Short Abstract Presentations

Tuesday, May 24, 2016
8:00 am – 9:00 am

Locations:
Wyndham Grand Pittsburgh Hotel

Grand Ballroom 1

Kings Garden 1

Kings Garden 2/3

Advanced Pathology Informatics
Computational Pathology

Grand Ballroom 1

Development of A Digital Pathology Database for Annotation and Quality Management of a Brain Tumor Biorepository

Jeremy Molligan, MD (jeremy.molligan@jefferson.edu); Robert Stapp, DO; Miraj Patel, Jack London, PhD; Chirayu Goswami*, PhD; James Evans, MD; Stephen Peiper, MD

Content
Biorepositories play a crucial role in biomedical research on human disease. A richly annotated biorepository was designed and implemented at Thomas Jefferson University through the integration of clinical data, pathologic data, whole-slide imaging, and CAP-modelled quality control procedures. Specimens with corresponding clinical and genomic data are accessible via a single web portal.

Technology
Open-source software was utilized, including: Open Specimen v1.2 (database application for specimen inventory, tracking, and annotation), i2b2 (a patient data analytics platform), NeuroDB (custom clinical data acquisition web application), the Aperio whole-slide imaging system, Cerner CoPath Plus v2012 (A/P LIS) [Kansas City, MO, United States], and inventory management utilizing BioTillion [Skillman, NJ, United States] RFID enabled specimen vials.

Design
After consenting the patient, clinicians enter clinical data into NeuroDB. Blood and tissue samples are collected directly from the OR, processed, and annotated within Open Specimen. Specimens are aliquoted for FFPE and routine histology (1), nucleic acid extraction (2), and cryopreservation in LN2 (6). The QC slide is examined by a pathologist, electronically documented for quality control, and scanned utilizing the Aperio imaging system. The pathology files are accessible for quality management and to investigators via an i2b2 query tool link and a web-based image viewer. Pathology reports are transferred from Cerner CoPath Plus to i2b2 via an HL7 interface, where the information is parsed, de-identified, and uploaded into the i2b2 research data mart.

Results
Researchers are able to search for banked tissue samples via the i2b2 query tool. These specimens have been managed utilizing strict quality control standards, are annotated with relevant clinical history, morphologic diagnoses, molecular diagnoses, and digital images.

Conclusions
Current medical research requires access to large numbers of tissue samples consented for broad research use, that have extensive annotation, and whose collection and preservation conserves the integrity of macromolecules. Past archived specimens obtained rarely meet all of these requirements making it necessary to establish a state-of-the-art biorepository to for current and future research. The development of a complementary database of digital images will enhance quality management and provide scientists with knowledge of the composition of banked tissue.

Standardization of Electronic Templates for Cancer and Biomarker Reporting

Keren I. Hulkower PhD[1] (khulkow@cap.org), Thomas P. Baker MD FCAP[2], Michael A. Berman MD FCAP[3], James Dvorak MT(ASCP)[1], Jaleh Mirza MD MPH[1], Richard L. Moldwin MD PhD[1], and Samantha Spencer MD[1]

[1]College of American Pathologists, Northfield, IL, USA
[2]The Joint Pathology Center, Silver Spring, MD, USA
[3]Jefferson Regional Medical Center, Jefferson Hills, PA, USA

Content
The Cancer Committee and Pathology Electronic Reporting Committee of the College of American Pathologists (CAP) have developed data capture forms for cancer and biomarker reporting. The Pathology Electronic Reporting Template Standardization Project workgroup made suggestions for standardizing sections of the CAP electronic Cancer Checklists (eCC).

Technology
Issues regarding template standardization were captured from eCC end users using custom Project Tracker software. The eCC Template Editor was used for visual content modeling of question and answer sets, with models and metadata stored in a SQL Server database. A custom .NET tool was used with code generation software to generate an eCC schema-compatible C# object model to serialize the database records into the eCC XML format. Additional XSLT programs were used to transform the XML files into eCC HTML files suitable for Template Standardization Project review.

Design
The Template Standardization Project team worked with Cancer Committee authors to arrange questions and sections to fit common pathologist workflows for data entry and reporting. The workgroup reviewed eCC HTML output of many modeling options and reached consensus on template standards.

Results
A standardized eCC base template was created for consistent modeling. Tumor site, histological type, grade, size, extent and accessory tumor findings were consolidated in nested sections under a single "TUMOR" header to improve workflow and combine data elements related to pT classification. The "MARGINS" section was modified to improve workflow when reporting tumor distance from closest margins and non-invasive histologic types present at margins. The "LYMPH NODES" section was standardized for reporting laterality, nodal stations and sentinel nodes. Immunohistochemical and biomarker data elements were moved into separate biomarker templates. Finally, new metadata attributes were added to the eCC database and XML templates to facilitate known problem areas such as textual changes needed for synoptic reporting and reporting rules.

