by the first tier, and segments out nuclei from the background at high scale (40x). The third tier classifies cells as either tumor buds or non-tumor cells on small patches (40x) centered on segmented nuclei regions from the second tier. We used 18 WSIs from The Cancer Genome Atlas for training and 10 holdout WSIs for validation.

Results: The fist tier classifier has an average 0.85 sensitivity in identifying the three tissue regions, with a sensitivity of 0.89 for the stromal tissue (which we focus on for remaining tiers). The second tier segmentation classifier achieves an average dice coefficient of 0.60. The third tier classifier achieves 0.9 sensitivity and 0.92 specificity in classifying tumor bud nuclei.

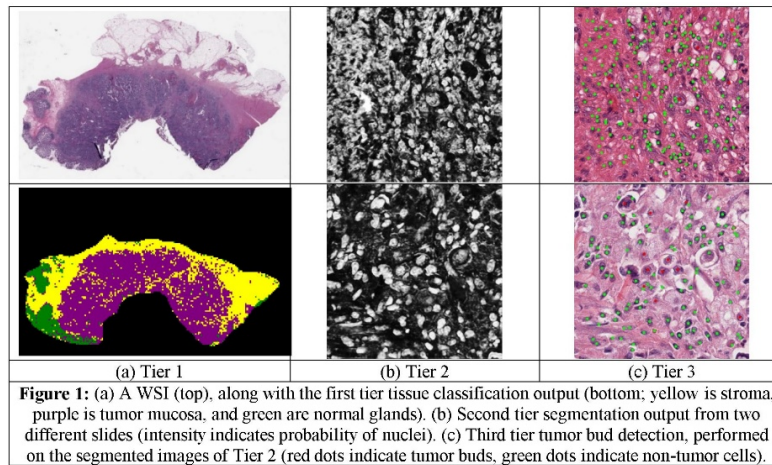|	gland	other	tumor	other	tumor bud
sensitivity	0.798±0.078	0.886±0.032	0.867±0.016	0.920±0.035	0.903±0.070
specificity	0.953±0.020	0.899±0.008	0.923±0.055	0.903±0.070	0.920±0.035
(a) Tier 1				(b) Tier 3	
Table 1: Sensitivity and specificity of the classes in tiers 1 and 3 (the two "classification" tasks).					

Figure 1 - 1585

Figure 1: (a) A WSI (top), along with the first tier tissue classification output (bottom; yellow is stroma, purple is tumor mucosa, and green are normal glands). (b) Second tier segmentation output from two different slides (intensity indicates probability of nuclei). (c) Third tier tumor bud detection, performed on the segmented images of Tier 2 (red dots indicate tumor buds, green dots indicate non-tumor cells).

Figure 2 - 1585

Figure 2: Training (blue) and validation (orange) curves for (a) first, (b) second, and (c) third tier classifiers. Shown are accuracy (top row) and loss (bottom row); the loss for each tier is calculated as cross-entropy loss, a standard metric used to drive deep learning training.
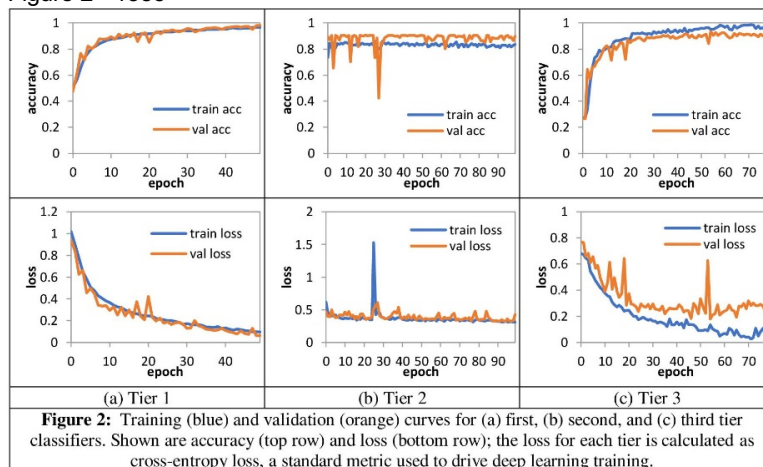
Conclusions: The multi-tier design efficiently processes WSIs while improving the specificity of tumor bud detection. The first tier reduces the search space and eliminates confounding regions such as epithelial and mucosal layers. The second tier focuses on a relatively simple task of segmenting nuclei, which allows the third tier to classify tumor from non-tumor cells with over 0.9 sensitivity and specificity. This multi-scale framework can be applied to any pathology task where small objects in WSI are critical for accurate diagnosis.

