Advancing Practice, Instruction, and Innovation Through Informatics (APIII 2008) Conference

Scientific session presentations (http://apiii.upmc.edu/abstracts/2008_sci.cfm) and electronic poster presentations (http://apiii.upmc.edu/abstracts/2008_ep.cfm) were conducted at the 13th annual international conference on Advancing Practice, Instruction, and Innovation Through Informatics (APIII 2008) on October 19–23, 2008, at the Marriott City Center Hotel, Pittsburgh, Pennsylvania. The course director was Michael J. Becich, MD, PhD, University of Pittsburgh School of Medicine. The course codirectors were John R. Gilbertson, MD, Harvard Medical School, and Daniel L. Rubin, MS, MD, Stanford University Medical Center. The scientific session director was David J. Foran, PhD, Robert Wood Johnson Medical School. The practical informatics course directors were John Sinard, MD, PhD, Yale University School of Medicine, and Anil Parwani, MD, PhD, University of Pittsburgh School of Medicine.

SCIENTIFIC SESSION PRESENTATION ABSTRACTS

Color Correction of Pathologic Images

Tokiya Abe, PhD[1] (tokidoki321@gmail.com); Pinky A. Bautista, PhD[1]; Yukako Yagi, PhD[1]; Yuri Murakami, PhD[2]; Masahiro Yamaguchi, PhD[2]; Nagaaki Ohyama, PhD[2]; Hideaki Haneishi, PhD[3]; John Gilbertson, MD.[1] [1]Department of Pathology, Harvard Medical School, Boston, Massachusetts; [2]Imaging Science & Engineering Laboratory, Tokyo Institute of Technology, Japan; [3]Research Center for Frontier Medical Engineering, Chiba University, Chiba, Japan.

Context: The goal of multispectral imaging research is to develop a decision support system for pathology. The color correction for hematoxylin-eosin (H&E)–stained images by means of multispectral technique was proposed previously. However, it is difficult to introduce a camera system with a large number of bands to general pathologic institutions from the point of view of cost and operability. Therefore, this paper proposes a color correction method using a camera system with a small number of bands.

Technology: In this experiment, the multispectral imaging system that was used to capture the 16 band images is composed of 2000 × 2000 pixels for the charge-coupled device camera, 16 band rotation filters, a conventional optical microscope (Olympus BX-62) with an objective lens of ×20 where the light source is a halogen lamp, and a PC-based image capturing and displaying unit. The images with a small number of bands, for example, 3 band images and 6 band images, were generated from 16 bands images.

Design: In the color correction method, the transmittance spectra are estimated from multispectral images, and the amount of stain color pigment is calculated based on the Beer-Lambert law. It is possible to correct the color of tissue images of various staining conditions into an image of optimal staining condition by correcting the amount of coloring pigment calculated from multispectral images. Because the spectral characteristics of staining dye is affected by the staining condition, a sufficient number of bands, probably greater than 3, is necessary to appropriately estimate the spectral transmittance of stained tissue components. Thus, this paper proposes a color correction method wherein the matrix used for the Wiener estimation is adaptively selected depending on the staining condition of the image.

Results: Various H&E images of different staining conditions were corrected. The colorimetric error between the optimally stained tissue images and the corrected images reveals that the color of the corrected images closely matched that of an optimally stained tissue image.

Conclusion: The performance of the proposed method was investigated for H&E-stained images with respect to different staining condition and number of camera bands. The results indicate that the proposed method performs better than the color correction method with conventional Wiener estimation.

An Update on the Informatics Supported Annotation and Integration of Datasets From the Gynecological Disease Program (GDP)

Waqas Amin, MD[1] (aminw@upmc.edu); Sambit K. Mohanty, MD[1,7]; Anil V. Parwani, MD, PhD[3]; Sharon B. Winters, MS[1]; Nancy B. Whelan, BS[1]; Althea M. Schneider, BS[1]; John T. Milnes, BS[1]; Charma D. Chaussard[4]; Gail Harger, MS[2]; Katherine Farrow, BS[2]; Debra Bass, MS[2]; Tim D. Fennell, MS[1]; Hai Hu, PhD[6]; Thomas C. Krivak, MD[4]; Rajiv Dhir, MD[3]; Robert P. Edwards, MD[4]; Larry Maxwell, MD[5]; Michael J. Becich, MD, PhD.[1,3] [1]Departments of Biomedical Informatics and [2]Epidemiology, University of Pittsburgh, Pittsburgh, Pennsylvania; [3]Departments of Pathology and [4]Obstetrics and Gynecology, Magee-Women Hospital, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [5]Department of Pathology, Walter Reed Army Medical Center, Washington, DC; [6]Windber Research Institute, Windber, Pennsylvania; [7]Department of Pathology, Winthrop University Hospital, Mineola, New York.

Context: The Gynecologic Disease Program (GDP) is a tissue banking initiative funded by the Department of Defense. The primary objective of this program is to collect endometrial and ovarian tumor biospecimens and data from multiple institutions. The associated data collected from each site are submitted to the University of Pittsburgh and Ohio State University for quality review before being exported to a central data warehouse at Windber Research Institute (WRI), Windber, Pennsylvania.

Technology: The bioinformatics structure formulated for GDP aims to develop a well-characterized and high-quality biospecimen repository for ovarian and endometrial malignancies that facilitates the collection and transfer of well-annotated datasets to the central data warehouse at WRI with simultaneous storage of data in an internal storage system using a Java-based clinical trial management application (CTMA) developed at the University of Pittsburgh, Pennsylvania.

Design: Each site is responsible for data collection at the local level. At

Accepted for publication March 12, 2009.

Corresponding author: Michael J. Becich, MD, PhD, Department of Biomedical Informatics, University of Pittsburgh School of Medicine, UPMC Cancer Pavilion, Suite 301, 5150 Centre Ave, Pittsburgh, PA 15232 (e-mail: becich@pitt.edu).

the University of Pittsburgh, epidemiology data is collected using the HOPE Ovarian and Endometrial Questionnaire. Pathology data are acquired from CoPath Plus with treatment and outcomes data from the cancer registry.

Results: Patients have been consented for the GDP study since November 2003. The workflow of clinical annotation and data integration is as follows. At first encounter, patients are consented and enrolled in CTMA. Subsequently, the preoperative (case/control) questionnaire is administered by the research nurse coordinator and each case is allocated a deidentified number. Meanwhile, the surgically accessioned specimen is retrieved for pathologic examination and CoPath synoptic report is generated. Afterward the postoperative questionnaire is completed by the research nurse coordinator and copies are sent to the data managers. Lastly, the collected datasets are assembled, processed, and transferred electronically to the central data warehouse at WRI.

Conclusion: The GDP acts as a central repository for clinically annotated endometrial and ovarian tumor tissues for the research community. This tissue banking initiative provides an infrastructure of a joint multiinstitutional bioinformatics network that facilitates the sharing of clinically annotated data and high-quality biospecimens to support important research activities.

Quantitative Detection and Stratification of Lymphocytic Infiltration in Breast Cancer

Ajay Basavanhally, BS[1] (abasavan@eden.rutgers.edu); Shannon Agner, BS[1]; Anant Madabhushi, PhD[1]; Shridar Ganesan, MD, PhD[2]; Gyan Bhanot, PhD.[2] [1]Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey; [2]The Cancer Institute of New Jersey, New Brunswick.

Context: Molecular changes in breast cancer (BC) are often accompanied by changes in the tumor microenvironment. One such example is the presence of lymphocytic infiltration (LI), a histologic feature that has been recently been demonstrated to correlate with prognosis in early-stage human epidermal growth factor receptor (HER) 2+ breast cancer. The evaluation of LI in BC histology is usually measured in a qualitative manner that can have high intraobserver variation. A robust quantitative scheme would create a more precise stratification and allow for objective analysis of the extent of LI present in a tumor and its relation to clinical outcome measures. In this study, we develop computer algorithms to measure the extent of LI in BC.

Technology: We used 40 images of hematoxylin-eosin–stained histology obtained from 7 HER 2+ breast cancer samples. Each image was examined by an experienced clinician and scored as having low, intermediate, or high level of LI. All samples were digitized via a whole-slide scanner and analyzed with the MATLAB software package.

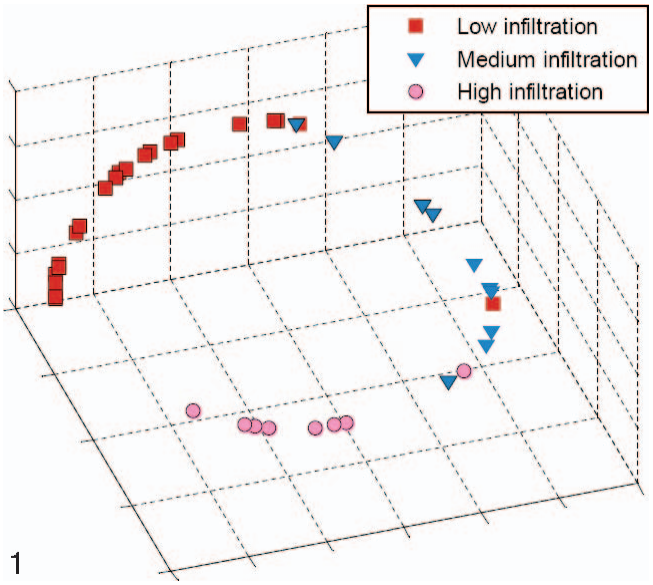
Design: We applied an automated segmentation scheme to find the centroid of each lymphocyte. We calculated graph-based features (Voronoi Diagram, Delaunay Triangulation, and Minimum Spanning Tree) from these centroids to quantify the architecture (ie, relative spatial placement) of lymphocytes. We calculated several features, including lymphocyte density, area disorder, and nearest neighbor statistics. Manifold learning techniques, such as graph embedding, allow us to visualize the stratification of samples in a low-dimensional feature space. We further apply a classification algorithm, known as support vector machine, to provide statistical results.

Results: The result of applying manifold learning to the graph-based features from manually segmented lymphocytes is shown in Figure 1. Each point represents a sample, and proximity between samples denotes similarity in reduced feature space. The low-dimensional manifold reveals the underlying structure of the data by showing the progression from low to high degrees of infiltration. Furthermore, randomized classification demonstrates an accuracy of 89.50% ± 6.22% for discriminating between low and high LI samples using manually segmented graph-based features.

Conclusion: We have presented an automated algorithm for detecting and classifying LI in high-grade BC. We have used graph-based features to exploit the architectural differences in arrangement between BC nuclei and lymphocytes. The graph embedding manifold shows the potential for using graph-based features in determining the degree of LI, rather than simply the presence or absence of LI.

A Multispectral Image Enhancement Approach to Visualize Tissue Structures

Pinky A. Bautista, PhD[1,3] (pinkybautista@gmail.com); Tokiya Abe, PhD[2,3]; Yukako Yagi, PhD[3]; John Gilbertson, MD[3]; Masahiro Yamaguchi, PhD[1]; Nagaaki Ohyama, PhD.[1] [1]Imaging Science & Engineering Laboratory, Tokyo Institute of Technology, Tokyo, Japan; [2]Research Center for Frontier Medical Engineering, Chiba University, Chiba, Japan; [3]Department of Pathology, Harvard Medical School, Boston, Massachusetts.



Context: Visualization of tissue structures necessitates the application of chemical stain. Popular for routine staining is the hematoxylin-eosin (H&E) stain, which generally differentiates acidophilic from basophilic structures. Initial investigations on the application of multispectral imaging to pathology has shown that with such imaging technology it is possible to visualize tissue structures that are ordinarily emphasized with special stains from an H&E-stained image.

Technology: In contrast to conventional RGB imaging, multispectral imaging uses more than 3 spectral filters to capture images such that salient information that can be useful for image analysis is contained in the images. An enhancement technique for stained multispectral images is introduced whereby tissue structures that are not clearly differentiated in H&E-stained images can be emphasized.

Design: PC vectors are derived from the transmittance spectra of tissue components that are not desired to be enhanced. Using only m-PC vectors the transmittance spectra of the multispectral image pixels were reconstructed wherein greatest error occurs for those tissue components whose transmittance spectra samples were not included in the set from which the PC vectors were derived. The spectral errors are then used to modulate the original transmittance spectra of the image pixels to produce varying color impression on tissue structures. A change in color from the original H&E-stained impression can be specifically observed on tissue structures that acquired high spectral errors.

Results: In our experiment multispectral images captured from the slides of kidney, liver, and heart tissue samples were used. The resulting enhanced images were compared with the real MT-stained images with respect to the delineation of collagen fiber from the rest of the tissue components. Initial results show the possibility of differentiating collagen fiber by applying the current enhancement method.

Conclusion: A multispectral enhancement framework has been introduced whereby the utility of multispectral imaging to visualize tissue structures with subtle spectral difference has been shown. In our future work we will consider the enhancement of other tissue structures that are emphasized by stains other than Masson trichrome and also the practical implementation of the current multispectral enhancement procedure.

A Histotechnology and Microscopy Curriculum and Web-Based Instructional Resource: Bridging Gaps

Philip J. Boyer, MD, PhD (philip.boyer@uchsc.edu); Yao Xu Schmidt, MS; B. K. Kleinschmidt-DeMasters, MD. Department of Pathology, University of Colorado Denver, Aurora, Colorado.

Context: The specialty of anatomic pathology employs a wide spectrum of histologic and microscopic techniques and tools with which few residents have familiarity upon entering residency. We have designed a curriculum that incorporates both Web-based instructional materials and reference to a standard textbook. The materials are compiled to provide a coherent curriculum to (1) meet the educational needs of medical students, residents, and fellows and (2) satisfy Accreditation Council for Graduate Medical Education (ACGME) guidelines with respect to docu-

menting and assessing formal instruction. Originally developed to supplement and provide an infrastructure for a series of 6 formal lectures, the curriculum should also provide an opportunity for focused self-learning and reference during diagnostic work.

Technology: Resources were compiled on a departmental Web server with access through standard Web pages with information stored as HTML documents, Adobe Acrobat PDFs, Articulate QuizMaker pretests and posttests, and Articulate Presenter voice-annotated Powerpoint files.

Design: Materials were grouped into 8 major topic areas: (1) basic histotechnology, (2) stains and histochemistry, (3) immunohistochemistry, (4) ''bug'' identification, (5) electron microscopy, (6) basic microscopy, (7) image capture, and (8) Photoshop for pathologists. Organized around defined knowledge and proficiency objectives, each topic area has a pretest and a posttest, reading assignment, Powerpoint file(s), text-based instructional content, PDFs of key references, and additional resources. Materials are accessible through a standard Web page. To comply with fair use restrictions, resources are currently accessible only through the University of Colorado Denver Intranet.

Results: This online resource is being deployed for use beginning with the 2008–2009 academic year, initially as part of a neuropathology curriculum, to supplement lectures, for independent study, and as a basic resource. Initial resident feedback on a β testing site has been overwhelmingly favorable.

Conclusions: This curriculum provides both instruction and easy access to resources with the goal of optimizing the learning experience and facilitating different learning styles. Knowledge-based objectives are matched with proficiency objectives. The curriculum is compiled within an ''ACGME wrapper'' and provides documentation of instruction and evaluation in a core pathology competency.

Software for the Laboratory: The Blood and Bone Marrow Counter

Alexis B. Carter, MD (abcart2@emory.edu); Stephanie Schniederjan, MD; Karen Mann, MD, PhD. Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia.

Context: Examination of bone marrow aspirates and peripheral blood with abnormal automated cell counts require manual differential counts of nucleated cells. Many laboratories still use push-button manual counters, the largest of which accommodates only 8 cell types and has widely spaced buttons that are not conducive to one-hand counting. Electronic counters are often expensive and have display screens of limited size, poor ergonomic design, and relatively few counting buttons.

Technology: A simple bone marrow and peripheral blood counting program was developed that could be executed on any computer running Microsoft Windows XP (Microsoft Corporation, Redmond, Washington). The program was constructed using Microsoft Visual Studio 2005 in Visual Basic.

Design: The program design allows the reader to manipulate the glass slide on the microscope using one hand while counting with the other hand on a nearby computer keyboard. Up to 20 different cell types may be counted, 5 of which are customizable. Cell counts are performed in 100-cell increments, and up to 10 sets of 100 cells may be stored in memory. Patient information can be entered, and printouts contain the raw data, calculated averages, and calculated ratios. For training purposes, the program includes help information, images, and descriptions of cell types.

Results: The software was compared with both manual and electronic counters in the laboratory at Emory University, Atlanta, Georgia. The software was faster and more user-friendly than counting with manual punch counters.

Conclusion: We have developed a peripheral blood and bone marrow cell counting program that is easy to use and has more capabilities than currently available manual and electronic counters. We have additionally demonstrated that simple software development can improve and facilitate routine laboratory procedures. Potential disadvantages of this software include an inability to interface with some laboratory information systems and lack of database functionality to maintain storage of differential counts after the program is closed. However, as laboratory information systems become more sophisticated and database functionality improves, the data could potentially be imported and stored. Advantages include those already described as well as distribution of the software as freeware.