Conclusions
The Template Standardization Project has improved modeling of electronic cancer and biomarker data entry forms to better fit the pathologist workflow and reporting needs.

Selection of a Laboratory Information System for the Molecular Pathology Laboratory: Unique Aspects and Key Decision Factors

Roy E Lee, MD (leer3@ccf.org), MS, and Walter H. Henricks, MD

Cleveland Clinic Foundation

Content
Many molecular laboratories rely on paper forms and spreadsheets for information management, instead of information systems such as for anatomic/clinical pathology.

Functional requirements for a molecular laboratory's information system (LIS) are not adequately addressed by information systems designed for the clinical laboratory or anatomic pathology. This is due to unique aspects of molecular diagnostics such as data models specific to molecular testing, varied testing workflows involving multi-step method protocols, and complexities of molecular reporting.

This study reports a structured approach for evaluating and selecting a molecular LIS solution, and identifying key decision factors.

Technology
Word processor and spreadsheet programs (Word and Excel, Redmond, WA)

Design
A multi-step approach was designed, including 1) complete inventory of current information management practices in the molecular laboratory, 2) definition of key functional requirements and priorities for a laboratory information system, 3) development and distribution of RFP, and 4) evaluation of vendor responses and system demonstration visits.

Results
A comprehensive document was generated that catalogued all information practices in the laboratory, compiling examples of paper forms, spreadsheets, and other laboratory information practices. This document laid the foundation for further system evaluation, including gap analysis, RFP development, and demonstration playscripts. In addition to assessing best fit for workflow/operations support, this approach yielded several key decision factors crucial to selection:

- Clear definition of scope – for example, initial de-emphasis on next generation sequencing tests
- Compatibility with preanalytic processes such as specimen identification/ tracking
- Interoperability with other IT systems
- Testing procedures defined in system by laboratory users, not LIS managers or vendor
- Commitment by vendor to develop novel interfaces between LIS and molecular instrumentation
- Input and buy-in from laboratory technologists and technicians in a discipline where paper and spreadsheets are the standard methods.

Conclusions
Initial creation of a comprehensive inventory of information management practices was a key success factor in laying down the foundation for molecular LIS selection. The most significant considerations identified were compatibility with existing preanalytic processes in the laboratory and interoperability with other systems. Understanding the unique considerations and needs in selecting a molecular LIS can be facilitated through this structured approach.

Computational Models of Oncogenic Networks: A Potential Tool for Computational Pathology

Edward C. Stites, MD, PhD (estites@path.wustl.edu)

Department of Pathology and Immunology, Washington University in St. Louis

Content
Although cancer genomics has identified the commonly mutated cancer driver genes, clinical cancer genomics struggles to deal with all of the different mutants that can be found within a given gene. Few have been sufficiently well studied to be more than a VUS – variant of uncertain significance. There is evidence that all hotspot mutations are not equivalent, which suggests that the set of "VNS" –variants of known significance – may be even smaller than has been assumed. Informatics approaches are needed that can better explain and predict mutant specific behaviors. This abstract focuses on a computational model of oncogene signaling with demonstrated an ability to make mutant-specific predictions.

Technology
A computational model of Ras signaling was developed that is based upon mass-action kinetics, ordinary differential equations, and existing knowledge of the Ras pathway. The model makes mutant-specific predictions by incorporating biochemical rate constants for specific oncogenic Ras mutants and their interactions with regulatory proteins. The model has been developed, simulated, and analyzed in MATLAB. The model includes twelve oncogenic mutants and ten Ras mutants found in Noonan syndrome.

Design
The model was then applied to problems where different mutants to the same KRAS gene have been observed to have different biological and clinical phenotypes. The model was then used to investigate whether the model could predict these differences and to investigate what the key parameters were that could explain the different behaviors.

Results
The model successfully distinguishes between oncogenic and non-oncogenic mutants. The model suggests that mutant strength is relative, which suggests that knowledge learned about specific Ras mutants in one type of cancer may not apply to other cancers. The model also reveals that an unexplained KRAS response to targeted therapies is actually consistent with known information about Ras signaling; this logical conclusion was not apparent without our computational model.

Conclusion
Mass-action based models of disease promoting variants have the ability to relate biochemical data that describe distinct mutants to phenotypes important to clinical medicine. Models like that that incorporate a mechanistic understanding of biological processes should play an increasing role in Computational Pathology.

