

Original Article

The 2013 symposium on pathology data integration and clinical decision support and the current state of field

Jason M. Baron^{1,2}, Anand S. Dighe^{1,2}, Ramy Arnaout^{3,4,5}, Ulysses J. Balis⁶, W. Stephen Black-Schaffer^{1,2}, Alexis B. Carter^{7,8}, Walter H. Henricks⁹, John M. Higgins^{1,5,10}, Brian R. Jackson^{11,12}, JiYeon Kim¹³, Veronica E. Klepeis^{1,2}, Long P. Le^{1,2}, David N. Louis^{1,2}, Diana Mandelker^{2,14}, Craig H. Mermel^{1,2}, James S. Michaelson^{1,2,15}, Rakesh Nagarajan^{16,17}, Mihae E. Platt^{1,2}, Andrew M. Quinn^{2,14}, Luigi Rao^{1,2}, Brian H. Shirts¹⁸, John R. Gilbertson^{1,2}

Departments of ¹Pathology and ¹⁵Surgery, and ¹⁰Center for Systems Biology, Massachusetts General Hospital, MA, ²Harvard Medical School, MA, ³Department of Pathology and ⁴Division of Clinical Informatics, Department of Medicine, Beth Israel Deaconess Medical Center, MA, ⁵Department of Systems Biology, Harvard Medical School, ⁶Division of Pathology Informatics, University of Michigan Health System, MI, Departments of ⁷Pathology and Laboratory Medicine, and ⁸Biomedical Informatics, Emory University School of Medicine, GA, ⁹Cleveland Clinic, Center for Pathology Informatics, Pathology and Laboratory Medicine Institute, OH, ¹¹ARUP Laboratories, UT, ¹²Department of Pathology, University of Utah School of Medicine, UT, ¹³Regional Reference Laboratories, Southern California Permanente Medical Group, CA, ¹⁴Department of Pathology, Brigham and Women's Hospital, MA, Departments of ¹⁶Pathology, Immunology and Laboratory Medicine, ¹⁷Genetics, Washington University School of Medicine, MO, ¹⁸Department of Laboratory Medicine, University of Washington School of Medicine, WA

Email: *John Gilbertson - jrgilbertson@partners.org, *Anand Dighe asdighe@partners.org

*Corresponding authors

Received: 14 October 13

Accepted: 08 December 13

Published: 31 January 14

This article may be cited as:

Baron JM, Dighe AS, Arnaout R, Balis UJ, Black-Schaffer WS, Carter AB, et al. The 2013 symposium on pathology data integration and clinical decision support and the current state of field. *J Pathol Inform* 2014;5:2.

Available FREE in open access from: <http://www.jpathinformatics.org/text.asp?2014/5/1/2/126145>

Copyright: © 2014 Baron JM. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Background: Pathologists and informaticians are becoming increasingly interested in electronic clinical decision support for pathology, laboratory medicine and clinical diagnosis. Improved decision support may optimize laboratory test selection, improve test result interpretation and permit the extraction of enhanced diagnostic information from existing laboratory data. Nonetheless, the field of pathology decision support is still developing. To facilitate the exchange of ideas and preliminary studies, we convened a symposium entitled: Pathology data integration and clinical decision support. **Methods:** The symposium was held at the Massachusetts General Hospital, on May 10, 2013. Participants were selected to represent diverse backgrounds and interests and were from nine different institutions in eight different states. **Results:** The day included 16 plenary talks and three panel discussions, together covering four broad areas. Summaries of each presentation are included in this manuscript. **Conclusions:** A number of recurrent themes emerged from the symposium. Among the most pervasive was the dichotomy between diagnostic data and diagnostic information, including the opportunities that laboratories may have to use electronic systems and algorithms to convert the data they generate into more useful information. Differences between human talents and computer abilities were described; well-designed symbioses between humans and computers may ultimately optimize diagnosis. Another key theme related to the unique needs and challenges in providing decision support for genomics and other emerging diagnostic modalities. Finally, many talks relayed how the barriers to bringing decision support toward reality are primarily personnel, political, infrastructural and administrative challenges rather than technological limitations.

Key words: Clinical decision support, genomics, interpretive reporting, machine learning, test utilization

Access this article online

Website:

www.jpathinformatics.org

DOI: 10.4103/2153-3539.126145

Quick Response Code:



INTRODUCTION

Pathologists and informaticians are becoming increasingly interested in the application of electronic clinical decision support to pathology and laboratory medicine. Application of decision support has great potential to optimize laboratory test selection, improve test result interpretation and permit the extraction of enhanced diagnostic information from existing laboratory data.^[1-3] This may in turn help to transform laboratory medicine from primarily an observational field to one more centered on interpretation and definitive, comprehensive and precise diagnosis. Anatomic pathology has similar potential to enhance the diagnostic information it delivers.