1586 Prostate TMA Classification Using Weakly Supervised Labels via Graph Convolutional Networks

Jingwen Wang[1], Richard Chen[2], Ming Lu[2], Faisal Mahmood[2]
[1]Brigham and Women's Hospital, Boston, MA, [2]Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Disclosures: Jingwen Wang: None; Richard Chen: None; Ming Lu: None; Faisal Mahmood: None

Background: Histopatholoy-based patient stratification is often of clinical importance for many cancer types including prostate cancer. The commonly used Gleason score, which informs the aggressiveness of prostate cancer, is based on the architectural pattern of tumour tissues and the distribution of glands. However, its manual assignment is extremely time-consuming for physicians as it requires detailed pixel-level annotation. We propose a deep learning based-approach for automatic patient stratification on TMAs using Graph Convolutional Networks (GCNs), which learns from the global distribution of cell nuclei, cell morphometry and spatial features without requiring pixel-level annotation. Using this weakly-supervised method, we hope to stratify out low-risk patients (Gleason score below 6) from actionable cases (Gleason score 6 or above) and alleviate the burden of manual annotation by physicians.

Design: The patient's tissue microarrays (TMAs) are first stain-normalized to remove the color variation. Then, for segmenting the nuclei in the TMAs, we used image-to-image translation with a Generative Adverarial Network (GAN). Based on the segmentation results, we are able to construct a graph for each image. Each nucleus is connected to its top-k nearest neighbor nuclei if they are with in a certain distance with each other. For graph convolution, we use several spatial and morphological features for each nucleus, including the coordinates, roundness, area, eccentricity, etc. We also extracted other features from 64 x 64 windows centered at the nuclei centroids,

including the grey level co-occurrence matrix and embeddings acquired from self-supervised learning. Then, we used graph convolution to extract the higher representation of the whole graph and classify each image as normal or abnormal.

Results: We aggregate the TMAs with gleason score higher than 6 as abnormal tissues, in contrast to normal tissues. Preliminary results show that our method can achieve 82.47% validation accuracy and 0.9039 AUC.

Figure 1 - 1586
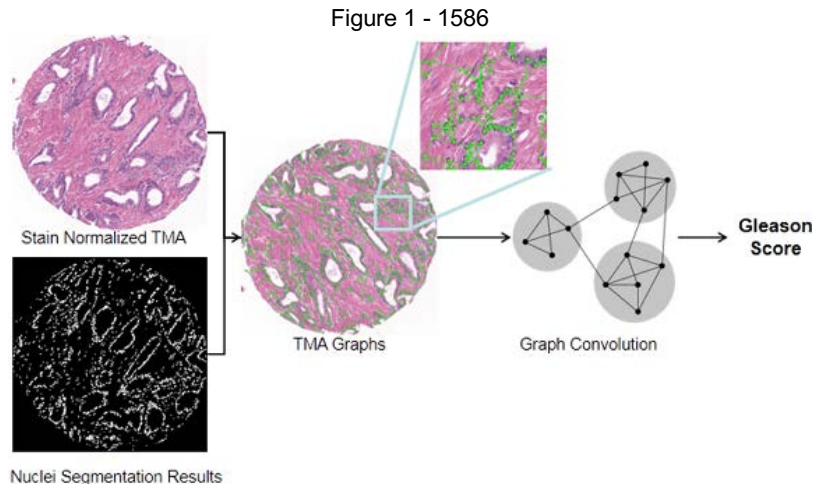
Conclusions: We demonstrate that our GCN-based deep learning pipeline can accurately stratify patients at the TMA level without requiring exhaustive pixel-level annotation. Our work offers a new paradigm for weakly-supervised TMA-level patient stratefication, with the potential to be extended to whole slide images.