Computational Pathology

Kings Garden 1

Comparison of Applied Machine Learning Tools for the Prediction Of Myelodysplastic Syndromes Using Complete Blood Count Parameters

Amrom E. Obstfeld, MD, PhD[1] (obstfelda@email.chop.edu), Philipp W. Raess, MD, PhD[2], Stephen R. Master, MD, PhD[3], Adam Bagg, MD[1]

[1]Department Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine
[2]Department of Pathology, Oregon Health & Science University
[3]Department Pathology and Laboratory Medicine, Weill Cornell Medical College

Content
Patients with myelodysplastic syndromes (MDS) frequently present with cytopenias of uncertain etiology. A gap in their care may reside in the decision as to when the suspicion for MDS is sufficient to warrant a bone marrow biopsy. The aim of this study was to compare the performance of several machine learning tools in the prediction of the presence of MDS using cell population data parameters from an automated complete blood count (CBC) analyzer as predictors.

Technology
Cell population data was collected from a Beckman Coulter LH780 analyzer. The randomForest, rpart, ada, nnet, kernlab, and bartmachine packages were used to generate 7 different models within the R statistical program. Using the ROSE package, an additional 6 models were generated using these packages after compensating for imbalance in the dataset.

Design
CBC data from outpatients were collected and patients with MDS were identified by screening the electronic medical record for patients with an ICD-9 code corresponding to MDS; MDS was confirmed by review of clinical data. CBCs were randomized into independent, training and test sets. The training set consisted of 39 CBCs from patients with MDS and 3,294 CBCs from controls, and the test set consisted of 20 CBCs from patients with MDS and 1,686 CBCs from controls.

Results
Classifiers were compared on the basis of routine assay performance criteria; results are presented in Figure 1. As would be expected given the heavily imbalanced training set, several of the models failed to predict any cases of MDS in the test set, resulting in sensitivities of 0% and specificities of 100%. Much of these shortcomings were overcome by random over-sampling the cases of MDS in the training set, however this came at the expense of a decrease in specificity. The best performing classifiers for the purpose of screening a general outpatient population for MDS were the generalized linear model, random forest classifier, and the neural network after random over-sampling.

Continued on next page

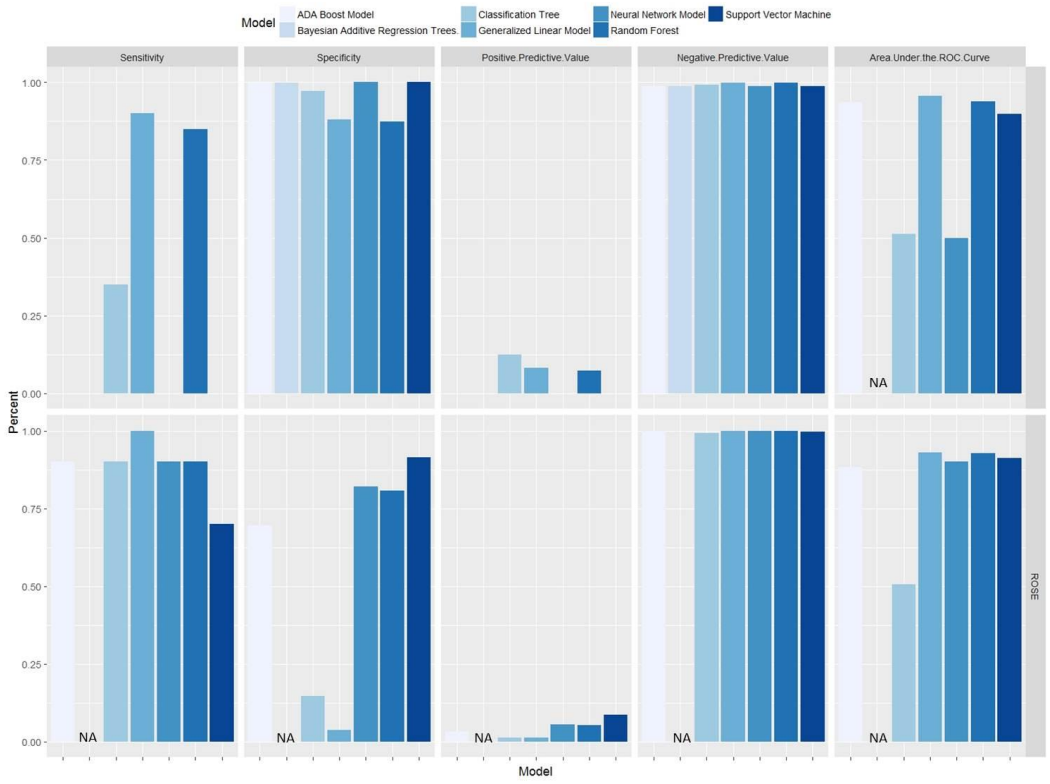

Figure 1: Performance characteristics for seven models without (top) and with (bottom) random over-sampling

Conclusion

Our data highlight the differences in accuracy of commonly used machine learning techniques in a specific use case. Future work will investigate the utility of these techniques in other populations at risk for MDS.