Diagnostically Lossless Compression of Pathology Image Slides

Alexis B. Carter, MD[2] (abcart2@emory.edu); Saunya M. Williams, MS, PhD[1]; Nikil Jayant, PhD.[1] [1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta; [2]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia.

Context: Image compression enables efficiencies in the storage and communication of digital pathology content. While mathematically lossless compression provides a guarantee of high fidelity, pervasive telepathology may depend on a higher level of compression, for which we propose the guiding criterion of diagnostic losslessness (DL).

Technology: A SPOT Insight digital camera was used to capture 8 uncompressed images from slides of morphologically distinct lesions. MATLAB was used to perform JPEG compression on each digital photomicrograph. Blackboard Academic Suite was used to create and administer blind evaluations of images by 11 surgical pathologists and 8 pathology residents with at least 5 months of surgical pathology training.

Design: The subjective test was divided into 4 sessions. Each session included 2 different sets, with each set representing 5 differently compressed images. Each set was presented twice, with its individual stimuli in random order. Images were rated using an opinion scale: acceptable, somewhat acceptable, or unacceptable. The ratings were converted to a numerical scale of 3.0, 2.0, or 1.0, respectively. Subjects accessed the Blackboard via Web browser and also provided free-format qualitative inputs. All of the subjects, except one, completed the sessions on a single computer of their choice.

Results: An empirical threshold of 2.5 or higher was equated to DL. Of the 8 test images, 4 images (anaplastic thyroid carcinoma, brain blastomycosis, paraganglioma, Warthin tumor) achieved DL at one or more of the milder compression levels in the range (15:1–43:1). The uncompressed photomicrograph represented a digital rate of 24 bits per pixel. Poor color contrast, defocusing, and pixelation were common among the feedback associated with low ratings. The remaining 4 images (ovarian yolk sac tumor, pilar cyst, endometrial cystic atrophy, uterine leiomyosarcoma) failed the DL criterion even at the mildest compression level.

Conclusion: Based on the observed consistency in subjective inputs, the DL criterion seems to be a viable tool in optimizing image compression for pathology slides. Television-based or teleconferencing-based compression ratios of 20:1 to 50:1 seem to be too severe as a general rule for telepathology. A compression ratio of 10:1 seems to be a pragmatic and image-independent criterion to be validated in future testing.

Flexibility and Searchability: Support for Quantitative Analysis and Content-Based Image Retrieval of Tissue Microarrays

Wenjin Chen, PhD (chenwe@umdnj.edu); Vicky Chu; Jun Hu; Lin Yang; David J. Foran, PhD. Center for Biomedical Imaging & Informatics, The Cancer Institute of New Jersey, UMDNJ - Robert Wood Johnson Medical School, New Brunswick.

Context: High-throughput tissue microarray (TMA) technology is gaining wide acceptance by investigators throughout the cancer research community. Future progress in TMA-based research is reliant upon the capacity to archive, manage, and share TMA-related information among research groups and institutions. We recently developed a tissue microarray repository (TMR) module to organize and manage TMA-related information and images. The new TMR interface facilitates the population of distributed databases with new datasets including image metrics and correlated profiles in multiuser environments. While we are integrating the existing and emerging TMA data exchange standards into the TMR we are also establishing new meta data sets to support a range of image-based feature measurements that are being developed as part of a National Institutes of Health–funded research project.

Technology: The TMR framework is being developed to support seamless extensions, or plug-in applications, for performing quantitative image analysis while providing access to flexible database structure and image archive for storage and retrieval. The system is developed using Java and an Oracle 10 gigabyte database with a denormalized structure design.

Design: The TMR module is composed of a physical specimen layer, a digital specimen layer, and a quantification layer. The quantification layer is designed to provide flexibility in accommodating heterogeneous data forms that are generated from clinical, biologic, and image-based TMA research. In the development stages of each project, the researchers have maximal flexibility in designing and exploring different meta data forms. As the datasets and analytical tools continue to mature, these flexible structures can be easily migrated into their permanent forms so as to provide optimal efficiency in allowing Grid-enabled sharing and searching of these data and meta data.

Results: We have recently migrated into TMR, a content-based image retrieval application which enables one to query a gold-standard database of more than 3000 previously diagnosed breast cancer TMA discs and provide decision support for scoring and comparative analysis.

Conclusions: As the TMA-related common data elements are still being developed and evolving, the TMR framework provides a reliable solution for accommodating heterogeneous needs of the TMA research commu-

nity. Flexibility and shareability of the data are both taken into consideration in the design of the framework.

Use of Self-Configuring Lexical Analysis Approaches for Automated and Semiautomated Anatomic Pathology Data Extraction and Transformation

Jerome Cheng, MD (jeromech@umich.edu); Ulysses J. Balis, MD. Department of Pathology, University of Michigan Medical School, Ann Arbor.

Context: Legacy anatomic pathology (AP) information systems often lack sophisticated search, lexical extraction, and cross-load tools, thus creating barriers to effective utilization of archival information. While numerous technical approaches and solutions have been proposed and demonstrated for this need, none so far have made use of recent advances in self-configuring lexical analysis algorithmic/heuristic approaches. These are compelling, in that they allow for minimization or even elimination of incremental development and/or customization of standard lexical packages, in order to render a fully operable extraction pipeline. To overcome these historical limitations, we propose and demonstrate a self-configuring extraction-transformation-load (ETL) tool suite that avoids the complexity associated with requisite customization inherent in the conventional AP ETL turnkey solutions that have been reported to date.

Technology: Active state Perl, Visual Basic 6 (VB), PHP (hypertext preprocessor), HTML, SQL server 2005, with the predicate data source exemplar being AP data with Cerner Pathnet v3.06, were used.

Design: A CCL (communication command language) script was utilized to extract approximately 8900 AP cases as unformatted streaming text. Upon initial heuristic parsing of these datasets, self-configuring/dynamically adaptive heuristic lexical analysis methods were employed to identify an optimal/near-optimal set of hierarchical regular expressions that would allow for (1) separation of the text stream into case-level granularity and subsequently (2) further reduction of this case-level data into concept-level atomic data elements. The data pipeline included use of regular expression matching modules in Perl and heuristic lexical analysis and regular expression generation modules in VB. Resultant granular data was converted to CSV (comma separated values) file format and bulk inserted into the SQL database. Both VB and PHP/HTML-based search portals were created to peruse and validate the integrity of the extracted datasets.

Results: Heuristically derived regular expression patterns were successful in driving a PERL-based lexical extraction engine with this module being applied to the test set (~40 megabytes). Extraction times for the entire set were less than 10 seconds. Resultant dataset integrity was validated by use of a plurality of diagnostic term queries, initiated from both front-ends.

Conclusions: Self-configuring lexical analysis approaches hold significant promise for simplifying and automating the process of reliably extracting hierarchical, structured data from highly variegated and bulk text legacy AP repositories.

Tools for Manipulating JPEG-2000—Based Whole-Slide Image Formats

Toby C. Cornish, MD, PhD (tcornis3@jhmi.edu). Department of Pathology, Johns Hopkins University, Baltimore, Maryland.

Context: Whole-slide imaging produces large image files that are typically compressed and require specialized viewers. The popular Aperio SVS file format (Aperio Technologies, Inc, San Diego, California) is based on a tiled TIFF file structure and uses JPEG-2000 codestreams for image compression. JPEG-2000 is a standard (ISO/IEC 15444) for wavelet-based image compression and higher image quality at a given compression rate than most image compression algorithms, including JPEG. Unlike JPEG, though, the TIFF specification does not natively include JPEG-2000 image compression, and no widely available, nonproprietary utilities for directly decoding JPEG-2000 encoded TIFF files (ie, SVS files) exist. The native JPEG-2000 file format (JP2) is also used for whole-slide imaging, and programs for viewing and manipulating JP2 files, although rare, are available. This project aims to create free and open source tools for reading and writing SVS files, including a utility for the interconversion of SVS and JP2 file formats.

Technology: The software tools were implemented in C using these libraries: JasPer was used for encoding and decoding JPEG-2000 codestreams and JP2 files, LibTIFF was used to read and write TIFF files, and ImageMagick was used for image-processing functions.

Design: Only libraries with free and open source-compatible licenses were used. Additionally, portable libraries were selected that support compilation on all common, modern operating systems. The SVS file format was read by accessing the TIFF image file header and locating the image file directory for whole-slide image. The image file directory provides byte offsets and data lengths of the small tiles, compressed as JPEG-2000 codestreams, that comprise the full-size image. The tiles were copied directly or decoded at the desired resolution as needed.

Results: The software tools created here are capable of JP2 and SVS (version 9.1) file interconversion and direct access to SVS image tiles and meta data.

Conclusions: The availability of open source tools for manipulating whole-slide image formats facilitates the wider interchange of imaging data and enables the development of novel third-party software for the processing, analysis, and visualization of whole-slide imaging data.

Automated Detection of Basal Cell Keratinocytes for Quantification of Immunohistochemistry Biomarkers

James D. Deeds[1] (james.deeds@novartis.com); Lance Ostrom[1]; Dan He[1]; Christine Miller[1]; Colleen Conway[1]; Rebecca Mosher[1]; Jonathan T. Barron.[2] [1]Novartis Institutes for Biomedical Research, Cambridge, Massachusetts; [2]University of California at Berkeley.

Context: Skin is an increasingly popular surrogate sample for the evaluation of pharmacodynamic biomarker response following drug therapy. The basal keratinocyte population of cells is frequently the target of interest. However, this population makes up a small amount of a typical human or mouse skin biopsy. Because of this dilution effect and possible contamination by other cell types, histologic methods including immunohistochemical biomarkers are routinely used. Quantifying staining of this cell population can be done by manual scoring but is subject to intraobserver and interobserver variation. To achieve useful automated scores, the population of cells must be carefully selected to include only the population of interest. Accurately selecting this population by manual tracing is laborious and typically requires significant input by a pathologist.

Technology: Genie (Aperio/LANL), MATLAB, and ImageScope (Aperio) were used.

Design: We examine 3 methods to define basal keratinocyte regions of analysis used subsequently for automated scoring—manual tracing of regions as well as automated selection of regions using the commercially available Genie software and an in-house MATLAB program called PathTRAC. The 2 automated tools are general purpose pixel classifiers that use an initial training set of regions to develop algorithms that are then applied to the test set of whole-slide images to identify regions for analysis versus exclusion. The same training regions were used as input for each automated tool and different samples were used to test the accuracy of the resulting algorithms.

Results: We compare the relative sensitivity and specificity of each method using a set of images defined as ''ground truth,'' which have had regions for analysis and exclusion carefully selected by a pathologist (R. Mosher, MD). Skin slides taken from mice treated with a compound known to induce apoptosis were stained for the proteins Ki-67 and cleaved-caspase 3. The accuracy of pixel classification and the utility of nuclear counts for each method were assessed compared with pathologist-defined ground truth. A larger additional set of images without ground truth mark-up was also examined to compare the 3 methods to each other. In addition, we measured the amount of hands-on time required to perform each method.

Conclusions: The 2 tools provided variable sensitivity and specificity for the detection of keratinocyte containing regions when compared with pathologist-determined ground truth. The PathTrac application showed higher sensitivity (87% vs 68%) and higher specificity (44% vs 31%) with the cleaved caspase stained samples, whereas the Genie application yielded higher sensitivity (84% vs 65%) but lower specificity (46% vs 58%) with the Ki-67–stained slides. However, both methods provided positive nuclei counts comparable to those obtained with ground truth regions of analysis. Automated detection methods require significantly less hands-on time than manual selection approaches. Therefore, the use of general purpose pixel classification tools that can be rapidly trained by end users (eg, pathologists) could result in a significant increase of productivity for pathology image analysis. We are currently examining whether the addition of preprocessing and postprocessing operations (such as component size and binary ''erode/dilate'') to the existing analysis tools may improve specificity.

Use of Active Learning for Selective Annotation of Training Data in a Supervised Classification System for Digitized Histology

Scott Doyle[1] (scottdo@eden.rutgers.edu); Anant Madabhushi, PhD[1]; Michael Feldman, MD, PhD[2]; John Tomaszeweski, MD.[2] [1]Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey; [2]Department of Surgical Pathology, University of Pennsylvania, Philadelphia.

Context: Annotation of ''ground truth'' in a dataset is necessary for proper training of a supervised classification system. However, in the field of digitized histopathology, obtaining labeled ground truth is costly and time-consuming. Active learning is a method that intelligently chooses informative samples (rather than random samples) from a database for annotation. In this work, we present an active learning paradigm for the automated classification of prostate cancer from images of digitized histopathology. By selectively annotating training samples based on their uniqueness, the number of training samples necessary for accurate classification is reduced when compared with traditional randomized techniques.

Technology: A set of hematoxylin-eosin–stained prostate tissue slides are scanned into a computer using a whole-slide digital scanner. These images are analyzed using a set of routines implemented in the MATLAB software package. Ground truth is labeled by an expert pathologist using standard image-editing software (Aperio ImageScope).

Design: Our dataset consists of digitized high-resolution images of prostate histology. The goal of the supervised classification system is to detect which pixels in each image correspond to cancerous growths. From each digital image, a set of features is extracted, including Haralick cooccurrence features, Gabor filter features, and statistical greylevel features. Ground truth is annotated on each image by an expert pathologist, and from this pool of annotated data we construct a training set. In this study, we are comparing the accuracy of 3 different experimental setups: (1) the *randomized training set*, (2) the *active learning set*, and (3) the *control set*. To construct the randomized training set, we randomly sample the annotated data to generate 10 sets of training data, each of which is used to train a decision tree classifier. Each of the 10 decision trees casts a vote for the testing samples, so that each of the samples has between 0 and 10 votes for the ''cancer'' class. To construct the active learning set, we append the randomized set with the testing pixels that received an intermediate (between 2 and 6) number of votes. In this way, we are choosing to annotate ''informative'' or difficult-to-classify samples to improve classification accuracy. These are used to retrain the 10 decision tree classifiers and reclassify the images. Finally, we construct the control set by selecting samples that have either very low (0 or 1) or very high (7 through 10) votes for inclusion into the training set. By selecting these uninformative samples, we can ensure that any increase in accuracy is because of the inclusion of informative samples rather than the increase that would be expected when any data (informative or otherwise) are added to the training set.

Results: The results of the analysis of 3 test images are shown in the Table. We find that the classification accuracy increases when active learning is used to select additional training samples and that the increase is because of the informative nature of the samples themselves, as noninformative samples (the control group) show a smaller increase in accuracy. Shown in the Table is the number of training samples used in each of the 3 setups, along with the accuracy obtained using that training set, for each of the 3 images. Note that our goal here is to show trends in accuracy rather than perfect classification; clearly, a more robust classifier can be built with a larger dataset to annotate.

	Randomized		Active Learning		Control	
	Training Samples	Accuracy	Training Samples	Accuracy	Training Samples	Accuracy
Image 1	481	0.477	922	0.553	922	0.524
Image 2	750	0.610	1402	0.672	1402	0.620
Image 3	75	0.729	108	0.778	108	0.728

In each case, the classification accuracy goes up when active learning is employed to annotate informative samples for training. Further, the control group uses an equal number of training samples to the active learning group but achieves lower classification accuracy. This indicates that annotation of informative samples, rather than uninformative samples, is necessary to maximize classification accuracy.

Conclusions: In this study we have presented an active learning paradigm for the training of a CAD system. By selective annotation of difficult-to-classify samples, we can increase the accuracy of the system using fewer training samples than would be necessary in a traditional training paradigm. Because labeling digitized histopathology is costly and time-consuming, we must choose to annotate only informative samples in order to maximize the ratio of accuracy to training. Implementing active learning reduces the cost of obtaining labeled ground truth samples from a pathologist by reducing the overall number of training samples necessary to obtain high classification accuracy.

Preliminary Validation of a Multispectral Image Analysis Application for Confirmation of Isolated Tumor Cells in Axillary Lymph Nodes From Breast Cancer Patients

Jeffrey L. Fine, MD (finejl@upmc.edu); Kimberly McManus, HT(ASCP); Amber Luketich; David J. Dabbs, MD. Department of Pathology, Magee-Womens Hospital of the University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

Context: When examining axillary lymph nodes for breast cancer, small metastases can be difficult to confirm using traditional immunohistochemistry (IHC) methods, as the focus may not be present on the stained slide. Although alternative methods exist (eg, destaining), morphologic context of the original hematoxylin-eosin (H&E) stain may be lost. Multispectral image analysis is a technique that permits ''demixing'' of multiple stains on a slide. This study details preliminary validation of a special IHC procedure that generates virtual H&E and IHC images for the ordering pathologist.

Technology: Multispectral image analysis was carried out using the Nuance system (CRi, Woburn, Massachusetts). IHC staining for cytokeratin AE1/AE3 (Dako, Carpinteria, California) was performed using a Benchmark XT automated stainer (Ventana, Tucson, Arizona).

Design: H&E recuts were ordered on 5 axillary lymph nodes with macrometastases. IHC was performed directly on these H&E recuts, and additional stains were performed: traditional IHC, diaminobenzidine (DAB) only IHC (eg, IHC without counterstain), negative IHC controls, hematoxylin only, and eosin only. The eosin, DAB, and hematoxylin only slides were used to build a spectral library for image analysis. Composite H&E/IHC slides were spectrally ''demixed'' then recombined as false-color ''virtual'' H&E and IHC images.