1587   Utilization of Deep Neural Network in Recognition of BCR/ABL Gene Rearrangements in Fluorescence In Situ Hybridization Images

Junyan Wu[1], Mustafa Deebajah[2], Zongshan Lai[3], Mark Micale[4], Limin Yu[4]
[1]ZKShuangHe LLC, Beijing, China, [2]Henry Ford Health System, Detroit, MI, [3]Beaumont Health, Royal Oak, MI, [4]Beaumont Health-Royal Oak, Royal Oak, MI

Disclosures: Junyan Wu: None; Mustafa Deebajah: None; Zongshan Lai: None; Mark Micale: None; Limin Yu: None

Background: Interphase dual-color fluorescence in situ hybridization (iFISH) has been used for identification of BCR/ABL gene rearrangements in Chronic Myeloid Leukemia (CML). Artificial intelligence, particularly deep neural network (DNN), has achieved major breakthroughs in image analysis and classification. The purpose of this study is to see if DNN can be successfully trained to recognize of BCR/ABL gene rearrangements in FISH images.

Database of single-cell images (101 positive, 278 negative), and original multi-cell images (33 positive, 118 negative). The classification model was built on single-cell images and use to test the end-to-end performance only on multi-cell images.
The fully automatic analysis pipeline consists of single-cell detection and single-cell classification modules (Fig 1). The detection pipeline following systematic steps; starting with the RGB images are first converted into greyscale. After which we set an intensity threshold to remove the text and apply median blurring to de-noise the image. Referring to the topological structure, we detect the closing contours and generate a bounding box on the contour region. The small and overlapping regions are removed or merged. We crop and resize the original image in the bounding box regions into a dimension of 255 and rescale the RGB value into -1 to 1 as the input to the classification network. The deep neural network architecture for single-cell classification is based on VGG, which consists of 16 layers of convolution, max pooling and fully connected operations. The network outputs a binary vector indicating positive and negative. We use Adam optimizer and Cross-entropy loss to optimize the training process. During the training time, we apply flip operation as data augmentation.

Results: Our end-to-end performance matrices showed a total f1-score of 98% and recall of 98%.

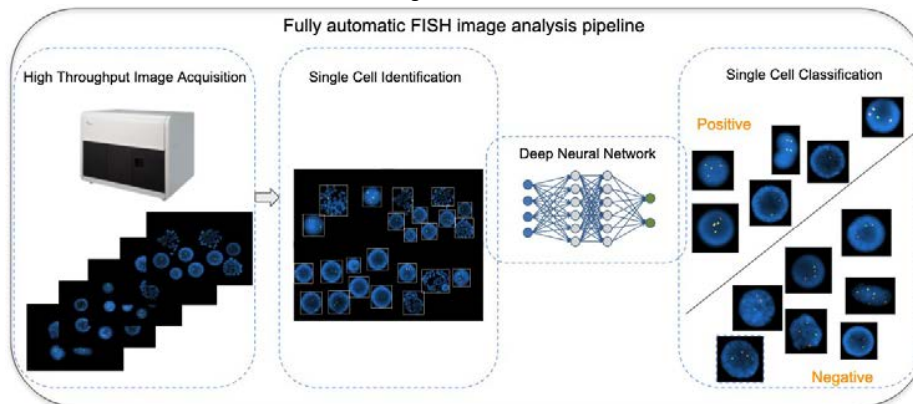|	precision	recall	f1- score	support
negative	0.99	0.99	0.99	85
positive	0.97	0.97	0.97	33
weighted avg	0.98	0.98	0.98	118
Table 1. end-to-end performance report on multi-cell images				

Figure 1 - 1587

Fig 1. Fully automatic FISH image analysis pipeline. The workflow takes high throughput FISH images, then detects the single cell. The deep neural network classifies each cell into positive and negative.

Conclusions: Our study shows the deep neural network can be trained to reliably recognize BCR/ABL gene rearrangements in FISH images with pathologist-level of accuracy.