Computational pathology framework for segmenting and classifying clinically relevant regions of interest in whole slide images of breast tissues

Luong Nguyen (lun5@pitt.edu), Akif Burak Tosun, Adrian V. Lee, D. Lansing Taylor, Jeffrey L. Fine, Chakra Chennubhotla

University of Pittsburgh

Content

The current manual practice of pathology is time consuming, error-prone, and subjective. Computational pathology with advanced image analysis is crucial to increasing the productivity of pathologists and improving the accuracy of cancer diagnoses. We propose a comprehensive computational framework for identifying and characterizing various regions of interest (ROIs) in breast biopsies, to help pathologists quickly navigate whole slide images (WSIs). We parse WSIs of H&E stained breast tissues into ROIs, characterize the ROIs with spatial image statistics, then rank them based on their clinical relevance into distinctive phenotypes: invasive cancer, carcinoma in situ, atypia, and normal.

Technology

Breast tissue slides were scanned into WSIs at 0.5um/pixel resolution, then viewed with provided software (Aperio XT, Leica, Vista CA USA). Ground truth data for ROIs was collected with Prospector (PMC4466790).

Design

Our dataset includes 30 WSIs of breast tissues from TCGA. First, color normalization is done on WSIs to alleviate effects of stain variability on tissue appearance. Second, superpixels in purple (nuclei), pink(stroma), and white (lumen, fat, tissue tears) channels are generated on the color normalized images. Third, coarse-grained ROIs are proposed using context and distance distributions of superpixels in all three channels. Fourth, the ROIs' boundaries are further refined using a state of the art spatial statistics based segmentation algorithm. Finally, spatial statistics from each ROI are used to rank them from the most suspicious (invasive cancer) to the least (normal tissue).

Results

Fig. 1 top row shows the computational framework proposed here, while the bottom row provides more details on the specific modules within our framework. The framework
is highly accurate in segmenting and classifying ROIs in breast tissue images with segmentation scores on par with state of the art algorithms and a classification accuracy averaging at 77%.

Conclusion

Our paper shows an end-to-end framework for navigating whole slide breast tissue landscapes using ROIs as tissue landmarks. In the future, we will incorporate an active learning strategy that seeks feedback from pathologists to further customize and improve the accuracy of the computational pathology pipeline. Not only will such work lead to computer-assisted pathology signout, it will continually improve based on prior experience.

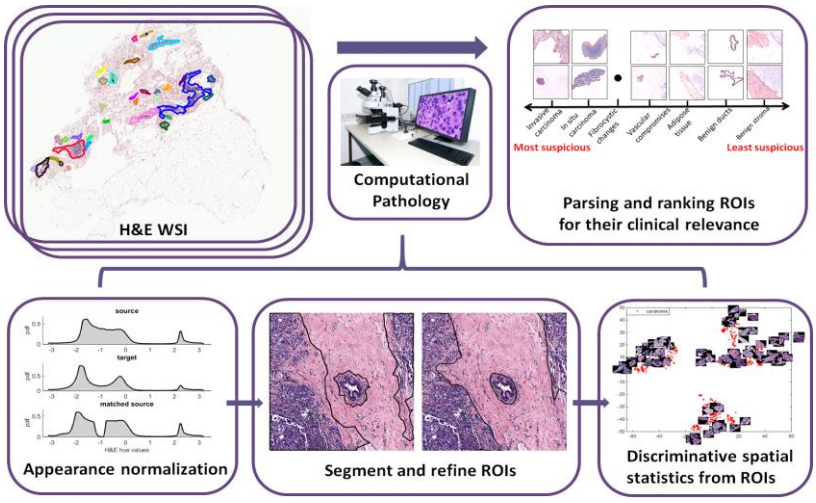

Figure 1: Computational pathology framework for navigating whole slide images of breast tissue with clinically relevant regions of interest

Extracting context-aware diagnostically relevant patterns from H&E stained lung tissue images

Akif Burak Tosun, PhD, (tosun@pitt.edu), Chakra Chennubhotla, Ph.D.

University of Pittsburgh, Department of Computational and Systems Biology

Content
With digital pathology rapidly evolving, image analysis tools have become inevitable for facilitating the needs of current pathology workflow; e.g. decreasing the workload from normal gestalt, easy handling of huge image repositories like TCGA, assisting on diagnostic decisions. In this study, we present an image analysis algorithm to detect structural patterns in whole slide images (WSIs) using context-aware features extracted from representative tissue components. Specifically, we worked on tissue images of Lung Adenocarcinoma (LUAD) and Lung Squamous Cell Carcinoma (LUSC), which are the most common types of lung cancers. Since these two cancer subtypes have different prognosis and treatment directions, it's crucial to differentiate between them.