Results: All IHC stains on all 5 metastases were validated manually. Composite H&E/IHC stains did not retain sufficient eosin for high-quality virtual H&E image reconstruction. Reapplication of eosin permitted virtual H&E and IHC stain images to be generated; images were validated against H&E and IHC slides. Even with all 3 stains (hematoxylin, eosin, and DAB), composite H&E/IHC slides did not require image analysis for validation.

Conclusions: This is a preliminary validation of our procedure to multiplex H&E and a single IHC stain as a clinical application. Following further validation with 20 known axillary breast cancer metastases, this service will be offered to pathologists for clinical use. Because of the ease of manual validation, this is an ideal introduction for this technology. Future related applications will focus on other very small foci that are difficult to stain (eg, microinvasion) and multiplexing of multiple IHC stains.

An Algorithm to Guide the Selection of Specific Biomolecules for Future Wet-Lab Experiments

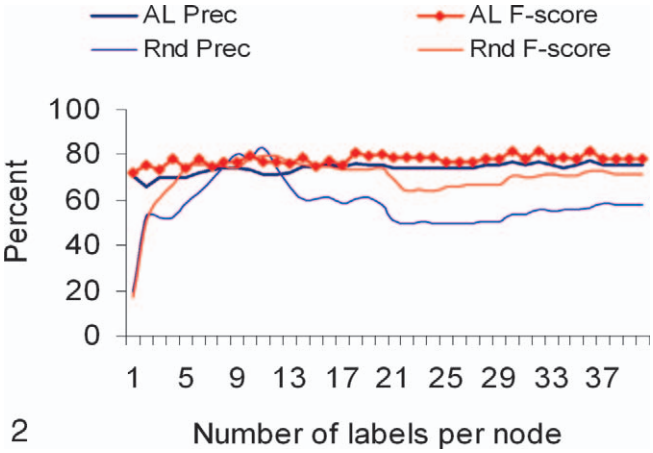
Madhavi Ganapathiraju, PhD[1] (madhavi@pitt.edu); Jessica Wehner, BS.[2] [1]Department of Biomedical Informatics and [2]Bioengineering and Bioinformatics Summer Institute, University of Pittsburgh, Pittsburgh, Pennsylvania.

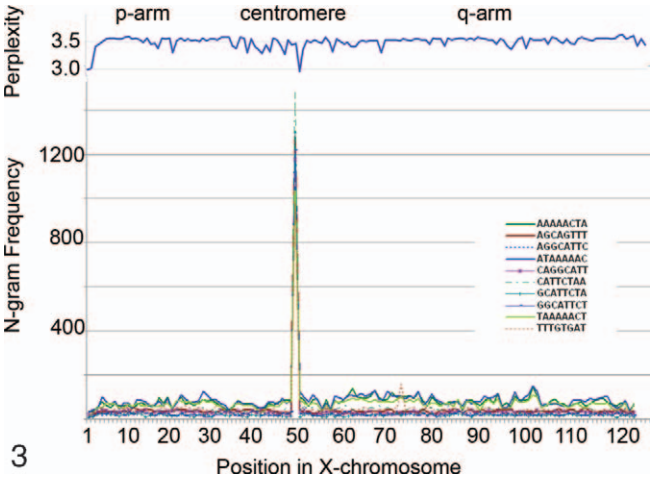
Context: With the large amounts of biomedical data available today, the few instances of labeled data are insufficient to confidently characterize remaining unlabeled data. Conversely, data characterization using wet-lab experiments is expensive in expert manpower, time, and resources. When choosing a specific biospecimen or molecule to be studied by wet-lab experiment, its redundancy with previously annotated (*labeled*) data is usually not taken into account. It is desirable to have algorithms that guide selection of data that is best suited for improving accuracy and confidence of labeling the remaining data. We designed an algorithm with active learning in application to a problem relevant to structure-based drug design. Membrane proteins (MPs) comprise 60% of drug targets. Prediction of transmembrane helix locations in MPs serves as the first step in computational modeling of their structure and thereby in structure-based drug design. However, experimentally determined structures are available for less than 1% of MPs. Using standard modeling techniques it is not possible to predict structures of novel MPs that lack a representative structure.

Technology: The algorithm is developed in MATLAB and the MATLAB Neural Network Toolbox.

Design: Feature vectors of MP sequences are derived as in TMpro. The data are clustered using a neural network-based self-organizing map of 5×8 nodes. The active learning algorithm is implemented for 2 scenarios: one in which localized data points within a molecule are selected to maximize disambiguation while minimizing redundancy, and a second scenario in which selection is permitted only at a whole molecule level, in which case coverage on unlabeled data is maximized.

Results: A novel active learning algorithm has been successfully de-





signed for this domain for nonredundant data selection with a gain in F-score and precision compared with unguided selection of training data (Figure 2). Using only 1% of available labels, an F-score of 80% has been achieved.

Conclusions: The algorithm can be translated to other domains, such as medical informatics and bioimage informatics to choose data selectively for manual or wet-lab annotation so as to accurately characterize the complete data while ensuring there is minimal redundancy.

Open Access Toolkit for Nonparametric Explorative Pattern Mining to Detect Events Relating to Disease in Large-Scale Genome Sequences

Madhavi Ganapathiraju, PhD[1] (madhavi@pitt.edu); Thahir P. Mohamed, BS[1]; Asia D. Mitchell, BS.[2] [1]Department of Biomedical Informatics and Intelligent Systems Program and [2]Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania.

Context: Events such as gene duplication, variations in tandem repeats, and abnormal methylation of CpG islands are often markers of diseases (eg, Alzheimer's disease, coronary heart disease, and cancer, respectively). Locating these events in individual genomes can allow diagnosis of these disorders, including the severity and age of onset. Currently there are no flexible tools for pattern mining on large-scale genome sequence data. We developed a suite of tools for large-scale explorative pattern mining on genome sequences and applied them to study human chromosomes X and 19. The tools allow the biomedical and genomics community to study genome and proteome sequences efficiently at many resolutions and in many flexible ways to draw inferences at a fast pace.

Technology: Algorithms have been developed in C and may be compiled and run on any Unix platform or on Windows platform with Cygwin software.

Design: The genome sequence is preprocessed into an efficient data structure called suffix array, and its well-known augmentations, the longest common prefix array and rank array. The tools rely on computing perplexity, a measure of how predictable the nth nucleotide is, given the $n-1$ preceding nucleotides. Repeat rich regions are indicated by a drop in perplexity.

Results: The tools can perform different types of analyses: (1) find locations and characteristics of repeats; (2) compute complexity along a sequence; (3) compare how similar or dissimilar 2 sequences are overall; and (4) perform other innovative n-gram analyses. When applied to the human X chromosome, it revealed the repeat rich p-arm and centromere (indicated by dips in perplexity in Figure 3). Upon closer analysis of these regions (windows 47–50), it revealed highly abundant n-grams in this region (see high peaks of n-gram counts in this region in Figure 3), which are rare in the rest of the chromosome. Analysis of chromosome 19 showed richness of repeats, a possible location of centromere, and at least one pair of reverse complements containing repeats. The tools are nonparametric, efficient, and scalable and can aid discovery of patterns that have biomedical significance.

Conclusions: The toolkit, which will be released in open access with complete functionality, can be applied to discover a number of genomic events.

Validating Currently Available Digital Measuring Tools for Melanoma

Gabor Hertz, MD (gabor.h.hertz@kp.org). Department of Pathology, The Permanente Medical Group, Sacramento Medical Center, Sacramento, California.

Context: Skin cancers account for the most common neoplasms seen in general clinical practice. Unlike most of these cancers (basal cell carcinoma and squamous cell carcinoma), melanomas have a significant mortality rate and have a reported incidence of 16 to 24/100 000. Treatment and prognosis of melanomas depends on the pathologic stage. Pathologic staging depends on measuring the Breslow tumor thickness. Traditionally this has been obtained by using an eyepiece reticule calibrated by stage micrometer. Pathology informatics/digital imaging may provide a more reproducible measurement. To that end, this study attempts to validate digital measuring tools available in an off-the-shelf product for use in staging melanomas.

Technology: Olympus BX40 microscope, Olympus DP 71/DP 25 digital cameras using either FireWire or PCI connections, Olympus 0.01-mm stage micrometer, and Olympus Micro-Suite 5 software (Melville, New York) were used.

Design: Validation study used a stage micrometer standard and retrospective review of 18 randomly selected cases from our files during the last 2 years. The slides were reviewed, digitally imaged, and measured by a pathologist. Digital measurements were obtained at varying optical magnifications, and the mean measurement was compared with the reported clinical Breslow measurement.

Results: Validating the digital measurements to the stage micrometer found a disagreement averaging 25% (up to 0.3 mm) from the expected digital measurement. In reviewing the 18 cases from our files, a digital measurement was successfully obtained on all 18 cases with an average discrepancy of 0.12 mm between the analog and digitally obtained measurements.

Conclusions: This validation study showed that the digitally obtained results varied up to 25% from the stage micrometer standard. This discrepancy may be a result of the stage micrometer or technical problems with either the initial software installation or a mismatch between the optical system and software. These factors need to be further elucidated prior to clinical use of this measuring tool. Interestingly, in the 18 retrospective clinical cases, there was tighter correlation between the analog and digitally obtained measurements (averaging 0.12 mm). This tighter correlation is probably due to multiple additional confounding factors inherent in pathologic staging that should be mitigated by the documenting abilities of digital imaging.

Tipping Points in the Adoption of Digital Pathology in Clinical Settings

Jochen Lennerz, MD, PhD[1] (JLennerz@path.wustl.edu); Michael Isaacs[1]; Lloyd J. LaComb, PhD[2]; Michael R. Descour, PhD.[2,3] [1]Department of Pathology and Immunology, Washington University, St Louis, Missouri; [2]DMetrix, Inc, Tucson, Arizona; [3]College of Optical Sciences, University of Arizona, Tucson.

Context: Whole-slide scanning systems have enabled the introduction by early adopters worldwide of digital pathology into the practice of surgical pathology and thereby the clinical environment. The tipping points related to digital pathology in this environment are improved patient care, interaction with a laboratory information system, scanning hardware and software integration, pathology laboratory workflow integration, available information-technology infrastructure, data management, storage and mining, and cost of ownership.

Technology: Digital pathology presents a unique challenge in microscopy because it departs from the conventional microscopy paradigm of trading image detail for field of view. This new challenge requires novel approaches to high-throughput image capture that can be divided into serial and parallel methods. We have experience with both methods. In this presentation, we extrapolate from today's state of the art instruments into the near future. Our extrapolation is based on careful review of a wide range of imaging systems that deliver fine detail over a large area, for example, lithography projection cameras, and the overall technical system requirements dictated by rapid, sustained image capture.

Design: The review of current imaging and related technologies is performed against a backdrop of a 1000 slides per day benchmark. We present a scaleable analysis of implementing a digital pathology solution with such capacity. While holding the benchmark constant, we analyze the impact of a digital pathology solution extrapolated from technology trends observable today.

Results: Our review of imaging technologies shows that significant advances in throughput can be expected to occur in the next 2 to 5 years. Faster imaging also means that digital slides can be a more lifelike representation of the glass slide, without an impractical sacrifice in throughput. However, technology constraints suggest that control of the imaging ''robot'' by a laboratory information system through slide-specific imaging instructions will be required.

Conclusions: High throughput and integration with information systems (eg, laboratory information systems and picture archiving and communications systems) are required to promote the further adoption of digital pathology in the clinical environment. Throughput in excess of 60 slides per hour is feasible and can be expected to advance adoption of digital pathology through reduced cost of ownership.

Assigning Putative Protein Identifications to Selected Lung Cancer Biomarkers From Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry of Blood Serum

Jonathan L. Lustgarten, MS[1] (JLL47@pitt.edu); Vanathi Gopalakrishnan, PhD[1]; David Malehorn, PhD[2]; William Bigbee, PhD.[2] [1]Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania; [2]University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania.

Context: We present a computational method for assigning putative protein identifications to mass spectral features to enable selection of biomarkers (mass-to-charge values) for further experimental validation. We test the method on proteomic mass spectra derived from surface-enhanced laser desorption/ionization time-of-flight analysis of blood serum samples from healthy controls and patients with lung cancer. We demonstrate that along with the identification, we can provide information on sequence, disease association, fragments, and other proteomic mass-to-charge peaks that may have been found using different technologies for the corresponding protein. It also could be used as a guide toward designing and prioritizing validation experiments.

Technology: A dataset of proteomic mass spectra consisting of 322 samples from the Lung Specialized Program of Research Excellence (SPORE) registry were scanned into the computer and then processed using our in-house rule learning program. Discriminative markers of lung cancer as detected by this program were provided to the Empirical Proteomics Ontology Knowledge Base (EPO-KB) (http://www.dbmi.pitt.edu/EPO-KB) to obtain putative identifications based on the literature.

Design: We used rule learning to extract mass-to-charge values that the algorithm deemed statistically significant from each of the 10 runs of 70/30 split of the training data, where 70% was used for training and the remaining 30% was used for testing. We calculated the steadiness of each discriminative biomarker by a simple fraction of how many times it appears in the 10 different models. We then assigned identifications to those mass-to-charge values that appeared at least twice by using EPO-KB.

Results: The results of the analysis chose 10 mass-to-charge values of which 4 appeared in 6 or more models. We were able to assign 9 identifications using the EPO-KB. The assigned putative identifications consisted of proteins that were well known for their appearance within cancer.

Conclusion: We have presented a system that uses an advanced machine learning technique to choose significant biomarkers and then assign putative identifications from prior knowledge to assist researchers in the selection and validation of biomarkers for lung cancer.

This research is partly funded by grant P50 CA090440 07 from the National Cancer Institute and GM071951 from the National Institute of General Medical Sciences.

Use of a ''Mathematical Microscope'' to Understand Radiologists' Errors in Breast Cancer Detection

Claudia Mello-Thoms, MSEE, PhD (mellothomsc@upmc.edu). Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania.

Context: Excluding cancers of the skin, breast cancer is the most common type of cancer for women in the United States. Mammography screening is the only method proven to reduce mortality from this disease, but because of several factors, between 10% to 30% of breast cancers are still missed at screening. These misses have been divided into (1) search errors (lesions did not attract any amount of visual attention); (2) perceptual errors (lesions attracted visual attention but not long enough for object recognition); and (3) decision-making errors (lesions attracted visual attention for a long period but were ultimately dismissed). Perceptual and decision-making errors account for 70% of all misses. However, we need to determine how these lesions differed from correctly reported breast cancers.

Technology: Using a ''mathematical microscope,'' a wavelet packets transform, we have derived a model that characterizes, like the human visual system, each area fixated. The basic steps of the decomposition are (1) area segmentation into squares measuring 5° of visual field (size of the fovea); (2) wavelet packets filtering; (3) feature extraction; and (4) normalization. Each miss and correctly reported cancer is characterized by this local representation and by a representation of the background areas used to support the decision.

Design: Four MQSA (Mammography Quality Standards Act)-certified radiologists read 40 two-view digitized mammogram cases, of which 30 cases contained malignant masses. Eye position and decision response were recorded. Cancers that were correctly reported were contrasted with those that attracted visual attention but were not reported.

Results: Analysis of variance of local representation of cancers that were correctly reported and those that yielded perceptual or decision-making errors do not yield statistically significant differences, although differences between correct reports and perceptual errors reach borderline significance for certain orientations/spatial frequency ranges (eg, Scheffe posthoc test, $P = .05$). Analysis of background areas used for decision support show no difference in search strategy in case of perceptual errors (indicating no object recognition) and statistically significant differences after fixating both correctly reported cancers and decision-making errors.

Conclusions: Local feature analysis often cannot explain why lesions were missed, but a global analysis that takes into account support information can characterize why errors occur.

Token Swap Contingency Tables in 3 Dimensions: Paradigm for Biomedical Data Analysis
http://www.netautopsy.org/tokncube.htm

G. William Moore, MD, PhD[1,2,3] (George.Moore4@va.gov); Grover M. Hutchins, MD[3]; Lawrence A. Brown, MD.[1,2] [1]Pathology and Laboratory Medicine Service, Veterans Affairs, Maryland Health Care System, Baltimore; [2]Department of Pathology, University of Maryland Medical System, Baltimore; [3]Department of Pathology, The Johns Hopkins Medical Institutions, Baltimore, Maryland.

Context: Contingency tables are commonly used for organizing frequency data on biomedical databases. Classical statistical methods applied to contingency tables include the χ^2 test and Fisher exact test, based upon squared-normal and binomial distributions. In the token swap method, patients, or tokens, in the contingency table are randomly swapped to determine whether observed data deviate from a preset null hypothesis.

Technology: We used Perl programming language and theory of statistics.

Design: The simplest contingency table is a rectangular table, consisting of 4 cells, 2 rows by 2 columns, that measures association between row and column variables in a misclassification space. The null hypothesis predicts expected values for each cell; tokens are randomly swapped until they match observed values. More generally, a 3-dimensional contingency table has rows, columns, and depths, representing a variable for ultimate biomedical outcome.