1588 Computer-Assisted Morphometrical and Statistical Analysis to Differentiate Well-Differentiated Hepatocellular Carcinoma, Hepatocellular Adenoma, and Non-Neoplastic Liver

Rong Xia[1], Xuchen Zhang[2], Dhanpat Jain[3]
[1]SUNY Downstate Medical Center, Brooklyn, NY, [2]Yale University School of Medicine, Orange, CT, [3]Yale University School of Medicine, New Haven, CT

Disclosures: Rong Xia: None; Xuchen Zhang: None; Dhanpat Jain: None

Background: Distinguishing well-differentiated hepatocellular carcinoma (WD-HCC), hepatocellular adenoma (HA) and focal nodular hyperplasia (FNH) from non-neoplastic liver (NNL) is challenging solely based on histology. Some of these cannot be reliably classified on morphology either as a HCC or HA are often called "well-differentiated hepatocellular neoplasm of uncertain malignant potential (HUMP)". In our previous study, we performed the computer-assisted morphometrical study to quantify the nuclear features of the hepatocytes on H&E stained sections, and used the Chi-square Automatic Interaction Detector (CHAID) analysis to develop an algorithm to distinguish WD-HCC from NNL and HA (Figure 1). Here, we used regions of interests (ROIs) of the whole slide scan images to validate the algorithm.

Design: Forty-five core liver biopsy slides (WD-HCC=10; NNL=23; HA=10; HUMP=2) were retrieved and evaluated. The NNL group includes 11 FNH, 5 cirrhosis and 7 normal liver cases. H&E stained slides were scanned at 400X magnification using the Leica AT2 scanner. ROIs were selected and saved as .png, excluding areas with severe fibrosis, cholestasis, artifacts, or portal tracts. The ROI images were deconvoluted to subtract background colors. After noise reduction and Otsu thresholding, morphologic opening, connected component analysis was applied to quantify the nuclear density (D), median nuclear-perimeter (P), median nuclear-sphericity (S), and SD nuclear-eccentricity (E) of each image using ICY. The decision tree model was then used on the whole slide scans to validate the CHAID algorithm (Figure 1). The sensitivity and specificity of the model were evaluated by binary classification analysis.

Results: The cellularity was higher in WD-HCC compared to HA and NNL and significant difference in P were noted between the WD-HCC and HA (Figure 2, $p < 0.05$), while there was no significant difference in S and E identified in the three groups. Using CHAID algorithm the WD-HCC is detected with the sensitivity of 90.0% and the specificity of 93.9% (Tab. 1). Two HUMP cases were both classified as HA, which is also supported by the limited follow-up on both cases. No difference in the nuclear features were identified between the FNH, and the other NNL.

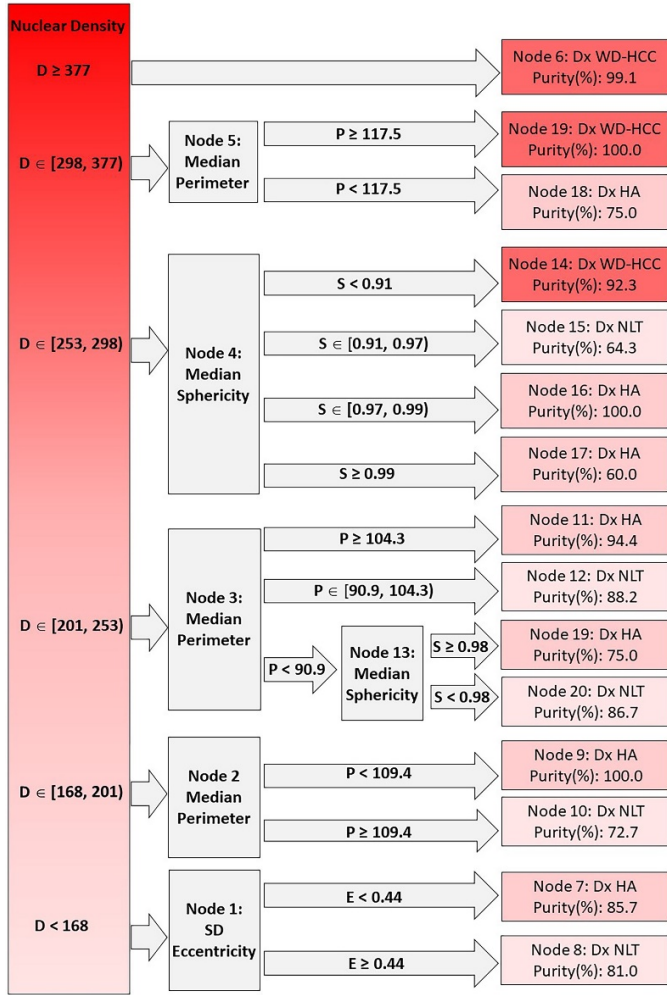| Computer Classification	Histopathologic Diagnosis				Sensitivity	Specificity
		NNL	HA	HCC	Count	
	NNL	16	2	0	18	69.6%	90.0%
	HA	7	6	1	14	66.7%	75.8%
	HCC	0	2	9	11	90.0%	93.9%
	Count	23	10	10	43		