Technology
Phenotypical structural patterns were extracted from H&E stained WSIs of 156 LUAD and 188 LUSC patient cohorts from TCGA. The Pittsburgh Supercomputing resources have been extensively used to process WSIs. The average processing time for a single WSI image, roughly of size 15K x 15K, is about 15 minutes.

Design
We first generate structural elements that roughly correspond to tissue components, nuclei, stroma and lumen regions, of the WSI (Fig. 1A). Next, we encode spatial context of the tissue image by building co-occurrence matrices around each tissue component. Finally, we cluster the co-occurrence matrices into a small number of *phenotypical* spatial patterns (~10). We report the total tissue area occupied by each of the structural pattern as a feature vector for any given WSI.

Results
Fig. 1B shows how encoding for spatial context enhances the differences between LUAD and LUSC. The spatial spread of the phenotypical structural patterns changes dramatically between LUAD and LUSC, as seen in the example prototypes of Fig. 1C.

Conclusions
Our experiments prove the existence of context-aware diagnostically relevant image phenotypes that can discriminate the two types of lung cancers, LUAD and LUSC. For future work, we will build a more discriminative framework for associating tissue phenotypes to genomic signatures that are readily available in TCGA. In parallel, we will also expand our algorithm to encode *global* or long-range spatial dependencies by building hierarchical Markov random fields.

Continued on next page

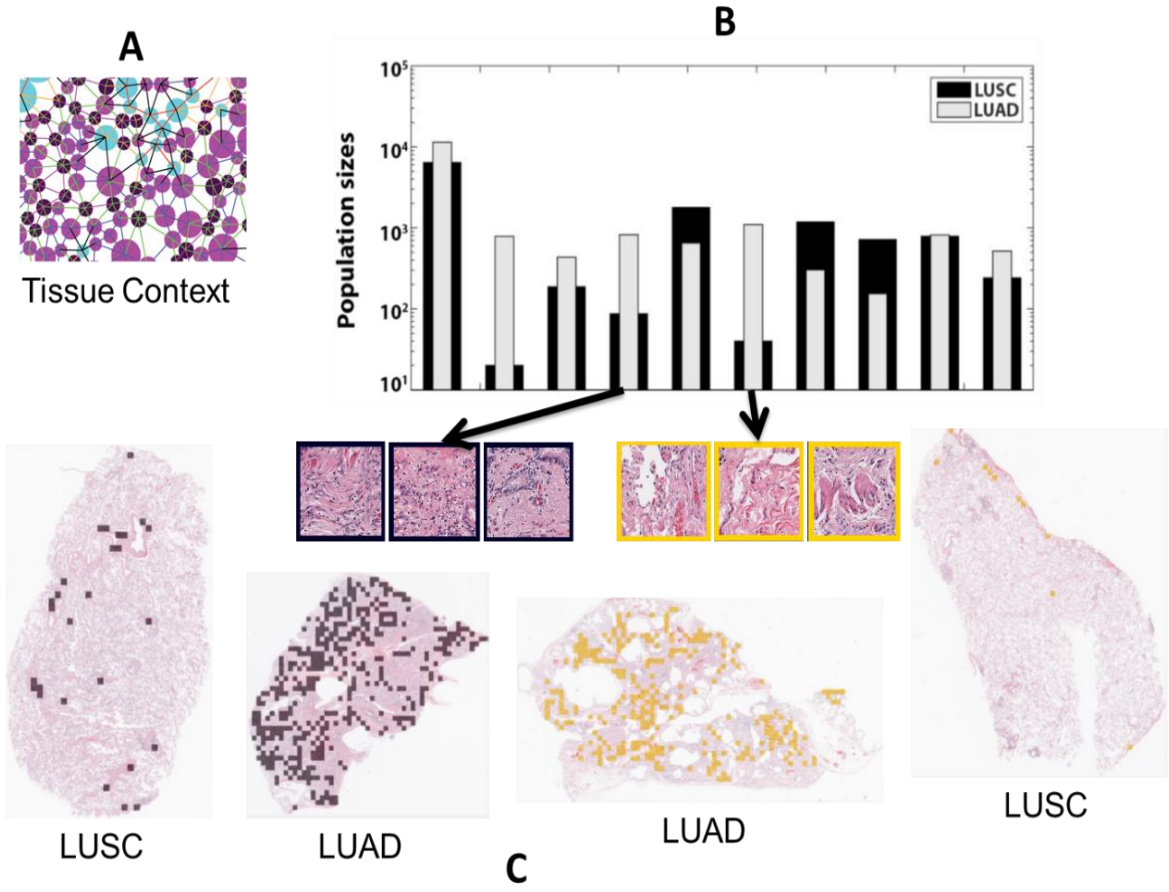

Figure 1 - **Extracting context-aware diagnostically relevant patterns from H&E stained lung tissue images** (A) Tissue components in the form of nuclei, stroma and lumen are spatially connected as nodes in a graph over the WSI. (B) Relative distribution of context-aware phenotypical spatial patterns, (C) two examples of which are shown outlined in black and yellow boxes. Their spatial distribution on the WSI can discriminate LUAD from LUSC.