Results: The 2- and 3-dimensional token swap methods satisfy the Neyman-Pearson condition for power of the alternative hypothesis. Unlike classical methods, the token swap method supports a range of null hypotheses, including those with zero cell totals.

Conclusion: The present model extends the range of existing contingency table analysis to incorporate additional clinicopathologic information and to explore customized null hypotheses.

Implementing Rapid Real-Time Process "Defect" Collection to Optimize Anatomic Pathology Workflow: A 2-Pronged Approach Based on End-User Work Setting

Michael Riben, MD[1] (mriben@mdanderson.org); Leslie Nesbitt[1]; Shibu Ninan[3]; Mark J. Routbort, MD, PhD.[2] [1]Departments of Pathology and [2]Hematopathology and [3]Section of Clinical Laboratory Informatics, Division of Pathology and Laboratory Medicine, University of Texas, M. D. Anderson Cancer Center, Houston.

Context: Our workflow optimization project leverages manufacturing-based quality improvement methods, such as Lean and Six Sigma, which strive for zero ''defects'' as a measure of quality. We have implemented a change management infrastructure that utilizes data collection and analysis to aid decision making, targeting sources of waste and process defects that affect efficiency and quality. We define defect using a definition from D'Angelo et al published in 2007. Like these authors, we developed a defect collection system that met several criteria: it was easy to use and provided real-time data capture, equal access to all, standardized, menu-driven, defect capture closest to discovery, visual presentation and public exposure, blameless participation, compliance with participation, and reusability.

Technology: We utilized paper-based customized Post-it (3M) notes with a 4-quadrant design and developed a custom data entry form in PathStation, our VB.Net/SQL based workflow integration software application.

Design: A coded defect classification and location/process list was developed and distributed to all employees. The data collection event lasted for 12 working days. For non–computer-based workstations, defects were captured in real-time and completed Post-it notes were posted in each lab at designated wall locations. Notes were collected every other day for data entry into the electronic module. For computer-based settings, an integrated defect collection module automatically set user context, case context, date/time, and user location and utilized an identical defect classification. A running total was always visible in the module. All entered defects were visible on a display tab. Ultimately, all defects were recorded into the electronic module to facilitate visibility to the widest possible audience and allow for data analysis.

Results: We collected 1291 defects. The range of defects were then analyzed and classified by multiple schemes, including location, process steps, reporting locations, staff type, and workflow (preanalytic, analytic, and postanalytic).

Conclusions: The ability to document defects in real-time enhances reporting opportunities and increases compliance with data gathering events critical to designing improvements that optimize efficiency and quality. We demonstrate the ability to document the defect identified in less than 5 seconds in settings with and without a computer workstation.

Use of a Repetitive Task Scheduling Engine for Workflow Automation and Rare Event Detection in a Clinical Environment

John Sinard, MD, PhD (john.sinard@yale.edu); Peter Gershkovich, MD, MHA; Wolfgang Freis; Agatha Daley. Department of Pathology, Yale University School of Medicine, New Haven, Connecticut.

Context: Maintaining the information technology infrastructure necessary to support the clinical operations of pathology departments is becoming increasingly difficult for already overburdened support staff. In addition, systems analysis and quality assurance create a growing number of assertions that have to be manually verified. The ability to off-load repetitive tasks to automated systems can free up needed support staff, provide more timely notification of actionable failures, and detect rare events that would not be practical to explore manually.

Technology: We have built a repetitive task scheduling engine based on the open source quartz scheduling engine from OpenSymphony. The system is written in Java and is deployed as a Web application running in a Tomcat servlet container under a Macintosh operating system. The Web interface is built using Google Web Toolkit.

Design: Each individual task is structurally defined as a job. It has a trigger, a job type, and an action. The trigger determines how frequently the job runs. The job type corresponds to a Java class, which performs the function of the job, determines the result, and activates the appropriate action, which is typically sending an e-mail or a page.

Results: We have successfully used this system to monitor daily backups, monitor the online status of various departmental Web servers, detect billing batch failures, file and manage images and scanned documents, distribute voice dictation files, run and distribute overdue case reports, detect video server failures, verify the online status of other Web applications, and detect the rare accessioning of procedures to cases associated with inactive clients. Errors needing immediate intervention are communicated to information technology support staff by paging. Less serious errors are e-mailed to support staff. Other events are logged in a log file.

Conclusions: Automating several repetitive tasks and targeted notification has resulted in more rapid response to events, more reliable handling of routine processes, and detection of rare occurrences that may otherwise have gone undetected. This solution is highly scalable. Support staff can concentrate their efforts on solving problems rather than on performing routine tasks or looking for problems of which they are not yet aware.

Semiautomated Archiving of Scanned Requisition Documents in Anatomic Pathology

John Sinard, MD, PhD (john.sinard@yale.edu); Peter Gershkovich, MD, MHA. Department of Pathology, Yale University School of Medicine, New Haven, Connecticut.

Context: The vast majority of anatomic pathology laboratories still receive paper requisition forms with specimens, in part because many specimens originate in physician offices that are not electronically connected to the pathology laboratory. Cataloging, sorting, copying, distributing, resorting, and storing requisition forms can consume many hours of labor. Scanning these documents into electronic images can facilitate their distribution, availability, and storage but requires properly indexing and electronically filing each image so that it is associated with the correct specimen and patient.

Technology: We streamlined the filing and indexing of scanned document images using barcoded labels and custom software modules written in Java and Perl, which integrate with a C-based barcode decoding library from Softek Software. The software runs in the background processing incoming documents and passing them to our repetitive task scheduling engine for filing. All components run under a Macintosh operating system.

Design: A customization in our laboratory information system (CoPath Plus; Cerner, DHT) prints barcoded specimen labels. These are affixed to any paperwork that is to be archived. Bulk paperwork is then scanned, without prior sorting, to a network hard drive using a scanner-copier with a sheet feeder. Custom background software then opens each document file, locates and reads the barcode, converts the image to a JPEG file, renames the file according to departmental file naming conventions, and files the document into a repository that is available through a Web-based interface.

Results: We have successfully used this system to store more than 300 000 specimen requisition forms, frozen section forms, and outside consult letters. To the best of our knowledge, no barcode has ever been misread. Initially, ~5% of the scanned documents failed to file automatically because the barcode could not be read. In most of these cases, no barcode label was placed on the document before scanning. This number has significantly decreased with training, and the current failure rate is about 1%.

Conclusions: Automating the scanning and filing of requisition images has significantly decreased the need for repeated sorting of paperwork and has eliminated the need for sending this paperwork for microfilming. Insurance claims reconciliation has also been greatly facilitated.

Multispectral Whole-Slide Imaging System

Mitsuyoshi Tashiro[1] (tashiro.m.ac@m.titech.ac.jp); Noriaki Hashimoto[1]; Yuri Murakami[2]; Takashi Obi[1]; Masahiro Yamaguchi[2]; Nagaaki Ohyama[2]; Yukako Yagi.[3] [1]Interdisciplinary Graduate School of Science and Engineering and [2]Imaging Science & Engineering Laboratory, Tokyo Institute of Technology, Tokyo, Japan; [3]Department of Pathology, Harvard Medical School, Boston, Massachusetts.

Context: Applications of multispectral imaging such as quantification of dye amount, adjustment of stain condition, digital stain, and color enhancement processing have been developed for pathologic diagnosis support. Recently, whole-slide imaging (WSI) has been widely promoted and is expected to be installed for pathologic image analysis based on multispectral information.

Technology: A multispectral image (MSI) is one of the spectral images captured by using N color filters, $N > 3$, where each image pixel contains spectral information that can be used for image analysis. Apparently the data size of a MSI is larger than that of a RGB color image. In addition, there is a need for conversion processing to display a MSI as a color image for ordinary observation. For this reason, it is difficult to implement interactive WSI simply by replacing the stored RGB images with MSIs.

Design: To address the setback in implementing interactive WSI, we proposed the following system. First, the equivalent RGB image and other

spectral information extracted from an MSI are saved in the server separately. The RGB images are used for ordinary observation of the image data and the additional spectral information is only sent when detailed image analysis is needed. This process therefore enables the implementation of multispectral analysis.

Results: We implemented the adjustment of stain condition of hematoxylin-eosin–stained images through MSI processing using the additional spectral information. In this case, we ensured the fast implementation of dye adjustment by modifying the saving format of the additional spectral information.

Conclusions: We proposed multispectral WSI that enables the application of image processing techniques based on spectral information. In our future work, we would evaluate other image processing applications, that is, digital staining, color enhancement, and examine the compression ratio and accuracy of the results using additional information.

The Development of Workflows for Image Analysis of Tissue Stained With Specific Antibody Markers

Mark R. Verardo, PhD (mverardo@dako.com); Lee Ryan; Tony Lacroix; Julie M. Smith, MA; Joachim Schmid, PhD. Imaging and Pathology Workflow, Dako North America, Carpinteria, California.

Context: The foundation of image analysis applications developed for routine diagnostic pathology markers are algorithms for quantitative image analysis. These applications are dependent upon tissue specimens stained using a standardized and automated workflow model. The workflow model is composed of antibody and visualization kits, automated immunohistochemistry staining protocols, and image analysis applications. Data output, or image measurements, are captured in standardized reports for target antibodies. Marker-specific workflows can be developed as tools addressing a major issue in the pathology laboratory, standardization.

Technology: Automated staining was performed using a Dako Autostainer Plus Link (Dako, Carpinteria, California) instrument, and digital microscopy was performed with the Dako ACIS III Automated Cellular Imaging System (Dako). A marker-specific imaging application (not currently available worldwide) was used for image analysis.

Design: Tissue samples were provided by the Cooperative Human Tissue Network, which is funded by the National Cancer Institute. Tissue was stained with a specific antibody using an automated workflow model. Specifically, Ki-67 was detected on breast cancer specimens using the Dako Ki-67 RTU Antibody (Clone MIB-1, Dako) and visualized with the Dako Flex Detection Kit (low pH TRS). Images of the tissue, captured by digital microscopy, were scored or graded by a board-certified pathologist who identified marker-specific regions of interest. An algorithm was developed that was antibody- and stain-specific and a pathologist's report was created.

Results: An automated workflow model was successfully used to develop an antibody-specific application that allowed standardization of immunohistochemistry staining and imaging resulting in reproducible, quantitative measurements that could then be incorporated into a final report. Important components in the development of this application included staining protocols, image analysis algorithms, and interactions with a board-certified pathologist.

Conclusions: The development of an automated and efficient marker-specific workflow can aid in quantitative measurements from images captured by digital microscopy. An example of this model has been developed for Ki-67. Standardized antibody detection, image capture, and image analysis are key components to producing an effective, reproducible tool for routine diagnostic and prognostic pathology markers.

caTissue Suite: An Open-Access, Feature-Rich Tool for Biospecimen Annotation and Data Sharing

Mark A. Watson, MD, PhD (watsonm@wustl.edu). Department of Pathology and Immunology, Washington University School of Medicine, St Louis, Missouri.

Context: Advances in molecular technologies and clinical trial design have increased the importance of biospecimen resources and the requirements associated with the collection and storage of human biospecimens. caTissue Suite is a caBIG application designed to manage the associated complexities of biospecimen annotation data.

Technology: JBOSS 4.0.0 (Red Hat Middleware, LLC), Java Development Kit 1.5 (Sun Microsystems, Inc), Apache Ant 1.6 (The Apache Software Foundation), and either MySQL 4.1.19 (MySQL Inc.) or Oracle 9i (Oracle Corporation) were used.

Design: caTissue Suite is an n-tiered application. A Web browser submits requests to the application server, which in turn persists or acquires data in a relational database. A Java-based application programming interface permits more advanced, automated, and customized access to all of the application's features. The application supports administrative functions (protocol and storage system definition), biospecimen accessioning (including provisions for consent tracking and iterative biospecimen derivation and aliquoting), and investigator queries (advanced query creation and specimen requisition system). A caTIES (Text Information Extraction System)-like interface allows for import and concept coding of textual based pathology data, and discrete pathology and clinical data entry is supported through customized form creation.

Results: caTissue Suite is scalable, configurable, and flexible enough for broad deployment across repositories of varying size and function. To date, at least 8 institutions have adopted the application and are using it in the daily operation of their biospecimen resource. A grid-based biospecimen data sharing initiative is underway. Continuous end-user feedback provides new requirements for each iteration of the application release. Planned enhancements in caTissue Suite v1.1 (December 2008) include the ability to use a single instance to independently manage multiple biospecimen repositories, to provide support for temporal queries (eg, find all biospecimens that were frozen within 30 minutes of collection), and functionality to create and use complex data forms dynamically from a new unified modeling language-based object model.

Conclusions: caTissue Suite is an open-access/open-architecture application that has been developed, formally tested, and successfully deployed by a number of National Cancer Institute–designated cancer centers and other biospecimen resources. Use of caTissue Suite by these institutions is providing a rapid and facilitated path toward standardizing biospecimen informatics both nationally and globally.

Pathtracker, a Software Program That Uses Commercial Digital Cameras to Track the Movement of Biospecimens and That Has E-Mail Reminder Capabilities

William H. Yong, MD (wyong@mednet.ucla.net); Philippe Taieb; Victoria C. DeGuzman; Steven S. Silver, BSc. Brain Tumor Translational Resource, UCLA Neuro-Oncology Program and Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California.

Context: One critical function of a pathology laboratory or a tumor bank is the tracking of materials and specimens to and from its site. Barcodes and radio frequency identification tag systems may be costly or not available. Our objective is to develop a relatively inexpensive method to improve biospecimen tracking and to provide prompts for completion of related tasks.

Technology: We used Visual Basic 6, Access 2003, Windows XP (Microsoft, Redmond, Washington); Intel Pentium dual core computer with 2 gigabytes of RAM (Dell Computer, Round Rock, Texas); and Canon Powershot SX100IS, Canon Powershot S2IS (Canon USA, Lake Success, New York).

Design: Visual Basic 6 is used to link a Microsoft Access database with a data entry graphical user interface and an inexpensive commercially available camera.

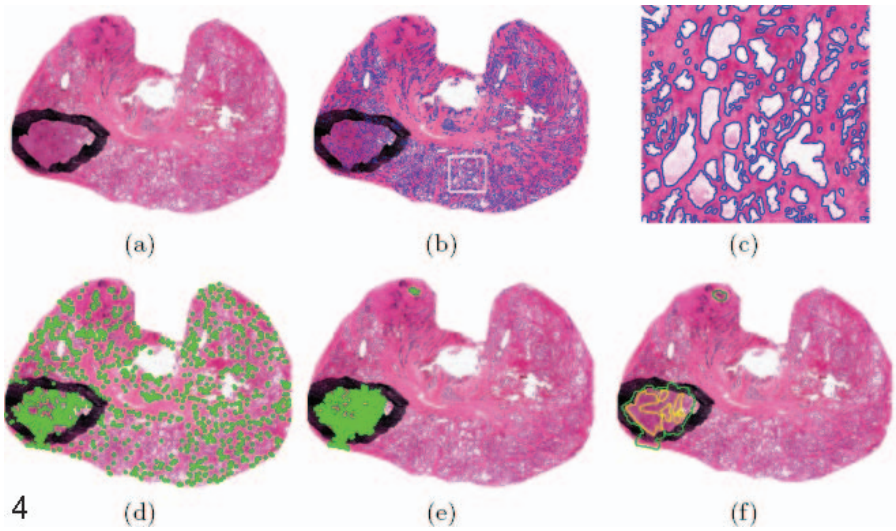
Results: A Canon Powershot S2IS camera and subsequently a Powershot SX100IS have been successfully linked to the software. The digital cameras can capture alphanumeric labels on slides and blocks. One image can document up to 20 slides in a tray or 90 blocks. The images are linked to a case. A search function can find cases using specified parameters including name, destination, and date range. Automated e-mail reminders are sent when materials or tasks are due. The software can also send sequences of reminder e-mails leading up to a due date. This feature has been adapted for prompting completion of autopsy reports. This image-based tracking has been successfully implemented in the UCLA Brain Tumor Translational Resource, a brain tumor biorepository.

Conclusion: We have developed a flexible software program that uses relatively inexpensive cameras to document the movement of materials. It is useful as an adjunct to other tracking technologies or as a substitute. The e-mail reminders can prompt return of materials but is easily adaptable to prompting completion of autopsy reports.

Outcome Assessment of Computerized Speech Recognition to Validate its Application in Surgical Pathology

Lanjing Zhang, MD, MS[1] (lanjing.zhang@mssm.edu); Divya Seth[2]; Margaret Huynh[2]; Jeffery S. Van Vranken, BA[2]; Zhenhong Qu, MD, PhD.[2] [1]Department of Pathology, Mount Sinai Medical Center, New York, New York; [2]Department of Pathology and Laboratory Medicine, University of Rochester Medical Center, Rochester, New York.

Context: Although computerized speech recognition (CSR) has reportedly matured into a practically useful tool in medical documentation,



4

widely diverse opinions on its usefulness or lack of it in pathology service still remain. This appears to stem from the lack of evaluation/appraisal based on objective outcome measurement. This study is designed to evaluate the utility and validity of CSR in routine surgical pathology service.