Figure 1 - 1588

Figure 1. Decision tree from CHAID analysis. Columns left to right: root nodes; decision nodes; terminal nodes. Shades of red from dark to light indicate WD-HCC, HA, and NNLT, respectively. Purity of each diagnosis is given in terminal nodes. [] = a set;) =number included in the set; [= number excluded from the set.
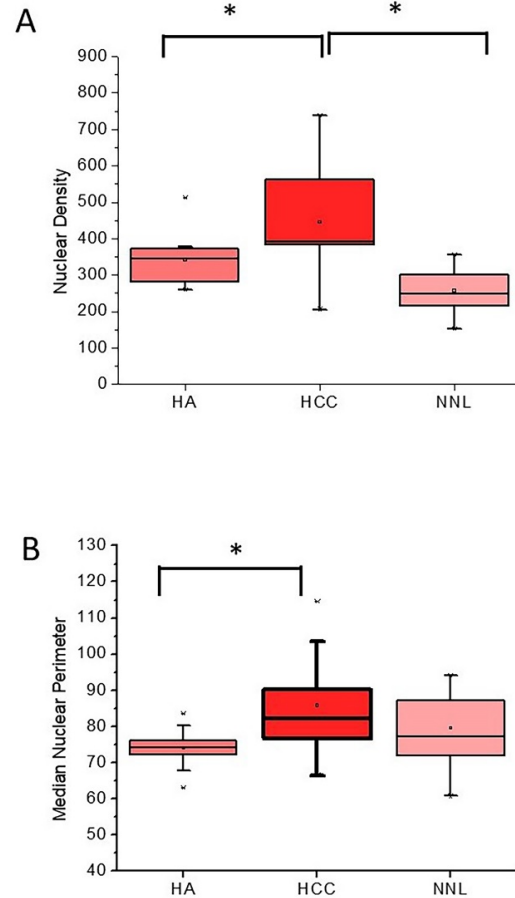
Figure 2 - 1588

Figure 2. Distribution of nuclear density (A) and median nuclear perimeter (B) ($p<0.001$; one-way ANOVA). Bars indicate 5th and 95th percentiles; boxes represent the 25th and 75th percentiles; lines inside the boxes are medians; small boxes indicate means. Asterisks (*) indicate significant ($p<0.05$) differences by Tukey post hoc test.

Conclusions: Computational image analysis of nuclear features can be used to differentiate WD-HCC from benign liver parenchyma with high accuracy and HA in most cases. Computational image analysis can be used to assist in the diagnosis of HCC especially in suboptimal specimens.

1589 Explainable Artificial Intelligence (xAI) for Safe Breast Core Biopsy Diagnosis Support

Mustafa Yousif[1], Akif Tosun[2], Gloria Carter[3], Esther Elishaev[4], Chengquan Zhao[4], Tiannan Wang[5], S. Chakra Chennubhotla[6], Jeffrey Fine[6]
[1]University of Pittsburgh Medical Center, Magee-Womens Hospital, Pittsburgh, PA, [2]SpIntellx, Inc., Pittsburgh, PA, [3]UPMC MAGEE, Pittsburgh, PA, [4]Pittsburgh, PA, [5]Department of Pathology, Magee-Womens Hospital, University of Pittsburgh Medical Center, Pittsburgh, PA, [6]University of Pittsburgh, Pittsburgh, PA

Background: Computational pathology has great potential for augmenting the accuracy and efficiency of Pathologists. Explainable AI (xAI) is a new computational approach that can justify its results to pathologists, to promote safety, reliability and accountability of machine learning for critical pathology tasks. Ground truth data labeling is necessary for machine learning training but has historically been a bottleneck. The best-qualified pathologists are often busy with clinical work, and a major challenge has been creating efficient image

annotation tools that they can easily use. Herein we report the initial phase of a validation study of a proprietary xAI platform for whole slide images (WSIs), HistoMapr-Breast.