Advanced Pathology Informatics
Computational Pathology
Imaging Informatics

Kings Garden 2/3

Morphologic Profiling of Erythrocytes Using Deep Convolutional Neural Networks

***Thomas J.S. Durant, MD[1]** (thomas.durant@yale.edu); *Eben M. Olson, PhD[1]; Richard Torres, MS, MD[1]

[1]Department of Laboratory Medicine, Yale School of Medicine, New Haven CT, 06520

Content:
Morphologic examination of peripheral blood smears remains a vital component of diagnosis in a large percentage of patients. Advancements in the field of automated object recognition have motivated the development of automated morphologic profiling of erythrocytes using machine learning-driven image classification. Most recently, convolutional neural networks (CNNs) has emerged as a superior machine learning technique for complex image recognition. Thus, we sought to evaluate the performance of this powerful approach when applied to classification of erythrocyte morphology profiles from peripheral blood smear images.

Technology:
Unlabeled images used for our training dataset were obtained using the CellaVision® DM96 platform slide scanner. Erythrocytes were assigned labels using an in-house developed web-application, with functionality dedicated to individual cell label assignment. The machine learning algorithm was based on a deep CNN architecture, which was implemented in Python 2.7. Computation was performed on a Linux environment using a standard high performance computer with programmable graphics card.

Design:
A limited label set was used for this initial study which included normal, echinocyte, and schistocyte. Classification of individual erythrocytes was done by a lab medicine resident and a hematopathologist. During the labeling process, cells which had already been shown to the user were randomly re-shown to capture intra-rater reliability. A subset of classified cells were then utilized as a training set for the CNN. A second 'naïve' test set of manually classified cells was also evaluated by the CNN and discrimination accuracy was calculated relative to manual classification.

Results:
The final training database was comprised of 2407 labeled erythrocyte images. The naïve dataset used to test accuracy was comprised of 495 cells. Inter-rater agreement demonstrated simple concordance of 95%, and a kappa concordance of 79% (71% to 87%; CI 95%)(Table 1). Sensitivity for automated detection of schistocytes and echinocytes was 83% and 91% respectively, while specificity was 99% and 96% respectively.

Conclusion:
Findings on this limited analysis suggest CNN performance for erythrocyte classification compares favorably to intra-human classification accuracy and demonstrate potential for improved specificity relative other machine learning algorithms based on available published data. Future work using CNNs on expanded datasets with diversified erythrocyte classifications is warranted.

Table 1.

		Manual analysis		
		Normal	Echinocyte	Schistocyte
Automated Analysis	Normal	389	2	0
	Echinocyte	14	32	3
	Schistocyte	3	1	15

The Putative Role of MALDI-MSI in the Study of Membranous Glomerulonephritis

Vincenzo L'Imperio, MD (v.limperio@campus.unimib.it), Fulvio Magni, Andrew Smith, Franco Ferrario, Elena Ajello, Fabio Pagni

Center of Nephropathology, San Gerardo Hospital, Monza, Italy

Content
Membranous Nephropathy (MN) can be idiopathic (iMN) or manifests as a result of systemic underlying conditions as a secondary epiphenomenon. For the prognostic and predictive consequences of this discrimination is crucial the discovery and development of reliable markers useful to assess these cases. The employment of proteomics and bioinformatics techniques directly on affected renal tissue can be helpful for these purposes, allowing the identification of new candidate biomarker for this disease.

Technology
MALDI-MSI, ScanScope CS digital scanner and various bioinformatics tools (SCiLS Lab 2014, FlexAnalysis3.4, FlexImaging 3.0) will be employed in an integrative fashion to extract morpho-proteomic informations directly from formalin fixed paraffin embedded (FFPE) tissue of the renal biopsies of patients affected by MN.