Technology: CSR is tested and established. Four trained users are recruited. Documentation time (DC; ie, the time required to complete the documentation task) for gross examination and reporting time (ie, the time required to review and sign out the case) by CSR and transcriptionist-mediated dictation (TMD) are compared. The efficacy of CSR is validated by time and effort (and accuracy) compared with that for TMD.

Design: For gross examination, the DC for a total of 83 commonly encountered surgical specimens by CSR (44) and TMD (39) were collated. No grossing templates were used for CSR. The DC by different methods for similar specimen types was compared. The documentation adequacy is measured by the presence of key words against a third-party's gross examination manual. The reporting times of 445 biopsy cases by these 2 methods were separately collected and compared. Unpaired 2-tail Student *t* test was used for statistical analysis.

Results: CSR and TMD required equal DC (n = 18 vs 6, mean ± SD = 13.20 ± 3.98 vs 12.20 ± 6.02 seconds per word) for large resection specimens. Interestingly, more or equal DC for small specimens was required for CSR compared with that by TMD (n = 26 vs 33, mean ± SD = 5.12 ± 2.13 vs 2.95 ± 1.50 seconds per word, $P < .01$). The mean reporting time of biopsy cases by CSR (n = 201) was 338.8 ± 75.4 seconds per case or 77.3 ± 20.2 seconds per slide, and that by TMD (n = 244) was 345.1 ± 81.9 seconds per case or 80.7 ± 18.3 seconds per slide. No statistical significance was established between reporting time by CSR and TMD ($P > .05$). A documentation accuracy of more than 95% was reached by either CSR or TMD.

Conclusions: The efficacy and utility of CSR depend on the types of surgical specimens to be documented. For gross examination, CSR appears to require equal or more DC than TMD to complete small resection specimens (mean, 5.12 vs 2.95 seconds per word), while equal DC is required to complete large specimens (mean, 13.20 vs 12.20 seconds per word). In reporting biopsy cases, CSR takes the same or even less time than TMD (mean, 77.3 vs 80.7 seconds per slide). In these 2 settings, CSR may significantly shorten report turnaround time and reduce cost by eliminating the transcription step by office staff.

A Computer-Aided Diagnosis System for the Automated Detection of Prostate Cancer on Whole-Mount Histology Images

James Monaco, PhD[1] (jpmonaco@rci.rutgers.edu); Anant Madabhushi, PhD[1]; Michael Feldman, MD, PhD[2]; John Tomaszeweski, MD, PhD[2]; Mehdi Moradi[3]; Parvin Mousavi, PhD[3]; Purang Abolmaesumi, PhD[3]; Alexander Boag, MD[4]; Chris Davidson, MD.[4] [1]Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey; [2]Department of Surgical Pathology, University of Pennsylvania, Philadelphia; [3]School of Computing and [4]Department of Pathology, Queens University, Kingston, Canada.

Context: We present an automated means for detecting prostate cancer from images of hematoxylin-eosin (H&E)–stained whole-mount histologic sections. The system is tailored to operate at low resolution (0.01 mm^2 per pixel). At low resolution glands are the primary visible structures. Accordingly, they are classified as either cancerous or benign using 2 discriminating features: (1) gland area and (2) the proximity of cancerous/benign glands to other cancerous/benign glands. The malignant glands are then aggregated into regions indicating the spatial extent of cancer.

Technology: A set of H&E-stained prostate whole-mount histologic sections are digitized using a whole-slide digital scanner. These images are analyzed using a set of routines implemented with the MATLAB software package.

Design: The cancer detection algorithm begins by segmenting the individual glands. Using these segmentations, the area of each gland is determined. A classifier then assigns a probability of cancer to each gland based on its area. Glands whose probabilities exceed an empirically chosen threshold are labeled as cancerous, while the remaining glands are labeled benign. A Markov random field iteration refines these labels by encouraging neighboring glands to be labeled similarly. Finally, the cancerous glands are aggregated into contiguous regions.

Results: Figure 4, a, shows a whole-mount histology section with the cancerous extent, as determined by a pathologist, roughly encircled in black. Figure 4, b, illustrates the result of gland segmentation (blue boundaries). The white box in Figure 4, b, is magnified in Figure 4, c. Figure 4, d, indicates the centroids (green dots) of those glands whose probabilities of cancer exceed the selected threshold. These labels are then refined by a Markov random field iteration, producing the labeling shown in Figure 4, e. Figure 4, f, illustrates the cancerous extent as determined by our automated system (green) and by an expert pathologist (yellow). Our algorithm detects cancer with a sensitivity of 0.8670 and a specificity of 0.9524.

Conclusions: We have presented an automated system for the detection of prostate cancer in whole-mount histology images. The method is specifically designed to operate at low resolutions where the primary visible structures are glands.

ELECTRONIC POSTER ABSTRACTS

A Web-Based Integration of Epidemiology and Human Papillomavirus Questionnaire Data to Head and Neck Neoplasm Database

Waqas Amin, MD[1] (aminw@upmc.edu); Harpreet Singh, MS[1]; Ann M. Egloff, PhD[2]; Althea M. Schneider, BS[1]; Jennifer L. Hetrick, RHIA, CTR[2]; Kerry R. Trent[2]; Anil V. Parwani, MD, PhD.[3] [1]Departments of Biomedical Informatics, [2]Otolaryngology, and [3]Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania.

Context: The University of Pittsburgh SPORE (Specialized Program of Research Excellence) Head and Neck Virtual Biorepository is a distinctive tissue bank informatics initiative that provides clinically annotated biospecimens through a Web-based interface that facilitates researchers' identification of available, clinically annotated biospecimens. Head and Neck Neoplasm Virtual Biorepository Epidemiology (Project 1 study) and human papillomavirus (HPV) study questionnaires capture patient epide-

miologic and sexual behavior information and allow integration with annotated clinicopathology data.

Design: The Project 1 interviewer-administered and HPV self-administered study questionnaires capture patient epidemiologic and sexual behavior information. The annotation tool provides for manual data entry and query on deidentified information obtained from these questionnaires within the data warehouse through a ''point and click'' interface that internally integrates with annotated clinicopathology data sources (cancer registry information system and coPath Plus).

Technology: The Web-based interface provides both data entry and electronic data load along with meta data–driven robust data query capabilities for study questionnaires, which is supported in a 3-tiered architecture and implemented on an Oracle application server on a Compaq DL360 server running Win2K with SP. The application utilizes the Oracle http server and mod_plsql extensions to generate dynamic pages from the Oracle database server to the users. The data warehouse is implemented using the Oracle 10g Enterprise Edition on a SunFire V880 server running Solaris 2.8.

Results: The integrated database provides subject epidemiologic information from the Project 1 questionnaire, subject sexual behavior from the HPV study questionnaire, clinical and pathologic data, and links to laboratory correlates including HPV status for authorized users.

Conclusions: The head and neck neoplasm virtual biorepository with robust translational biomedical informatics provides support to facilitate basic science, clinical, and translational research. The head and neck data query tool acts as a central source providing a mechanism for researchers to efficiently and effectively find clinically annotated datasets and biospecimens relevant to their research areas. The tool protects patient privacy by revealing only deidentified data according to the Health Insurance Portability and Accountability Act regulations. The data disclosure is strictly regulated by user's authorization.

System for Laboratory Image Management to Facilitate Digital Pathology for Multiple Cooperative Cancer Groups at the Biopathology Center

Thomas. J. Barr, BS[1] (Thomas.Barr@nationwidechildrens.org); Kathy Nicol, MD[2]; David Billiter, BS[1]; Mark Plaskow, BA.[1] [1]Research Informatics Core and [2]Department of Pathology, Nationwide Children's Hospital, Columbus, Ohio.

Context: Although the use of digital pathology continues to gain in popularity and acceptance by the research community, it remains a disruptive technology to integrate into existing laboratory workflows. The effective management and integration of digital pathology into the cooperative cancer group setting requires innovation and automation along the entire process including customer requests, priority of requests, order tracking, labeling, processing, quality assurance, and finally the safe return of the glass slides to the customer.

Technology: Utilizing ASP.NET, SQL server 2005, Aperio Spectrum Plus, and Web services, the Research Informatics Core is in the pilot phase of developing a system for laboratory image management (SLIM) that will fully automate the request, prioritization, processing, analysis, and return of glass slides and digital images to customers. SLIM also integrates multiple specimen banking applications and the virtual imaging for pathology, education and research, which captures pathology review data.

Design: SLIM is initially driven by the virtual microscopy request system, which captures necessary image order request information electronically from the customer. SLIM is also integrated with other operations and imaging systems developed by the Research Informatics Core including the Specimen Tracking and Receiving System and the Virtual Imaging for Pathology, Education and Research, which automates the pathology review process. This integration of systems allows for the capture of multiple data elements associated with both the specimen and the digital image.

Results: Although SLIM is not fully developed, early indications are positive as the system has significantly improved communication with customers as well as communication among members of the operations and imaging teams. Improved communications with customers has allowed the Biopathology Center teams to better plan and manage staff time and equipment resources.

Conclusions: We present a software application that integrates both specimen management and digital imaging applications utilized for multiple cooperative cancer groups to facilitate digital pathology into the existing laboratory workflow. The pilot phase of SLIM has already improved the effective use of digital pathology at the Biopathology Center. The complete system will likely accelerate further use of digital imaging for both existing cases and new cases received at our facility.

PathEd: An Online Digital Pathology Annotation and Teaching Tool to Further Pathology Education and Training

Melissa C. Castine[1] (castinem2@upmc.edu); Laura M. Drogowski, BS[2]; Goran Momiroski[4]; Shawn Moroney[4]; Ann R. Cecil[5]; Jonhan Ho, MD[2]; Anil V. Parwani, MD, PhD[3]; Drazen M. Jukic, MD, PhD.[2] [1]Departments of Biomedical Informatics, [2]Dermatology, and [3]Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania; [4]Innovative Medical and Information Technologies Center, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [5]Katz School of Business, University of Pittsburgh, Pittsburgh, Pennsylvania.

Context: PathEd is an online application developed by the Integrated Medical Information Technologies System Telepathology Project in order to augment pathology education in the Air Force and University of Pittsburgh Medical Center health systems. Typically, pathology education consists of slide sets in a given domain that are accessible to students that require them to review the slides and the given diagnosis. Many times there is little to no additional case information given for a slide.

Technology: PathEd provides complete case information and allows pathologists to annotate the digital images with any pertinent diagnostic and/or prognostic findings that may aid the pathology student in understanding why a case was given a particular diagnosis.

Design: Since PathEd is easily accessible via the Internet, care was taken during the application's development to ensure an extremely user-friendly design. The application uses simple drop-down menus that allow for quick and simple navigation through the system. Case input requires the user to fill out simple forms with any available case information. PathEd organizes cases according to the user-defined organ, category, and subcategory classifications and presents the cases in a tree structure through which the user can navigate to find relevant cases. Additionally, PathEd allows for several search options, including a keyword and classification searches.

Results: Currently, PathEd has more than 100 deidentified cases that have been entered into the general repository of cases, which can be utilized for educational purposes. These cases are richly annotated to further pathology education and training for residents, fellows, and medical students. Potentially, similarly annotated digital slides can be created for training other allied health professionals such as cytotechnologists and pathology assistants.

Conclusions: PathEd is an easily Web-accessible pathology education software application that allows users to input complete case information, append digital slides to the cases, and provide necessary annotations to those cases. This accessibility and added case information gives PathEd an advantage over traditional slide set study; thus, PathEd has great potential to further pathology education because of its accessibility through the Internet, ease of use, and detailed case information.

The Word Processor as a Data Collector for Quality Assurance in Anatomic Pathology Reports: Tracking the Use of a Text-Insertion Feature

Martin C. Chang, MD, PhD (mcchang@partners.org); Frank C. Kuo, MD, PhD. Departments of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Context: Anatomic pathology reports are documents generated with word processor software and are integrated into the laboratory information system. Commonly used words or phrases may be inserted into the word processor document using abbreviations, with the full text being subsequently substituted automatically. The purpose of text insertion is to reduce typing and errors in transcription.

Technology: At Brigham and Women's Hospital, the anatomic pathology information system (PowerPath 8.3, IMPAC, Sunnyvale, California) includes a built-in software interface to a word processor (Word 2003, Microsoft, Redmond, Washington) in which all anatomic pathology reports are generated. The text insertion feature and the recording of its use are built into this software interface.

Design: Each instance of text insertion using an abbreviation is recorded by the word processor software interface. We analyzed 1000 consecutive text insertion events and compared them with the final reports in which they were used, with particular attention to erroneous uses.

Results: The text insertion events varied from the replacement of a word to the construction of an entire synaptic report. Both simple text and other retrievable information were inserted using the function. In most cases, the function was used appropriately. Three main types of erroneous use were apparent: (1) instances of text insertion for which the inserted text was subsequently deleted from the final report; (2) incomplete use, in which retrievable diagnostic information was omitted; and (3) failure to use the text inser-

tion function in a field required for subsequent retrieval. All types of errors were more prevalent in longer synoptic reports.

Conclusions: The use of automatic text insertion in creating anatomic pathology reports saves time and allows important diagnostic data/statistics to be retrieved more readily. In longer reports, errors may result from unclear text or illegibility in the draft diagnosis and/or errors in transcription from this draft. Efforts to reduce errors should therefore aim for improvements in the transcription process.

HistoStitcher: An Interactive Software Package for Reconstructing Digitized Whole Histologic Sections From Fragmented Slices

Jonathan Chappelow, BSE[1] (chappjc@eden.rutgers.edu); Anant Madabhushi, PhD[1]; Michael D. Feldman, MD[2], PhD; John Tomaszewski, MD.[2] [1]Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey; [2]Department of Surgical Pathology, University of Pennsylvania, Philadelphia.

Context: We present a software program to accurately reconstruct whole histologic sections from digital slides of smaller adjacent tissue slices, such as quadrants of a whole prostate histologic section. The program allows interactive identification of anatomical landmarks on adjacent sections, automatic generation of the optimal transform with optional constraints, followed by automatic spatial transformation and combination of the individual fragments. We demonstrate the program in combining 2 adjacent histologic sections of a prostate gland, which was required to be sectioned into quadrants because of limitations in slide imaging hardware.

Technology: Digital images of quadrant histology sections obtained by digitally scanning tissue slides are first aligned and then reconstructed using a set of image processing routines implemented in the MATLAB software package.

Design: The developed graphical user interface (GUI) utilizes an interactive anatomic landmark-based approach to facilitate simple and accurate reconstruction of histology sections from smaller fragments. The user first identifies pairs of corresponding anatomic landmarks that exist along the incision separating 2 adjacent sections. The program then automatically solves for the optimal linear transform required to bring the landmarks into closest possible alignment. Finally, the images are conjoined via the determined spatial transformation and the resulting larger image is returned. The user may again execute the program with the previous result and another adjacent fragment to obtain a new larger semisection. The program also offers options to constrain the transformation to use isotropic image scaling, no scaling, or no reflection.

Results: A screen shot of the program being used to combine 2 histologic quadrant sections of the prostate is shown in Figure 5, A. Control points are first interactively selected by clicking directly on the image (red arrows). The reconstructed image is then generated on demand when the user is satisfied with the identified landmarks. An example of a reconstructed image is also shown in Figure 5, B.

Conclusions: We have presented an image reconstruction program that offers a powerful image alignment engine with flexible spatial transformation options. The program was used for reconstruction of fragments of prostate histologic sections into corresponding whole sections and required only the specification of visible landmarks for alignment via mouse clicks from the program operator. The GUI is friendly and intuitive enough to be used by anyone.

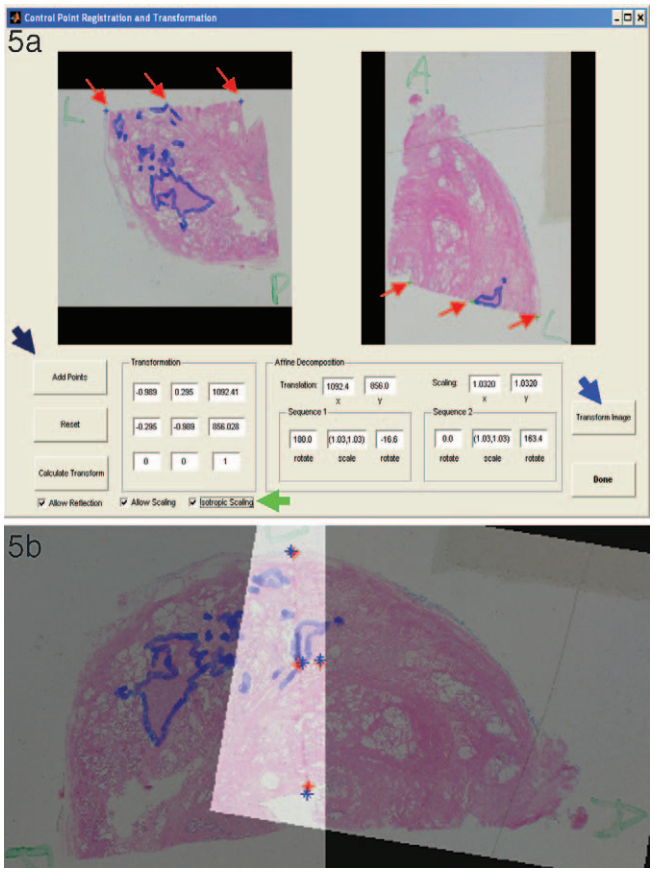
Data Capture and Analysis for Research Protocols

Michael Davis, MS (davismk@upmc.edu); Brenda Crocker; Bob Rubin; Kelli Richter; Julia Michel. Information Services Division, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

Context: Throughout the country, a number of research-based clinical trials are being conducted within various universities, medical centers, and pharmaceutical companies. With these trials, extensive amounts of data must be captured and managed for analysis. These data consist of a vast range of information, from clinical to regulatory to financial. Researchers are often in need of software to capture these data in a centralized and structured place, create a registry of patients involved in their studies, and capture data forms customized to their research.