This transformation of pathology may not only help reduce the waste or errors frequently associated with laboratory test selection and result interpretation,^[1,4-7] but may also enable a previously unattainable level of diagnostic precision. Advances in laboratory automation, next generation sequencing, mass spectrometry and other emerging data acquisition modalities will surely enhance laboratory efficiency and diagnostic value. However, substantial improvements in the diagnostic value of laboratory tests will also likely result from more effective use and interpretation of data from traditional assays and existing technologies. Clinical decision support for test selection^[1,8] and test result interpretation^[9] may help to avoid unnecessary testing, ensure correct tests are ordered and avoid misinterpretation of test results. Moreover, application of statistical, computational and machine-learning techniques to clinical and laboratory data may reveal key patterns and insights that manual interpretation of the data could not.^[2,3] In fact, many of the barriers to expanding the clinical application of next generation sequencing and other diagnostic modalities lie not in the data collection, but in the analysis and interpretation.^[10]

Transforming the focus of pathology and laboratory medicine to more broadly emphasize data analysis and personalized diagnosis may also be of key importance in maintaining the relevance of the specialty and enhancing the value of laboratory testing. In particular, expanding use of automation is leading to an increased perception that certain routine laboratory services can be thought of as commodities with cost being the only real consideration.^[11,12] A key strategy for overcoming such commoditization of laboratory services will be to understand that even if generation of certain laboratory data becomes routine, the process of extracting useful diagnostic information from this data will become increasingly complex.^[13] Laboratories and pathology services could shift their primary focus from creating data to generating diagnostic, prognostic and therapeutic information.^[1,9,13,14] Furthermore, particularly with

increased emphasis on cost-containment and utilization management, compounded by an increase in the complexity of testing (such as genomic analyses), it will be increasingly important for laboratories to assist with test selection.^[15,16] Computational data analysis and decision support systems will almost certainly be integral to this evolution in pathology and laboratory diagnosis.^[1,8]

Nonetheless, pathology decision support is still in its infancy. To facilitate the exchange of ideas and preliminary studies, we convened a symposium entitled: Pathology data integration and clinical decision support. Here we summarize some of the views shared in the symposium to provide an overview of the current state of the field and the challenges it faces.

METHODS/MEETING ORGANIZATION AND STRUCTURE

The symposium on Pathology data integration and clinical decision support, held at the Massachusetts General Hospital (MGH), Boston, MA on May 10, 2013, was sponsored by the MGH Department of Pathology and the Partners Fellowship Program in Pathology Informatics and was organized by JB, AD and JG. The speakers were selected to provide varied interests and backgrounds with half of the speakers coming from within the MGH Department of Pathology and half from institutions around the country. The 1-day meeting was divided into four subtopics and included 16 plenary talks, ranging from 20 to 25 min. The symposium also included three panel discussions during which questions from the audience were answered and discussed. This paper provides a brief summary of each plenary talk. In addition, the paper concludes with a synthesis of the key themes of the symposium and provides some “next-steps” for the field based upon the symposium presentations and discussions.

RESULTS

Block 1 Presentations: Big Picture Concepts and the Need for Pathology Data Integration and Decision Support

David Louis, MD

The first talk of the day was by David Louis, Pathologist-in-Chief of the MGH Department of Pathology and Benjamin Castleman Professor of Pathology at Harvard Medical School. Dr. Louis' talk, “The Skeleton of Computational Pathology,” described a vision for the future of pathology along with a “skeleton” of the “computational pathology” efforts ongoing within the MGH Department of Pathology. In particular, Dr. Louis first noted that laboratories and pathologists currently provide clinicians mostly with relatively discrete elements of diagnostic data, including

anatomic pathology interpretations, laboratory results and increasingly, “omics” data. Clinicians then must interpret this pathology data in the context of the clinical findings and medical knowledge to arrive at a diagnosis or treatment plan. However, moving forward, pathology services may add value by integrating within a single synthesized report an interpretation of the various elements of data they produce. The integrated report will interpret findings and results within the context of the particular clinical setting, utilizing medical knowledge derived from databases and medical literature.

Dr. Louis outlined an approach for reaching this goal and the final product may involve a pipeline containing at least six components: Clinical data integration, mathematical biologic modeling, clinical decision support, “omics and imaging” (algorithmically processing high complexity datasets), integrated reporting and performance analysis. In this pipeline, clinical data integration serves as the input step, integrated reporting is the output step and clinical decision support, mathematical modeling and “omics” and imaging comprise the “stuff in the middle,” processing, integrating and interpreting the data. Performance analysis will be used to monitor the entire process to identify areas for improvement, analyze findings in the context of health care resource utilization and demonstrate the clinical and economic value of the approach.