Design
For each patient, two 4 µm thick sections will be cut from a MN renal biopsy and thaw-mounted on the same conductive Indium Tin Oxide (ITO) glass slide (Bruker Daltonik GmbH, Germany). After the preparation of sample, slides will be analysed in linear positive mode in the mass range of 3000 to 20000 m/z using an UltrafleXtreme (Bruker Daltonik GmbH) MALDI-TOF/TOF MS equipped with a Smartbeam laser operating at 2kHz frequency. Resulting spectra information will be compared and localized on the virtual slide obtained after staining each biopsy with Trichrome and scanning them through ScanScope CS digital scanner (Aperio, Park Center Dr., Vista, CA, USA). Then, FlexImaging 3.0 (Bruker Daltonics, Bremen, Germany) data, containing spectra of each entire measurement region, will be imported into SCiLS Lab 2014 software (http://scils.de/; Bremen, Germany) after the acquisition. SCiLS will be used to perform a series of pre-processing steps on the loaded spectra: baseline subtraction (TopHat algorithm) and normalisation (total ion current algorithm). A series of further steps will be performed in order to generate an average (avg.) spectrum representative of the whole measurement region and of the primary GN sub-classes: peak picking (orthogonal matching pursuit algorithm), peak alignment (to align the detected ions with peak maxima) and spatial denoising (http://scils.de/; SCiLS Lab; 8.8 Spatial Denoising). Principal Component Analysis (PCA) will be also performed to reduce the high complexity of the data. Finally, Receiver Operative Characteristic (ROC) analysis will be performed, with an AUC of >0.8 being required, as an additional criteria to the $p<0.05$, for a peak to be considered as statistically significant. For MALDI-MS/MS spectra, baseline subtraction and smoothing will be performed using FlexAnalysis3.4 (Bruker Daltonics, Bremen, Germany). All MS/MS spectra will be searched against the Swiss-Prot database (Release 2015_05 of 29-Apr-15) with the Mascot 2.4 search engine (Matrix Science, London, UK).

Results
Applying the described study design, that already allowed us the identification of alpha-1-antitrypsin as a candidate biomarker responsible for the sclerosis deposition in many glomerulopathies (Smith et al, Proteomics. 2016 Jan 7), we will analyze 30 cases of MN (15 idiopathic and 15 secondary forms) with the aim of determine differences among primary and secondary forms in terms of type and distribution of protein expression spectra and, eventually, identify a candidate biomarker able to distinguish these two forms.

Conclusion
The employment of MALDI imaging technique directly on FFPE tissue specimens (a reliable substrate in a clinical setting) combined with the application of particular bioinformatics tools (SCiLS Lab 2014, FlexAnalysis3.4, FlexImaging 3.0) can allow the identification of candidate biomarkers for diagnostic and prognostic purposes in patients affected by MN.

Benchmarking Utilization with Population Prevalence Data: A Novel Use of National Census Data

Joseph Rudolf MD (jrudolf1@partners.org)[1, 2], Kristi Smock MD [3, 4]; Brian Jackson MD MS[3,4]; Robert Schmidt MD PhD MBA[3,4]

[1]Department of Pathology, Massachusetts General Hospital, Boston, MA
[2]Harvard Medical School, Boston, MA
[3]Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT
[4]ARUP Laboratories, Salt Lake City, UT

Content
Benchmarking is a useful tool for comparing provider or institution variation to that of a peer group, and can aid in identifying targets for laboratory utilization management. Some underlying level of variation is expected based on site or practice intrinsic factors including the population prevalence of a disease, and may require normalization for the interpretation of benchmarking results. This may be especially true for molecular assays where gene prevalence may be a significant driver of the positivity rate, and thus influence the yield of a given test. We sought to incorporate tailored population prevalence data to aid in a molecular test benchmarking initiative.

Technology
Millenium LIS (Cerner, North Kansas City, MO, US). Anaconda 3.5.1 (Continuum Analytics, Austin, TX, US). Microsoft Excel 2007 (Microsoft, Redmond, WA, US).

Design
De-identified results were obtained for two common molecular tests, Factor V Leiden (FVL) and Prothrombin gene mutation (PTGM), performed at a national reference laboratory. Test positivity rates were calculated for a selection of clients ordering more than 50 of the test of interest during the study period. Predictions for FVL and PTGM client prevalence rates were made by incorporating 2010 national census ethnicity data, based on the location (county level) of the client, using published gene prevalence observations by ethnicity. Prevalence predictions were merged with client positivity rates and other client level information including facility type and size.

Results
The final data set included data from over 100 different institutions with more than 25,000 results spanning 22 months for FVL and more than 40,000 results spanning 48 months for PTGM. The overall positivity rate was 11% for FVL and 5% for PTGM. Significant variation was observed between client positivity rates following normalization for client location predicted population prevalence as seen in Figure 1.

Conclusion
Normalization for intrinsic site or practice factors is necessary for the interpretation of a benchmarking analysis. Incorporation of national population data can aid in laboratory ordering pattern normalization and helps to highlight underlying utilization variation. We plan to incorporate other client factors in our model to further explain residual sources of variation and identify targets for utilization management.