Technology: The system was developed in Java using a 3-tiered, object-oriented approach. It uses a standard Web browser on a Microsoft Windows (Redmond, Washington) platform for the front-end interactions. The middle-tier is a remote method invocation server running on an Intranet machine using Windows servers connecting to back-end Oracle (Redwood Shores, California) databases residing on Sun Unix hardware.

Design: We have developed the Clinical Trials Management–Enterprise to fit the needs of this workplace. The application provides the clinical researchers, nurse coordinators, and supporting offices with an integrated set of tools for managing the administrative and clinical functions for both trial- and patient-based activities. Two options for the user interface are available: a ''lite'' version with a more tailored access to the studies and registration, and the more robust interface for handling other aspects of the study. One area of focus that we have for the software is to provide a tool that can easily create custom forms to capture questionnaires and data specific to the researchers' study and allow data entry in a user-friendly screen.



Results: The introduction of the lite Web-based interface offers a straightforward way to enroll subjects onto a registry and track patients involved in the project. The National Mesothelioma Virtual Bank utilizes Clinical Trials Management–Enterprise to keep a registry of the subjects involved. By capturing the Tissue Bank ID for the subject they are also able to link up with collected tissue and biopsy samples. Furthermore, the use of custom case report forms ensures that the necessary data can easily be captured and thus be available for analysis. Electronic versions of existing questionnaires and data collection forms can be created to store the data. In the case of the National Mesothelioma Virtual Bank, the prospective biospecimen collection information is captured to provide electronic access to the appropriate users and is also then available for analysis.

Conclusions: Utilizing software such as Clinical Trials Management–Enterprise enables researchers to collect structured data in a central environment so that ultimately the goal of analysis can be accomplished, new ideas can be generated, and the research process can continue successfully.

ConsulTHiS—Digitizing Dermatopathology Consultation to Improve Patient Care

Laura M. Drogowski, BS[1] (drogowskilm@upmc.edu); Russell A. Silowash, BS[2]; Drazen M. Jukic, MD, PhD.[1] [1]Departments of Dermatology and [2]Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania.

Context: The existing dermatopathology intradepartmental consultation process relies on pen and paper–based records and physical transport of materials. This inefficient, cumbersome process delays patient care, impedes pathologists' workflow, and is not easily incorporated into patient records. By creating a user-friendly database named ConsulTHiS we at-

tempt to make forms and digital images available for viewing and review by dermatopathologists.

Technology: ConsulTHiS is a Microsoft Access database that allows for creation and entry of image-enriched case information. Up to 5 images (jpeg, tif, bmp) can be added to a case using DBPix. Alert e-mails are sent to each consulting pathologist when a consultation is requested and to the requesting pathologist when consultations are completed.

Design: The referring pathologist logs into the database, generates a case, has the option to add images, and selects pathologists to consult on the case. Alert e-mails are sent to each pathologist who can then access the case from his or her queue of pending consultations. Once completed, the resulting form is locked and a digital signature is applied. A pdf file of the report can be generated by any pathologist involved in the consultation. Records are stored according to case accession number and requesting pathologist. Though names and patient data are not stored in the database, passwords and permissions are in place to preserve the exact diagnoses.

Results: The resulting database allows for rapid access to consultation information and immediate notification of changes in consultation status. As an unforeseen benefit, the opportunity for consulting pathologists to form an initial impression of the case allows them to request additional materials early in the process; therefore, these materials can be shipped with the case, rather than arriving after the case is evaluated once.

Conclusions: By implementing a user-friendly, fast system, we anticipate the number of pathologists seeking second opinions to increase. Pathologists are currently deterred by the inefficiency and disorganization of the process, and the delay in the diagnosis resulting from a need to ship materials and paperwork around prevents them from seeking expert help.

Utility of Telepathology in Diagnosis of Pigmented Lesions by Conventional and Immunohistochemical Techniques

Laura M. Drogowski, BS[1] (drogowskilm@upmc.edu); Steven Catinchi, MD[2]; Jill Buckthal-McCuin, MD[2]; Jonhan Ho, MD[1]; Arash Radfar, MD, PhD[1]; Drazen M. Jukic, MD, PhD.[1] [1]Department of Dermatology, University of Pittsburgh, Pittsburgh, Pennsylvania; [2]Educational Program (EP), University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

Context: Telepathology has not been explored in the evaluation of immunohistochemistry or specific diagnostic features relative to magnification. This study documents the feasibility of immunohistochemistry evaluation via telepathology, assessing its potential in remote diagnosis and determining its utility in observing diagnostic features.

Technology: The telepathology system consists of a high-resolution DP70 Olympus digital camera attached to an Olympus BX41 microscope. The camera is connected to a Pentium 4 (2.8 GHz processor, 2 GB RAM, 124 MB video card, Windows XP SP2) computer with Olympus Microsuite Basic with Netcam. The system enables real-time broadcasting of images to a static IP address through transmission control protocol/internet protocol (TCP/IP).

Design: Four pathologists evaluated 80 cases of dysplastic nevi and early melanomas on hematoxylin-eosin and immunohistochemical stained slides. A digital and glass set of cases were evaluated, each in 2 phases. The digital slides were hosted at a remote site and evaluated via a Java-enabled Web browser by consulting pathologists. Time to transmit, open, and review each case was documented in a survey completed by all reviewing pathologists. Magnification requirements for the evaluation of pigmented lesion features were documented.

Results: Phase 1 is complete and phase 2 is in progress. Preliminary results reveal a high concordance rate between telepathology and light microscopy diagnosis for dysplastic nevi with varying degrees of atypia. In the case of early malignant melanoma, diagnostic discrepancy occurred in few cases when evaluated through telepathology. Time spent on cases evaluated through telepathology is still significant as it takes more than twice the time as conventional light microscopy. Evaluators did not feel totally comfortable when the slide was moved at the remote site, because an area of interest was often omitted or out of focus.

Conclusions: A slow time to focus in telepathology evaluation elicited the most complaints. Regardless of level of comfort, preliminary data suggest that pathologists' evaluations were consistent between digital and glass slide evaluation. The second phase of the study will evaluate the use of telepathology in evaluating immunohistochemistry. Evaluators agreed that the robotic microscope provided excellent detail and would be useful in nonprimary diagnosis.

Workflow Tracking in Anatomic Pathology Consultation Cases

Laura M. Drogowski, BS[1] (drogowskilm@upmc.edu); Jamie Martina[2]; Laureen Conrad[1]; Drazen M. Jukic, MD, PhD.[1] [1]Department of Dermatology, University of Pittsburgh, Pittsburgh, Pennsylvania; [2]College of Wooster, Wooster, Ohio.

Context: The existing anatomic pathology second opinion consultation process relies on paper-based record keeping and physical transport of paperwork and glass slides. To date, no study documents the time required for each step of the consultation process (from initial accessioning to final sign-out). We hypothesized that most of the delay in case sign-out occurs during the diagnostic workup, and this delay is exacerbated by inefficiencies in the consultation process.

Technology: To evaluate case consultation efficiency, we tracked the time required for completion of each step of the consultation process. Information was recorded in an Excel spreadsheet and obtained from shipping documents, CoPath, and measurements by study participants.

Design: Consultation requests were tracked for 45 days, and during that period 59 consultations were requested. The receiving administrative professional recorded consultation-specific data, including (but not limited to) the referring site, date the consultation was shipped, method of transport, date and time received, number of parts, blocks and slides, diagnosis under which it was submitted, time to accession, time at which the specimen was signed-out in CoPath, final diagnosis, and final date and time it was faxed to the requesting office.

Results: The average time between referring office accession and consulting office accession was 13.5 days. The average time to transport a case was 1.2 days. Time to accession the consultation averaged 10.6 minutes and did not significantly contribute to the overall time to sign-out. The time from consultation receipt to sign-out by a consulting pathologist, pending further data, averaged 6 days.

Conclusions: The delay in case sign-out, when a consultation was requested, was related to the initial diagnostic workup and the delay in the consulting pathologist receiving additional requested material. By implementing telepathology to give the consulting pathologist an initial impression of a challenging case, additional materials could be requested early in the process. Furthermore, the organization of paperwork into a database could preserve the case history and impressions for easy reference. Both strategies will be implemented in a future study.

Literature Mining in Cytology Discovery Virtual Images

O. Ferrer-Roca, PhD (catai@teide.net); Francisco Marcano; Maria Esther Vidal, PhD; Edna Ruckhaus, MSc; Xiomara Santos; Enrique Iglesias, PhD. Department of Pathology, University of La Laguna, LaLaguna, Spain.

Context: In the present paper we introduce the innovative technique of literature mining to discover relevant cytologic findings to automatically detect regions of interest (ROI) in cytologic virtual slides that are being assessed remotely.

Technology: Four types of technologies are included: (1) literature mining techniques; (2) optimization algorithms to select relevant literature; (3) small size virtual slide digitization process; (4) selection of ROIs for remote diagnosis; (5) JPIP server availability of virtual slide for distant diagnosis; and (6) Internet services.

Design: The project was designed to ensure long distance diagnosis of cytologic specimens taken by a laboratory technician. The small size virtual slide is digitized and the ROI selected using literature mining.

Results: The results had been tested in a rural area in Venezuela. Cytology specimens were examined by the expert via the Internet and the ROI is selected with 90% accurate diagnosis.

Conclusions: Novel techniques in pathology such as literature mining and small size virtual slide or Web services have to be put in place in order to facilitate remote diagnosis.

Grid-Enabled Scoring and Assessment of Tissue Microarrays for High-Throughput Comparative Analysis

David J. Foran, PhD[1] (djf@pleiad.umdnj.edu); Lin Yang[1]; Jun Hu[1]; Vicky Chu[1]; Ryan Golhar, PhD[1]; Lauri A. Goodell[1]; Michael Reiss[1]; Wenjin Chen, PhD[1]; Tahsin Kurc, PhD[2]; Renato Ferreira, PhD[2]; Metin Gurcan, PhD[2]; Tony Pan, MS[2]; Steve Langella, MS[2]; Scott Oster, MS[2]; Shannon Hastings, MS[2]; Joel H. Saltz, MD, PhD.[2] [1]Center for Biomedical Imaging & Informatics, The Cancer Institute of New Jersey, UMDNJ-Robert Wood Johnson Medical School, New Brunswick; [2]Department of Biomedical Informatics, The Ohio State University, Columbus.

Context: Tissue microarray (TMA) technology provides insight regarding the underlying disease mechanisms and holds promise for advancing

cancer biology and drug discovery. Future progress in several key areas of research relies upon the capacity of assessing expression patterns in TMAs. The central objective of this research is to develop and evaluate TMA-Miner, a grid-enabled content-based image retrieval system for performing quick, reliable characterization and comparative analysis of TMAs. To test these technologies, a consortium of strategic sites is being established using the caBIG caGrid infrastructure.

Technology: The project is an extension of the team's prior work and leverages a Web-based, image-guided decision support system, a distributed telemicroscopy/image analysis system, an intelligent archival system, a framework for distributed execution of algorithms, and a model-driven, service-oriented architecture for secure federation of data and analytical resources.

Design: The team has developed software to automatically delineate the tissue discs comprising the arrays and decompose the discs into their constituent staining maps. Moreover, analysis methods are exposed to the environment as caGrid services. The back end of an analytical service can be a parallel machine capable of invoking multiple independent instances of the method simultaneously. Workflows can be composed by linking multiple analytical services using the grid-level workflow execution environment.

Results: A prototype system has been developed to automatically delineate tumor regions within a given imaged disc and formulate a query into a gold standard database of previously scored arrays to identify those tissue discs that contain lesions exhibiting similar expression signatures. A 4-stage analysis workflow has also been implemented with the following steps: (1) segmentation and color decomposition on TMA slides; (2) determination of region of interest; (3) texture filtering; and (4) texton computation. The services can be hosted on Grid machines for distributed execution. Each host may be the front end of a compute cluster, allowing multiple images to be processed simultaneously. Comparing the resulting segmentation with those hand-drawn by a board-certified pathologist, the false-positive rate averaged 6.62% and false-negative rate averaged 3.15%. Feasibility experiments conducted using 3744 imaged breast tissue discs demonstrated 89% accuracy in discriminating between benign and cancerous tissue discs and 80% when discriminating among 2 types of breast cancer and normal tissue. Only 30% of the images were used for training the system.

Conclusions: We have established a reference library of expression signatures for a mixed set of more than 130 000 tissue disc images. The feature extraction and indexing modules have been tested for distributed execution on a computer cluster at Ohio State University (OSU). The first-generation image-based query module has been deployed to strategic sites at the Cancer Institute of New Jersey and OSU and performance studies are underway. In the next phase of experiments we will continue to expand the database and image archives and investigate the use of the system in predicting clinical outcomes. We are also evaluating our overall framework and associated technologies for use in a range of other histopathology and hematopathology imaging applications.

This work was supported in part by the NCI caGrid Developer grant 79077CBS10, the State of Ohio Board of Regents BRTT Program grants ODOD AGMT TECH 04-049 and BRTT02-0003, the NHLBI R24 HL085343 grant, the NIH U54 CA113001 grant, the NIH through contract 1R01LM009239-01A1 from the National Library of Medicine, and the NSF grants CNS-0403342 and CNS-0615155.

Two-Year Evaluation of Barcode Identification System in Mortality and Autopsy Service With Integration of Demographic Database of Hospital Records and Microsoft Access Interface

Avneesh Gupta, MD (avneeshgupta@hotmail.com); Dan X. Cai, MD, PhD; Sandra Cota; Joseph F. Tomashefski, Jr, MD. Department of Pathology, MetroHealth Medical Center and Case Western Reserve University, Cleveland, Ohio.

Context: Linear barcodes are currently used in numerous hospital settings to improve efficiency and reduce error. We previously introduced an electronic mortality service database including a barcode system for body identification and Microsoft Access Interface. In this study we present a 2-year follow-up of this program, emphasizing the results of 2 separate user-satisfaction surveys of mortality and autopsy service personnel.

Technology: A barcode scanning process using a wireless barcode scanner (Intermec Technologies Corporation, Model No. ScanPlus 1802 Vista) and a Web site for body identification linked to demographic data from hospital electronic medical records were implemented in 2005 on our mortality service. Linked demographic information serves as a check on potential errors during manual data entry. The new database was transferred to a structured query language secure server.

Design: Using conditional formatting, an algorithm was developed for scanning barcodes on the decedent's body, the body bag, and mortality work sheet. For quality assurance, the barcode-scanned medical record number is used to avoid duplication of a death entry into the database. An employee identification number and a time-stamp are also entered into the database for accountability purposes.

Results: Two brief user-satisfaction surveys were conducted 9 and 22 months, respectively, after implementation of the system. The survey results were favorable, indicating that users had accepted the new barcode system. In the most recent survey, all users emphasized the system's ability to alert them of potential errors or identification mismatches and reduce the possibility of data entry errors or misidentification. Users graded barcode scanning as superior for increasing personnel accountability and efficiency. No respondent experienced shut-down, scanning failure, or data entry errors. The system has had an impact on the overall reduction of errors on the mortality service.

Conclusions: We have shown during a 2-year period, using a simple and inexpensive electronic monitoring system, that barcode technology is feasible and efficacious on the mortality service as in other areas of the laboratory. The application of similar systems in tracking mortality of mass casualties following terrorist attacks, in military operations, and for general forensic practice is feasible and currently under development.

Informatics Tools Useful in Solving the Credentialing and Competency Assessment Standards of the Joint Commission on Accreditation of Healthcare Organizations: A Comparison of Manual Slide Versus Digital Slide Methods

Lewis A. Hassell, MD (lewis-hassell@ouhsc.edu); Kenneth E. Blick, MD. Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City.

Context: The Joint Commission in 2008 proposed new standards for the credentialing of physicians, which pose challenges in pathology. We have elsewhere proposed a 3-tiered process for the assessment of competency of newly credentialed pathologists in anatomic pathology involving a test composed of locally representative materials, followed by proctored review of skills with ''live'' cases, and then, finally, selective review of specifically identified areas of weakness or unknown ability. We explore in this project the role of digitized slide images as an alternative to glass slides in the initial testing phase of this assessment.

Technology: We used whole-slide digitized images of complete cases captured using the Aperio slide scanner and viewed using the Imagescope software (Aperio Technologies, Vista, California).