Design: With IRB approval, 1931 de-identified WSIs of breast core biopsies (n=862 cases) were scanned at 0.5 microns per pixel (ScanScope AT2, Leica Biosystems, Buffalo Grove IL). Breast duct regions of interest (ROIs) (~80,000 ROIs segmented) were spatially extracted from the WSIs using pointwise mutual information maps. Further computational pipelines were used to generate diagnostically explainable features then classify the ducts into diagnostic categories: invasive carcinoma, ductal carcinoma in situ (DCIS), high-risk and low-risk (see Table 1). Three breast pathologists labeled 4500 ROIs using a novel annotation application.

Results: ROI labeling was rapid and completed in 15-20 minute sessions of 250 images (figure1), yielding diagnostic labels for 4462 ROIs. There was complete concordance for 3172 ROIs (71%) including 650 invasive carcinomas, 132 DCIS, 200 High-risk and 2190 low-risk diagnoses. Fleiss' (overall) kappa was K=0.6613, showing substantial agreement among pathologists.

Table 1: Remapping diagnostic subcategories to 4 main diagnostic labels.	
Invasive carcinoma **Ductal carcinoma in-situ (DCIS)**	Invasive/Infiltrating ductal carcinoma Invasive/Infiltrating lobular carcinoma Invasive/Infiltrating mammary carcinoma Ductal carcinoma in-situ (DCIS) Solid papillary carcinoma Encapsulated papillary carcinoma
High-risk	Atypical ductal hyperplasia (ADH) Atypical lobular hyperplasia (ALH) Lobular carcinoma in-situ (LCIS) Flat epithelial atypia (FEA) Atypical vascular lesion/proliferation Intraductal papilloma Radial scar Complex/radial sclerosing lesion Phyllodes Fibroepithelial
Low-risk/Benign	Fibroadenoma Fibrocystic changes Ductal epithelial hyperplasia Columnar cell changes and hyperplasia Sclerotic adenosis Pseudo angiomatous stromal hyperplasia (PASH) Apocrine metaplasia Benign / unremarkable / normal

Figure 1 - 1589
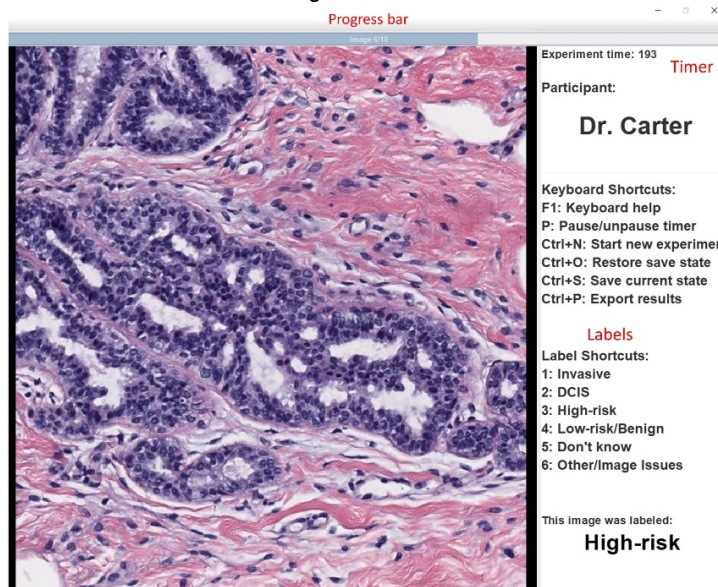
Conclusions: Computational pathology can revolutionize pathology practice, enabling pathologists to make critical decisions that only they can make and delegating tasks that can be automated. xAI is an important new technology that gives pathologists unparalleled situational awareness, providing for trust and confidence. Building such systems requires user-friendly and high-throughput ground truth image

annotation, as outlined in this study. The workflow developed here is similar to what can be used for computer-assisted diagnosis, and therefore is also useful for developing diagnostic xAI tools.