Figure 1 on next page

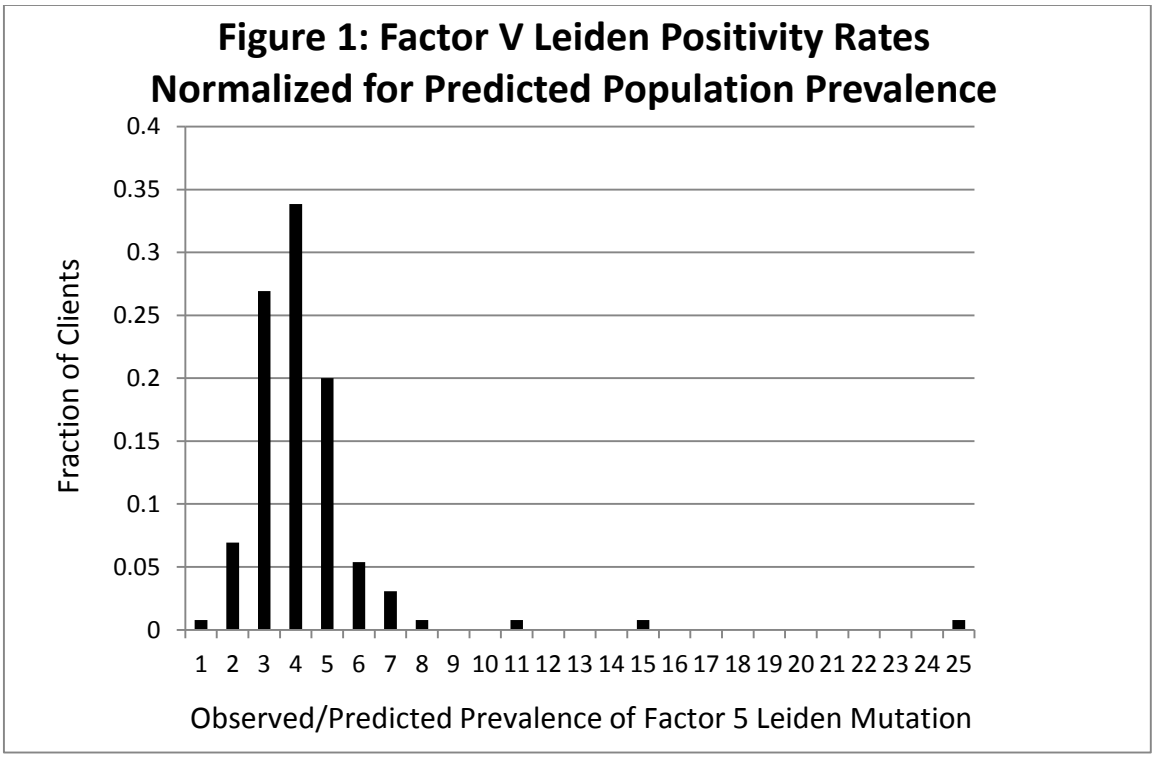
Figure 1: Factor V Leiden Positivity Rates
Normalized for Predicted Population Prevalence
Fraction of Clients
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Observed/Predicted Prevalence of Factor 5 Leiden Mutation

Enhancing Lupus Anticoagulant Reporting Workflow using Python

Ernest Chan, MD (ernest.chan@uchospitals.edu); Charles Van Slambrouck, MD; Jonathan L. Miller, MD, PhD; David S. McClintock, MD

University of Chicago

Content
Comprehensive lupus anticoagulant testing involves multiple test platforms, phases and outcomes, with diagnosis of a lupus anticoagulant requiring interpretation of laboratory results in the context of the patient's clinical history and recent medications. This complexity makes reporting standardization a challenge. Our objective was to create an automation tool that analyzes test data and makes logical deductions to suggest appropriate interpretive statements, thus aiding the pathologist in final interpretation and reporting.

Technology
Python 2.7.10 with packages xlrd 0.9.3, unidecode 0.4.18, python-docx 0.8.5

Design
In our clinical coagulation laboratory, technologists input each patient's test results into an Excel spreadsheet. A series of python scripts were developed to read the data in the Excel spreadsheet and generate a preliminary text report as a Word file. The clinical pathologist then reviews and modifies this report as appropriate based on the both the laboratory data and clinical information for each case.

Results
This program was well-received and rapidly adopted by clinical pathologists due to its ease of use, increased time-efficiency, and error reduction. Feedback showed the program was easy to use (average 4.6 of 5, with 5 being very easy to use), with all users noting either no change (40%) or an improvement (60%) in the quality of the reports. Further, users reported time savings ranging from 4 minutes (14.3%) to 10 minutes or more (57.1%) per report. Finally, this program enabled a successful transition to a paperless workflow in the coagulation laboratory for all lupus anticoagulant testing.

Conclusion
The development of automation tools to assist pathologists in their routine clinical work can result in more standardized reporting and significant time savings. These improvements allow pathologists to focus attention on the interpretation and correlation with clinical findings. Given this project's success, similar design concepts based in Excel and Python could be applied to other areas of the clinical laboratory.


SAVE THE DATE
May 22-25, 2017 • Pittsburgh, PA
PATHOLOGY
INFORMATICS
SUMMIT 2017