Dr. Louis further described the efforts ongoing within the MGH Department of Pathology to make this vision a reality. The department has organized six working groups corresponding to the six components noted above with over 80 members of the pathology department (faculty and trainees) participating in at least one working group. At the current time, the groups are carrying out analyses and designing projects in their assigned area. A steering committee that includes the chairpersons of each working group meets periodically to coordinate and integrate the work and activities across all six groups. Department-wide meetings in computational pathology occur about every 6 months and are intended to update the faculty and trainees, to get further input from members of all working groups and to encourage other faculty and trainees to join the overall effort.

Brian Jackson, MD

Following Dr. Louis’ talk emphasizing the “what,” Brian Jackson, Chief Medical Informatics Officer at ARUP Laboratories and Associate Professor of Pathology (Clinical) at the University of Utah helped to elucidate the “why.” His talk, “The Psychology of Decision Making: Why Don’t Those Darned Doctors Use Our Tests Properly?” introduced two important ideas:

1. The human brain is ineffective at deriving information from multidimensional data (i.e., many discrete elements such as with a set of many laboratory values from the same patient).

2. The brain is poor at “meta-cognition” and people often do not have accurate insight into why they make certain decisions. Similarly, the confidence that people place in the decisions or predictions they make may be impacted by factors not particularly relevant to the decision.

To illustrate these ideas, Dr. Jackson presented several studies. One small study,^[17] first surveyed two rheumatologists to discover which factors they each considered most important in assessing disease activity in rheumatoid arthritis patients. The rheumatologists were then asked to assess disease activity in a set of patients based upon chart review. The study found that the rheumatologists thought that they were putting significant weight on all five of the factors surveyed when making their assessments. However, in practice, each rheumatologist based the assessments almost entirely on a single factor. Dr. Jackson noted that, to the extent that this small study is generalizable, the results illustrate that the physicians may make medical decisions without really understanding the basis for these decisions.

Dr. Jackson next presented a study of “horserace handicappers” who attempt to predict the outcomes of horse races by using various known “predictors.”^[18] This study demonstrated that as the handicappers were presented with an increasing number of discrete data elements on which to make predictions (from 5-10 to 20-40) the accuracy of their predictions remained unchanged, but the confidence in their predictions increased. To the extent that this horserace study can be extended to clinical diagnosis, Dr. Jackson concluded that “flooding doctors with data does not improve diagnostic accuracy, but it probably contributes to overconfidence.”

At the conclusion of his talk and during a panel discussion, Dr. Jackson considered the questions of where the “bottlenecks” are occurring in laboratory-based diagnosis. In particular, Dr. Jackson argued that we produce more data than we can currently use; we need to shift some of our emphasis away from increasing the availability of new data and toward better applying and extracting information from the data we already produce.

John Gilbertson, MD

John Gilbertson, Associate Chief for Informatics in the Department of Pathology at the MGH and Associate Professor at Harvard Medical School, presented “Pathology decisions and decision support.” In this talk, he described some of the technical and administrative limitations that will need to be overcome to provide the next generation of decision support and some solutions currently in progress. Among these challenges is that laboratory information systems (LIS) traditionally emphasize the technical aspects of the laboratory while providing comparatively little support for professional or interpretive activities. For example, whereas current LIS

systems support aspects ranging from specimen tracking to test or special stain ordering to billing, these systems often provide little more than a word processor for pathologists to generate interpretive reports. Likewise, clinical pathology systems often have relatively limited calculation functionality, significantly hindering the ability to implement advanced interpretive algorithms. Molecular LIS modules tend to be particularly limited with regard to support of professional interpretation.

Dr. Gilbertson proposed that LIS vendors supplement their traditional “technical LIS” systems by adding to them “professional LIS systems” with expanded capacity to facilitate pathologists’ decision making, accurate and efficient sign out and advanced interpretive algorithms. With regard to the latter, certain functions may not need to be directly within the main LIS, but could exist in external systems; nonetheless, LIS vendors would need to generate highly flexible application programming interfaces to enable such external functionality.

Other needs as pathology related clinical decision support advances include systems for advanced analytics, robust electronic health records (EHRs), structured, high quality data, pipelines for molecular and genomic pathology and infrastructures for data extraction, warehousing and computation. Tissue registries and digital pathology systems are other factors that may support various aspects of research and decision support. A final and perhaps most significant need are personnel with the skills, knowledge and desire to advance pathology decision support. This includes not only highly trained and skilled pathology informaticians, but also non-informatician pathologists with informatics literacy. Information technology (IT) and technical staff with these skills and interests are essential.