1590 Automated High Resolution 3D Reconstruction of Multi-Modal Whole Slide Images with Application to Growth Patterns in Bone Marrow

Norman Zerbe[1], Michael Franz[2], Genevieve Crane[3], Tim-Rasmus Kiehl[4], Peter Hufnagl[5]
[1]Charité - University Hospital Berlin, Berlin, Germany, [2]University of Applied Sciences Berlin, Dept. Applied Informatics, Berlin, Germany, [3]New York-Presbyterian/Weill Cornell Medical Center, New York, NY, [4]Charité - University Medicine Berlin, Berlin, Germany, [5]Charité University Medicine Berlin, Berlin, Germany

Disclosures: Norman Zerbe: None; Michael Franz: None; Genevieve Crane: None; Tim-Rasmus Kiehl: None; Peter Hufnagl: None

Background: The bone marrow microenvironment is dysregulated in various disease states, leading to an abnormal number, localization and/or morphology of marrow elements as well as potential stromal alterations. The bone marrow morphology is a key criterion in the diagnosis of myeloproliferative neoplasms; however, these entities may show overlapping features on histologic sections, particularly at the early stages of a disease. We sought to develop a 3-dimensional model of the bone marrow that could be applied to clinical specimens to enable a better assessment of morphology. This tool may ultimately aid in the diagnosis of myeloproliferative neoplasms as well as insights into their underlying biology.

Design: The image data used are *ScanScope*- whole slide images (WSI) in SVS-format, where each case consists of multiple slides per stain (H&E, PAS or trichrome stains) having three sections on one slide. The application of an adaptive threshold-based segmentation algorithm on the WSI data produces an image series of the tissue particles' region of interest. For the subsequent alignment process, the *elastix/transformi*x framework is used to optimize and regularize the registration results. A python console application processes the image stack and call the *elastix* API with the current moving image, the corresponding fixed image and a parameter set for the desired transformation type. The registration produces a stack of aligned images, that are the basis for the 3D-model creation and visualized using 3D-Slicer.

Results: The aim of the processing pipeline is to enable pathologists to access histological image data and to support simple and comfortable exploration of these usually very large datasets. Therefore, the above-described pipeline is preprocessed, and the 3D-scenes are accessible from a standalone application that loads with user-defined preferences. Furthermore, the navigation and interaction with the 3D-model (fig.1) is intuitive. Included are functionalities for browsing through single image layers, to cut through the model from different angles and to interact using a 3D-mouse.

Figure 1 - 1590
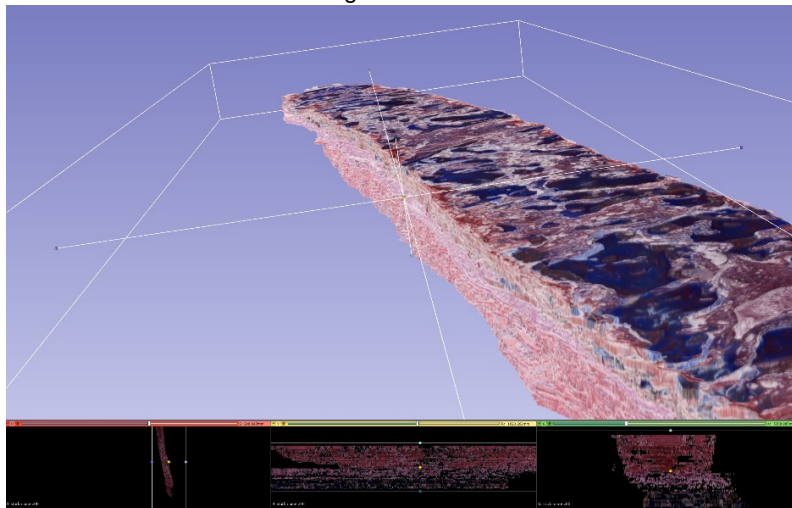
Conclusions: The registration of high-resolution histological image data is a very resource-consuming task. The large dimensions of the image data are not only a challenge with respect to capacity but also a tough problem in finding the proper mapping of consecutive sections. The pipeline introduced here provides an efficient and accurate method for cases and its automatic preprocessing, and provision to pathologists.