Design: Overall diagnostic error standards were determined for existing staff using a retrospective review of cases from each sign-out service. Mean and standard deviation from a perfect score was established using a weighted scoring system. Representative cases from each service were selected as a test group for newly hired members of the department and for potential use as an ongoing competency test for existing staff at the time of recredentialing. Slides from these cases were scanned. A test scenario was composed of a representative volume of cases from a single day's workload. New staff members were given the test, using validated but differing slide sets in both glass and digital formats. Diagnostic error and overall defect rates were collected and scored on the weighted scale and compared.

Results: Three newly hired and 3 current staff pathologists with anatomic pathology privileges were assessed during the study period. The results from digital slide review were comparable in diagnostic assessment overall, rare event detection, and subtle feature detection. Performance with glass slides was comparable to established standards by existing staff.

Conclusions: These results show the utility of using digitized slides for competency assessment measures in anatomic pathology. This facilitates the process measures needed to address the concerns driving expanded credentialing stringency. We continue to emphasize the need for these assessments to be based on local materials and standards, rather than from centralized digital slide repositories.

Ubiquitous Virtual Slide System of Pathology

Woo Young Jang, MD, MS (pathwyj@empal.com). Hangang Sacred Heart Hospital, Hallym University, Seoul, Republic of Korea.

Context: It is very hard to create a ubiquitous computing environment without Internet Protocol Version 6 (IPV6). The Korean Society for Cytopathology decided to build a ubiquitous virtual slide system of pathology. But the high price of hardware, software, and maintenance fees and some technical problems with the Internet make it difficult to create such a system. After searching the Internet to find helpful software and sign a

contract with a friendly company, a performance and money saving system was finally created.

Technology: Virtual slide machine, HP xeon server, Pentium (Quad Core) desktop computer, Pentium (Dual Core) laptop computer, combined cable/wireless Internet (speed: 100/54 Mbps), wireless Internet access point repeater type (speed: 54 Mbps), wireless Internet USB type (speed: 2 Mbps), fixed IP/DNS emulation software, free FTP software, free voice chatting software, and free antivirus/antispyware software were used for construction of this system.

Design: This system was designed for performance and money saving, along with clear vision diagnosis, consultation, conference, and teaching. We set this system on a HP Xeon server, then gave a public notice to all members. The free Virtual slide machine was installed on the HP Xeon server and connected with combined cable/wireless Internet. The retail or free control software was installed on the HP Xeon server. Every pathologist who installed the free viewer software to their Pentium computer system downloaded the virtual slide via FTP channel or Web browsing and exchanged their opinion about the case, using monitor and voice recognition system. When they reached agreement, they could take high-quality pictures.

Results: Prototype ubiquitous virtual slide system for pathology made it possible to get excellent diagnosis, consultation, conference, and teaching capabilities and high-quality pictures anytime and anywhere with ease.

Conclusions: The prototype ubiquitous virtual slide system had the following advantages: (1) easy handling of viewer; (2) high-resolution image quality; (3) easy to get snapshot; (4) discussion using voice recognition software; (5) access anytime and anywhere with ease; and 6) performance and money saving. More technical refinements would be desirable for the more advanced expanded tasks.

Structured Data Reporting: Implementation of a Next Generation Tool

Joy John Mammen, MD (jmammen1@hfhs.org); J Mark Tuthill, MD. Department of Pathology & Laboratory Medicine, Henry Ford Hospital, Detroit, Michigan.

Context: Evolution from traditional textual reporting to structured pathology reports has been slow to develop but is gradually gaining momentum as next generation tools emerge. We report progress on an ongoing phased deployment of a laboratory information system (LIS) integrated solution.

Technology: We used the following: LIS: Sunquest Copath Plus (2.5) (Sunquest Information Systems, Tucson, Arizona); structured data capture design tool: mTuitive Agile Author; and enduser application: mTuitive xPert Client (mTuitive, Centerville, Massachusetts).

Design: We have previously reported on the design of checklists using the mTuitive xPert Pathology application and benefits of the LIS integration. Prior to the implementation, we composed our customized checklists in the mTuitive Agile Author. These were extensively tested paying attention to technical and content details. All lists were approved by the appropriate expert pathologist prior to implementation. The staff and residents were given demonstrations followed by education and training, and were provided with procedures. Phase 1: The preexisting checklist content was available to the users in the new format without any logical control or calculations. Relevant help files were included. Phase 2: Four tumor checklists were enhanced utilizing logical control and tumor grade calculations besides standardized format for the lists and data collection, data integrity verification, and enhanced help files. Extension of the same features for other lists is in progress. Phase 3: This phase envisages automated calculation of stage and 2 such lists are undergoing testing. Different challenges have been identified and are being addressed in the test environment.

Results: All tumor cases are now signed out using this tool. The most important benefits were related to efficiency, namely functionality and workflow. Additional key benefits relate to reduction in errors of omission and data integrity. Recently, 5 new checklists have been introduced and are undergoing validation by experts. Four phase 2 lists are undergoing preimplementation testing. Six technical incidents were reported after going live that needed resolution by the vendor. Important disadvantages perceived by the users are the time required to open a case in the xPert Client and the inability to attach more than one checklist to a specimen.

Conclusions: Deployment of a new LIS integrated data structuring tool requires close collaboration between experts and users, along with consistent effort and support from leadership and vendors.

Cell Surface Tessellations in Malignant Growth: Morley's Theorem

http://www.netautopsy.org/cellmorl.htm

G. William Moore, MD, PhD[1,2,3] (George.Moore4@va.gov); Raimond A. Struble, PhD[4]; Lawrence A. Brown, MD[1,2]; Grace F. Kao, MD[1,5,6]; Grover M. Hutchins, MD.[3] [1]Pathology and Laboratory Medicine Service, Veterans Affairs Maryland Health Care System, Baltimore; [2]Department of Pathology, University of Maryland Medical System, Baltimore; [3]Department of Pathology, The Johns Hopkins Medical Institutions, Baltimore, Maryland; [4]Department of Mathematics, North Carolina State University, Raleigh; [5]Department of Dermatology, University of Maryland Medical System, Baltimore; [6]Department of Dermatology, George Washington University School of Medicine, Washington, DC.

Context: Tumors of surface epithelium are among the most common human malignancies. In benign surface epithelium, the cell surface exhibits a regular, repeated packing of cells, or tessellation, resembling a collection of equal cylinders resting side by side. Malignant transformation involves variably sized cells, a disorganized surface, and the tendency to invade surrounding tissues.

Technology: Ordinary and synthetic geometry were used.

Design: Mathematically, a tessellation is a periodic tiling of the plane by polygons or pace by polyhedra. An unbroken sheet of mucosal or epidermal cells viewed en face can be approximated as a collection of tangent circles. Any triple of tangent circles on a tessellation forms a triangle with vertices at the circle centers. Morley's theorem states that every tangent cell-triple has a unique internal Morley triangle formed by trisecting the angles of primary triangle vertices. This Morley triangle may serve as a communication hub for cell-to-cell interaction.

Results: It is demonstrated that the Morley triangle is maximal in an equilateral primary triangle.

Conclusions: Malignant surface cells are characterized by more size variation and less balanced packing. In this model, unequal cell size and decreased Morley triangle ratio are geometric features of the same underlying process. Therapy for smaller Morley triangle ratio might possibly control the malignancy process. Mathematical models can be used to explore alternatives to classical hypotheses in pathology and explore general paradigms.

Digital Staining Instrumentation

Jared M. Orrock, MD (orrock.jared@mayo.edu); Keith J. Kaplan, MD. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota.

Context: With light sources of increasingly broader ranges, spectral analysis of tissue sections has evolved from 2 wavelength image subtraction techniques to Raman near infrared micro-spectroscopic mapping, permitting discrimination of cell types and tissue patterns.

Technology: We are developing next-generation hyperspectral imaging systems for use in the physical and life sciences, specifically histopathology using unstained sections of routinely formalin-fixed, paraffin-embedded tissue. Previous investigators have shown previous success with high diagnostic accuracy for algorithms derived from hyperspectral imaging in hematoxylin-eosin (H&E)–stained sections. The use of unstained sections eliminates the cost and effort associated with staining techniques as well as artifacts that may be introduced by the staining process.

Design: A dedicated light source is integrated with a standard commercial research grade microscope and cooled charge-coupled device camera for obtaining a hyperspectral cube with each stack on the order of 512 $\times$ 512 pixels. Wavelength range is on the order of 5 nm covering the full optical range of 400 to 700 nm.

Results: Preliminary experience has shown that current available technologies can readily digitally reproduce an H&E-stained section while creating a digital dataset that can be subjected to algorithmic tools from these large datasets; this process may enhance the practice of pathology. We will show that ''gray'' areas of surgical pathology, subject to staining and intraobserver/interobserver variation, such as accurate grading of dysplasia in Barrett esophagus, determination of in situ versus invasive in a wide variety of organ systems, and rare event detection in lymph nodes or exfoliated specimens can be diagnosed with greater specificity beyond light microscope analysis alone.

Conclusions: Hyperspectral image analysis may enable pathologists to view an unstained sample and obtain the same information they would obtain if they looked at a stained sample under a traditional light microscope. A digital staining system would enable a pathologist to review a stained sample more effectively than one could visualize it using traditional light microscopy techniques. Image identification and classification could be used to prescreen samples or augment the screening capabilities of a pathologist. Digitally stained samples could yield information for pathologists to distinguish more details and enhance diagnostic accuracy.

The Effect of Antivirus Software on the Laboratory Information System

Liron Pantanowitz, MD (Liron.Pantanowitz@bhs.org); Andrew Ellithorpe, MHS; Christopher N. Otis, MD. Department of Pathology, Bay-

state Medical Center, Tufts University School of Medicine, Springfield, Massachusetts.

Context: Security is a key feature of any laboratory information system (LIS). The Clinical and Laboratory Standards Institute recommend that laboratory computer systems be monitored and protected against malicious intrusions (eg, spyware, Trojan horses, cookies, worms, viruses, and root kits) and attacks. With computers connected to the Internet forming an integral component of the modern LIS, the use of virus protection software is recommended. Antivirus programs perform audits (scans) of applications and data files on hard disks for viruses (eg, looking for viral signatures) and remove any that are found. It is also suggested that antivirus software be installed and run daily on servers and all end-user devices that access laboratory data. The impact of antivirus software on the LIS has, to the best of our knowledge, not been investigated. Therefore, the aim of this study was to determine the effect antivirus software has on LIS performance.

Technology: Anatomical pathology LIS (CoPath Plus version 2.5, Cerner), servers (database server Compaq ProLiant ML570; interface server Compaq ProLiant DL360), enterprise antivirus and antispyware software (McAffee version 8.5.0i), and application performance monitor (Citrix EdgeSight) were used.

Design: Our anatomical pathology LIS was installed and operating on an array of networked database and interface servers. Each server had antivirus and antispyware software installed that was continuously running scans involving all folders. Edgesight software was installed to collect data (errors, central processing unit usage, network activity) for trend analysis in order to monitor performance as experienced by the end-user.

Results: Increased network system activity spikes on our servers were detected that corresponded to antivirus and antispyware scanning activity. These activity spikes resulted in decreased LIS bidirectional data transmission over our network that accounted for LIS end-user feedback of delays (so-called episodes of hanging). Excluding LIS application folders (ie, .exe folders) in the directory from antivirus scans alleviated this problem.

Conclusions: Antivirus and antispyware software programs may affect LIS activity resulting in diminished performance. In order to avoid this problem we recommend excluding application folders during antivirus scans. Utilizing performance monitoring software to view LIS utilization and performance will allow such issues to be proactively addressed.

Decentralized Computer-Assisted Image Analysis for Estrogen Receptor (ER), Progesterone Receptor (PR), and Human Epidermal Growth Factor Receptor (HER2/*neu*) Immunohistochemical Analysis

Liron Pantanowitz, MD (Liron.Pantanowitz@bhs.org); Andrew Ellithorpe, MHS; Christopher N. Otis, MD; Giovanna M. Crisi, MD, PhD; Richard C. Friedberg, MD, PhD; Peter Marquis, BA. Department of Pathology, Baystate Medical Center, Tufts University School of Medicine, Springfield, Massachusetts.

Context: Accurate determination of estrogen recepter (ER), progesterone receptor (PR), and human epidermal growth factor receptor (HER2/*neu*) status in breast cancer is required to guide appropriate therapy. Computer-assisted image analysis (CAIA) has been shown to provide an accurate and reproducible method to score immunohistochemical staining. Several guidelines support the use of CAIA for breast marker quantification, with appropriate pathologist supervision. The aim of this study was to implement CAIA for this purpose in a surgical pathology department spanning distant medical centers to mimic daily practice.

Technology: Networked multimedia workstations (Dell Optiplex 745 personal computers), Spot Insight digital microscope cameras (Diagnostic Instruments, Sterling Heights, Michigan), Pathiam Web-based application (BioImagene, Cupertino, California), and server (Oracle application and image database) were used.

Design: Digital camera settings were standardized (calibrated) and 3 to 5 microscopic fields of view (29 cases), stored in jpeg format using ×20 eyepiece magnification, were acquired for ER, PR, and HER2/*neu* immunohistochemical stained breast (invasive and in situ) carcinomas. Nuclear immunostaining for ER and PR was analyzed using the Allred scoring system (proportion + intensity = total score) and membranous HER2/*neu* staining was analyzed per American Society of Clinical Oncology/College of American Pathologists recommendations (scored 0, 1+, 2+, 3+). CAIA was performed with defined control parameter sets established for each run. Fluorescence in situ hybridization (FISH) for HER2/*neu* was obtained in a subset of cases. Score result and analysis time (minutes) between manual (pathologist) and CAIA were compared.

Results: There was excellent concordance for ER, PR, and HER2/*neu* scores between pathologists and CAIA. For HER2/*neu* results there was good correlation among manual scoring, CAIA, and FISH. Rare discordant results (3 cases) for ER and PR were attributed to nonspecific cytoplasmic staining of tumor cells. CAIA, including image acquisition time, was considerably longer to perform than manual scoring.

Conclusions: Decentralized CAIA for immunohistochemistry designed to mimic daily surgical pathology workflow in practice is feasible. However, decentralized image acquisition from individual workstations requires vigilant standardization, entails a cumbersome workflow, and is time consuming. Many of these issues could be overcome by using a central system with whole-slide imaging.

Application of AutoHotkey to Enhance Data Entry Into the Laboratory Information System

Liron Pantanowitz, MD (Liron.Pantanowitz@bhs.org); Andrew Ellithorpe, MHS; Karen Murley-Kells; Carol A. Rauch, MD, PhD; Gary Poirier; William Lareau. Department of Pathology, Baystate Medical Center, Tufts University School of Medicine, Springfield, Massachusetts.

Context: Manual data entry into a laboratory information system (LIS) can be tedious, time-consuming, and prone to errors. This is especially true for systems requiring long sequences of keystrokes to accomplish specific tasks. Prior techniques to improve efficiency and accuracy of data entry into the LIS included barcoded phrase sheets and/or keyboard macros (fastkeys) to automate the entry of commonly used keystroke sequences. Unfortunately this technique was hindered by the limited number of fastkeys available. AutoHotkey is an open source macro-making utility for Windows used to create hotkeys (shortcut keystrokes) that perform a predefined function (sequence of computing instructions). Multiple hotkeys and combinations (hotstrings) can now be created and run simultaneously. The aim of this study was to determine if the use of AutoHotkey could improve LIS data entry in the microbiology laboratory.

Technology: Networked workstations (Dell OPTIPLEX GX620 or 745 computers), Windows operating system (Microsoft Windows XP Professional), and LIS (Sunquest v6.2; AutoHotkey, AutoHotkey.com) were used.

Design: AutoHotkey was downloaded from the Internet onto thick-client workstations in the microbiology laboratory at Baystate Medical Center, which performs testing and reports on clinical specimens using Sunquest LIS. Scripts (plain text files containing commands) were created to generate keyboard hotkeys that execute sequential data entry for culture workup and billing processes.

Results: Hotkeys demonstrated potential to improve productivity in the clinical microbiology laboratory in multiple ways. Data entry for routine laboratory culture work into the LIS was faster and easier to perform, with certain procedures that previously required up to 75 keystrokes being reduced to only 2 keystrokes. Hotkeys standardized culture workups by minimizing/eliminating variability previously associated with individual preferences of technologists. Hotkeys automated billing processes previously requiring manual entry of ''bill only'' codes associated with some culture workups.

Conclusions: AutoHotkey is free software that can be used in the clinical laboratory environment to script hotkeys and other macros for automating repetitive tasks in the LIS. The use of hotkeys for keystroke reduction offers an inexpensive and simple method to minimize errors and variability in data entry, as well as to minimize repetitive motion for technologists. This tool allows the efforts of skilled technologists to be most effectively utilized in today's busy laboratory environment.

Systematic Evaluation of Flatbed Scanners for Digital Gross Imaging

Liron Pantanowitz, MD (Liron.pantanowitz@bhs.org); Gabriel Caponetti, MD; John Hunt; Andrew Ellithorpe, MHS; Ann Briancesco; Pamela Passidakis; Christopher Otis, MD. Department of Pathology, Baystate Medical Center, Tufts University School of Medicine, Springfield, Massachusetts.

Context: Gross imaging is an important element of anatomical pathology. Photography of gross surgical and autopsy specimens using digital cameras is being widely adopted. Acquiring digital images of gross pathology with a flatbed scanner (FBS) is fast, reliable, and simple and has been proposed as an alternative method to using a camera. However, there has been no systematic evaluation of this technique. The aim of this study was to evaluate the FBS in gross imaging compared with a digital camera.

Technology: We used HP Scanjet 8270 flatbed scanner (Hewlett-Packard, Palo Alto, California), Sony digital camera DXC-390 (Sony Corporation, Tokyo, Japan), Dell Optiplex 745 desktop computer (Dell, Round Rock, Texas), and CoPath PicPlus (Cerner Corporation, Kansas City, Missouri).

Design: A variety of fresh and formalin-fixed surgical specimens (n = 25) including breast, viscera (colon, kidney, liver, thyroid, uteri, ovary, thymus), lymph nodes, bone, and soft tissue tumors were both photo-

graphed and scanned by 6 prosectors. Specimens were scanned with the FBS lid closed and open. All images were 24 bits per pixel and saved in JPEG format. For each specimen components of the photography (image acquisition time, focus, fitting the specimen in the image, ease of preparation and clean up) and image quality (focus, resolution, and presence of artifacts) were compared.

Results: Photography with a digital camera was preferred by 56% of prosectors, while 32% preferred the FBS, and 12% found both techniques comparable. These findings were unrelated to specimen type. Image acquisition with the FBS took 4 seconds and required no focus adjustment. Photography with the camera took longer, and prosectors experienced difficulty focusing the camera in almost half of the cases (48%). The FBS was able to focus at different depths in the field (3-D structure) in all specimens. Shadows were more commonly seen with the camera. Scanned images with the lid open produced a uniform black background. In only 1 case did a total colectomy specimen not fit onto the FBS plate. Image quality produced by the FBS was rated superior in 44% of cases. Enlargement of images up to 200% was better for those images acquired with the FBS. Fluid (blood, serous, formalin) on the scanner plate in some cases obscured specimen detail, especially in cross sections of solid organs. With compressible specimens the lid of the FBS flattened certain areas.

Conclusions: We recommend the use of a FBS as an inexpensive, complementary imaging tool for acquiring high-quality digital images of gross specimens. The FBS provides rapid image acquisition of all specimen types and is operator independent. Dry specimens without blood (eg, fixed tissue) that can easily fit on to the scanner are ideal for digital imaging with an FBS. To avoid tissue being compressed, gross specimens can be scanned with the lid open.

Breakthrough in Image Analysis Software for Cancer Research

Mark Edward Plaskow, BA (mark.plaskow@nationwidechildrens.org); Kathleen Nicol, MD; David M. Billiter, BS; Thomas Barr, BS. Research Informatics Core, Nationwide Children's Hospital, Columbus, Ohio.

Context: The Research Informatics Core Image Analysis software enables automated electronic scanning and analysis of digital pathology images to locate markers for both pediatric and adult cancerous diseases and offer enhanced and time-saving tools for pathologists and reviewers. Similarities and differences between cancerous diseases were analyzed, creating an ability to search across specimens from many protocols representing many different diseases.

Technology: The Image Analyzer utilizes the .NET technology, in addition to integration with virtual microscopy software, and a Microsoft SQLServer 2005 database for efficient location, persistence, indexing, and retrieval of key areas of interests on analyzed slide images. It runs as both a Web-enabled application and a Windows compiled application.

Design: The image analysis application was designed from a use case developed within the Children's Oncology Group (COG). The use case included defining a pediatric cancer protocol and scanning all cases in the protocol with associated slides. Consultation was conducted with solid tumor (sarcoma) specialists and hematology and leukemia experts.

Results: The first release of the Research Informatics Core team's image analysis software was dedicated to electronic location of abnormal nuclei diffusion within cancerous tumors. An advanced algorithm gives case pathologists and reviewers areas of interest when used against an individual slide specimen image, subset, or entire image archive. The reviewer can inspect the returned data and corresponding digital images using a virtual microscope/image viewer, locating, zooming/adjusting, and annotate desired parts of the image. Subsequent releases aim at a true pathologist dashboard, with the ability to specify criteria across cancerous diseases, including both solid and liquid infections, searching equally on nuclei or white blood cell clustering, cellular mitosis, infectious cellular growth, and so forth. Solid tumor specimens are searchable by tumor necrosis levels, while cancerous blood infections are searchable by white blood cell developmental stages, platelet and red blood cell levels, and so forth. The tool can search for custom cells that are definitive markers (ie, Auer rods in acute myelogenous leukemia).

Conclusions: In conclusion, our image analysis software is doing a great deal of good within cancer research, providing rapidly available tools for pathologists in the identification of specific case types and cancer characteristics using virtual microscopy.

Pathologist Assessments of Focus Quality Using Automated Whole-Slide Imaging for 30 Randomly Selected Surgical Pathology Cases

Russell Silowash, BS[1] (russ.silowash@silonet.com); Laura Drogowski, BS[2]; Drazen Jukic, MD, PhD[2]; Jonhan Ho, MD[2]; Robb Wilson, MA[1]; Leslie Anthony, MA[3]; Anil Parwani, MD, PhD.[4] [1]Departments of Biomedical Informatics and [2]Dermatology, University of Pittsburgh, Pittsburgh, Pennsylvania; [3]Innovative Medical and Information Technologies Center (IMITs), University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [4]Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania.

Context: Whole-slide image technology is increasingly being used in pathology for education, quality assurance, and consultation in the pathology community. There are limited data available on the focus quality of various anatomical structures when whole-slide imaging is used to capture the images. The focus qualities of anatomical structures have significant impact on a pathologic diagnosis. The aims of the current project were to assess the focus quality of images from multiple digital images.

Technology: Slides were scanned with an Aperio T2 scanner and served on a machine equipped with Microsoft Windows Server 2000. Focus quality surveys were collected using a self-reporting Microsoft Access database.

Design: Thirty randomly selected cases comprising 202 slides were scanned using the Aperio slide scanner and made available to participants. Six pathologists viewed digital slides with Aperio's Spectrum Web viewer and ranked overall focus quality using a 5-point Likert scale. Variables such as nuclear, cytoplasmic, noncellular, red blood cell, and lymphocytic detail were ranked with a 5-point Likert scale as well. Surgical pathology specimens were grouped into 13 groups. Focus quality ratings were recorded in an electronic database. Wilcoxon signed ranks tests were performed on all variables of focus quality for every surgical pathology biopsy group. Descriptive statistics for focus quality were also performed.

Results: Data collected from 6 pathologists within the University of Pittsburgh Medical Center resulted in 445 surveys using a screen resolution of 1024 × 768 pixels, and 279 surveys were completed with a screen resolution of 1280 × 1024 pixels. Screen resolution was not accounted for; pathologists used their everyday computers with their typical screen resolution as a default. Focus quality ratings ranged mostly from excellent to fair for each of the focus quality variables. Wilcoxon signed ranks test values had significant results for focus quality variables.

Conclusions: Focus quality ratings show encouraging results validating the use of digital slides as a viable technology for surgical pathology diagnostic purposes. Our data show that monitor screen resolution may have an effect on focus quality interpretation. Additional studies are underway to further characterize other variables that may have an impact on focus quality.

Diagnostic Accuracy in Intraoperative Consultations With Robotic Microscopy and Whole-Slide Imaging

Russell Silowash, BS[1] (russ.silowash@silonet.com); Laura Drogowski, BS[2]; Robb Wilson, MA[1]; Leslie Anthony, MA[4]; Jonhan Ho, MD[2]; Anil Parwani, MD, PhD[3]; Drazen Jukic, MD, PhD.[2] [1]Departments of Biomedical Informatics, [2]Dermatology, and [3]Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania; [4]Innovative Medical and Information Technologies Center, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

Context: Intraoperative consultation is an important component of anatomic pathology practice. Digital pathology is useful when a local pathologist is not available. Tissues analyzed during intraoperative consultation are often thicker than the permanent tissue preparations; therefore, these tissue samples may present challenges to current digital slide technologies in providing pathologists with diagnostic-grade images. Researchers at the University of Pittsburgh Medical Center conducted a study evaluating the capabilities of robotic microscopy and whole-slide imaging in the analysis of intraoperative consultation tissue samples.

Technology: Whole-slide images were created using a Trestle/Zeiss 50-slide loading robotic slide virtualizer paired with an Olympus BX51 microscope equipped with a Jai 3CCD RGB camera. These digital slides were served from a machine equipped with dual 3-GHz Xeon processors with 4 GB of RAM and Microsoft Windows 2000 Server. Air Force and Veterans Affairs' pathologists used an Olympus BX-41 microscope equipped with a Jai M7 camera. Digital slides viewed with robotic microscopy and whole-slide imaging used Trestle's MedMicro software. Data were collected using a Microsoft Access database.

Design: Forty cases randomly selected from University of Pittsburgh Medical Center hospitals were reviewed by the principle investigator, and 20 diagnostic intraoperative slides were identified. Pathologists from multiple sites analyzed these 20 cases using each technology and derived a diagnosis. Data were collected for case complexity, diagnostic confidence, and perceived time to complete. Diagnostic accuracy was rated by the principal investigator and statistically analyzed.

Results: Two University of Pittsburgh Medical Center pathologists, 4 United States Air Force pathologists, and 1 pathologist from a Veteran Affairs' hospital completed 109 total cases. These results are preliminary, and diagnostic accuracy results across digital technologies will be re-

ported as well as data regarding diagnostic confidence, case complexity, and perceived time to complete.

Conclusions: Findings for both technologies are promising. However, there may be differences in diagnostic accuracy when robotic microscopy is compared with whole-slide imaging. More validation work is required, and the University of Pittsburgh Medical Center is conducting studies involving digital slide interpretation.

Online Digital Atlas of Breast Pathology: Utility in Resident Education as Stratified by Postgraduate Year Level

Meenakshi Singh, MD (meenakshi.singh@uchsc.edu); Maxwell Smith, MD; Sharon Sams, MD; Philip J. Boyer, MD, PhD. Department of Pathology, University of Colorado Denver, Aurora, Colorado.

Context: Breast pathology encompasses a diverse range of histologic changes wherein many benign and malignant entities bear histologic resemblance and cause diagnostic dilemmas. Pathology residents find breast pathology to be one of the more challenging areas of surgical pathology with respect to attaining a high level of diagnostic accuracy during training.

Technology: Resources were compiled on a departmental Web server accessed through standard Web pages deploying JavaScript menus, which access multiple indexed JPEG images.

Design: We have developed an easy to use online atlas of breast pathology containing 500 images that cover the range of normal, benign, preinvasive, and invasive malignant entities in breast pathology. Short legends and multiple examples of entities are included. To validate the teaching potential of this atlas, we conducted this study to test the hypothesis that residents with limited surgical pathology experience can show an improvement in diagnostic abilities after they have reviewed the atlas. Breast biopsy sections containing a multitude of benign lesions, some that were diagnostically challenging, were selected. Residents recorded their diagnoses before and after completing review of the atlas. Residents were stratified according to postgraduate year level. Resident identity was kept anonymous.

Results: First-year residents with minimal surgical pathology experience (2–3 months) had a very high misdiagnosis rate before reviewing the atlas. After atlas review, they were able to correct a misdiagnosis of carcinoma, identify a papilloma, and use correct terminology for ductal hyperplasia of the usual type. Second-year residents (4–6 months surgical pathology experience) were able to correct a misinterpretation of lobular carcinoma in situ after atlas review and to correctly identify columnar cell change. Senior residents (fourth year) had a good correlation between their preatlas and postatlas review diagnoses and were able to identify most entities when compared with the breast pathologist's diagnoses, and they picked up additional diagnoses, for example, a radial scar, that none of the junior residents identified.

Conclusions: Online teaching tools, such as this breast atlas, can be useful in preparing residents, particularly junior residents, to make more accurate breast pathology diagnoses.

Electronic Cancer Reporting: Taking Cancer Reporting to the Next Level

J. Mark Tuthill, MD (mtuthill1@hfhs.org); Michael Czechowski, BS; Gary Kasperek, MS; Ron Brown, BS; Joy John Mammen, MD. Division of Pathology Informatics, Department of Pathology & Laboratory Medicine, Henry Ford Hospital, Detroit, Michigan.

Context: Federal and state regulations require centers diagnosing cancer to report results to designated local/regional cancer registry. For our health system, the local registry is at the Henry Ford Hospital (HFH) and the regional registry is the Detroit Metropolitan Cancer Surveillance System at the Karmanos Cancer Center. Traditionally, data originating from the Department of Pathology and Laboratory Medicine underwent a multistep process to be made suitable for transfer to the registries. We sought to increase the efficiency of reporting pathology diagnosis for cancer cases to local and regional cancer registries by implementing a paperless software solution, increase regulatory compliance for handling patient health information (ie, Health Insurance Portability and Accountability Act [HIPAA]), and increase the capture rate and decrease enrollment time for patients into clinical trials at the HFH cancer registry.

Technology: Previously, data were transmitted in the form of paper copies of the final pathology reports for all patients to the HFH registry at the Medical Records Department of the HFH. Each report was manually examined to identify cases that qualified for reporting to the registry. Cases were coded manually for topography and morphology and these data were entered into the registry database. The paper reports of cases that qualified for reporting were then transferred to the regional registry. At the regional registry, the registrar would extract similar data from the paper pathology report. Once the electronic data from HFH registry reached the regional registry, it was manually consolidated and verified against that already in their database. The new technology implemented included laboratory information system (LIS), Sunquest CoPathPlus v2.5 (Sunquest, Tucson, Arizona); servers, Windows Server 2003 (Microsoft, Redmond Washington); and a new application, E-Path (AIM Inc, Toronto, Canada).

Design: In the new paperless, electronic system, after the report is finalized in CoPathPlus, the E-Path software receives HL7 data from CoPathPlus. E-Path processes each report to identify key terms that are indicative of cancer, then formats and transmits the selected positive cases to the HFH and regional registries real-time over a fire-walled, secure virtual private network connection. Interface testing was performed ensuring no loss or distortion of data between all the provider and receiver units. Quality control measures included manual comparison of the reports transmitted online with the simultaneous facsimile transmission of the same case to ensure reliable transfer of identification data and appropriate content.

Results: Software calibration was performed to achieve 100% sensitivity and 98.9% specificity in case finding. The interface has been implemented and deployed in production since February 2007. This has resulted in elimination of paper reporting, saving money and time. Further, as only required cases are transmitted, HIPAA compliance has been achieved. The time for transmission and receipt of cases has been markedly reduced from weeks to minutes.

Conclusions: Implementation of IT solutions leads to savings of resources and help achieve better regulatory compliance. Deployment of the electronic reporting system has resulted in saving time, material resources such as paper and printing supplies, increased regulatory compliance, and improved the quality of cancer registry reporting. Since the system-wide LIS functions from a common database server, designing the interface once has made the same service available to other hospitals in the system. The next phase of deployment will involve studying the time required for enrolling patients into clinical trials at the HFH cancer registry and seeing if there is an increase in the capture rate. We also hope to develop interfaces between the local registries and E-Path to further data automation, data extraction, and report generation that would result in further efficiencies.

Clinical and Translational Research Centers Portal Page: A Web-Based Solution to Centralizing Services

Joyce Zelnis[1] (zelnisjb@upmc.edu); Brennen Flaherty[1]; Casey Holderfield[1]; Daniel Goldberg[2]; Eric Richie.[2] [1]Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania; [2]Information Services Division, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

Context: The University of Pittsburgh Clinical & Translational Science Institute CTRC (Clinical & Translational Research Centers) main portal page is the centralized point of access for CTRC personnel to various CTRC tools, as well as the entry point for requesting CTRC services by primary investigators and their staff. Useful online resource links have been provided as well as information that is important for completion of the CTRC service request form.

Technology: Important components of the software development were to provide a user-friendly interface, ease of navigation, and consistent user interfaces. The Web-based CTRC portal pages are written with ColdFusion/HTML and JavaScript. The service request form is a ColdFusion application that reads and writes to an Oracle database.

Design: The CTRC portal page was designed with multiple functions in mind. Members of the CTRC enter the portal to access their site-specific software tools, which include scheduling systems, protocol management systems, and visit tracking software. Additional design elements included providing the CTRC customers access to an online form that prompts them for all of the information necessary to request services from the CTRC. The CTRC annual reporting requirements were instrumental in the design of the service request form.

Results: The CTRC portal page had more than 90 users in its first 6 months of operation. Investigators and their staff submitted the information necessary for more than 70 service requests forms to the CTRC. A submitted service request form generates a .PDF file that is included in the submission to the University of Pittsburgh Institutional Review Board online submission process, also known as OSIRIS. Initial user feedback has been positive, with user-suggested enhancements scheduled to be implemented in the coming months.

Conclusions: The CTRC portal page has provided a centralized location for CTRC activities. CTRC customers and personnel alike go to this location for software applications, information, and service requests. As the protocols get scheduled at the CTRC the information entered by the CTRC customers will be updated into the appropriate software, taking the entered protocol data from requested services through patient visit.