Solutions to many of the challenges are being currently considered in the computational pathology initiative within the MGH Department of Pathology (see description of the presentation by David Louis for additional detail). Other strategies include the development of informatics faculty consisting primarily of practicing pathologists (or researchers) with a secondary practice in informatics and the creation of a large and robust pathology informatics fellowship training program to address the personnel needs. MGH pathology is involved in a co-development agreement with Sunquest Information Systems to help develop a more advanced LIS; Dr. Gilbertson is working with Sunquest to expand the professional LIS functionality of the system.

Stephen Black-Schaffer, MD

Stephen Black-Schaffer, Associate Chief for Education and Training in the Department of Pathology at the MGH and Associate Professor of Pathology at Harvard Medical School, presented the final talk of the first block: “Clinical decision support: Implications for Pathology Performance Assessment and Trainee Education.” In this

talk, he addressed how Pathology Departments can use the comprehensive clinical data they produce to assess the performance of both the operation as a whole and of individual pathologists, including trainees. Operational performance assessment is fundamentally quite similar to pathologist performance assessment and can use much of the same data.

In particular, Dr. Black-Schaffer conveyed the perception that Pathology Departments and Hospitals will increasingly need to assess and improve operational performance given increased emphasis on cost containment and shifting reimbursement models favoring efficiency. “Leakage” must be avoided. Dr. Black-Schaffer provided an example of how pathology and laboratory test performance can be analyzed to optimize the use of molecular testing on fine needle aspiration specimens taken from thyroid nodules.

Dr. Black-Schaffer also described how the data in the LIS could be used for performance assessment. In particular, resident sign-out could be compared with final attending review to objectively assess how well trainees are meeting key educational milestones and competencies. However, to do this in an automated fashion, LIS systems will have to store preliminary interpretations provided by trainees, in addition to final reports, in a structured form. A similar approach could also be used to evaluate peer-review of pathologist competency.

Block 2 Presentations: Clinical Decision Support for Resource Utilization and Pioneering Diagnostic Modalities Including Genomics and Systems Biology-Based Models

Rakesh Nagarajan, MD, PhD

Rakesh Nagarajan, Associate Professor, Pathology and Immunology and Associate Professor, Genetics at Washington University School of Medicine, delivered the first talk of the second block: “Clinical Genomicist Workstation: Analyze, interpret and report Nextgen based molecular diagnostic studies.” He began this talk by noting the challenges involved in converting genomic sequencing data into actionable information. Most of the talk was devoted to describing a platform developed and used at Washington University in St. Louis to help overcome some of these challenges. This platform, the “Clinical Genomicist Workstation,” helps transform raw sequence data into an interpretive report by processing data in several tiers. The first tier is “seamless” from the perspective of the clinical genomicist and automates basic data processing such as sequencing alignment and variant calling. Given the clinical application of this system, this first tier tracks key information that may be critical to future re-evaluation such as the version of tools used and parameters specified.

Whereas the first tier focuses on identifying and calling genomic variants, the second tier begins to

ascribe meaning to these variants. It applies data from various genomic databases and knowledge repositories to identify the subset of variants with known clinical significance (and which are relevant to the case at hand). It also provides phenotypic information for these relevant variants. This step helps to reduce the number of variants the clinical genomicist must review from an overwhelming number to a more manageable number and conveniently provides clinical knowledge about the variants of known significance. The genomicist workstation also provides tools for automated clinical report generation and data visualization.

While tiers one and two involve processing data for individual patients, the third tier serves to curate the knowledge repository and the rules that can be used to assign clinical significance to variants. The knowledge curated in this tier helps to determine which variants are displayed to the clinical genomicist and under which circumstances. For example, selection of variants to display may be tailored to the patient's test genomic test order (panel), phenotype, disease or medication list. Dr. Nagarajan also provided some planned or potential future enhancements for the system, including sharing of tier 3 knowledge repositories and rules across institutions.

John Higgins, MD

John Higgins, Assistant Pathologist in the Department of Pathology at MGH and Assistant Professor in the Department of Systems Biology at Harvard Medical School, delivered a talk entitled, "Systems Biology in Clinical Medicine: Mathematical Model-Based Diagnosis." In it, Dr. Higgins argued against the view held by some that as available data becomes more expansive, traditional hypothesis-driven mechanistic biologic and clinical research will become obsolete and will be replaced by data mining. He suggested to the contrary that models may be necessary to help make sense of the data. While statistical data mining routines by themselves are prone to overfitting and may make differentiating "signal from noise" difficult, placing data in the context of mechanistic physiologic models helps constrain the set of hypotheses from that which a raw statistical approach might use and make true "signal" easier to identify. Likewise, Dr. Higgins argued that traditional human clinical intuition alone is inadequate for optimal interpretation of the large datasets that are now being developed, because "human intuition doesn't scale to multidimensional data" but "human subjective bias does."

Dr. Higgins went on to define and describe systems biology as "the formulation and analysis of mechanistic models of systems of interacting biological components." Models need to be of appropriate detail and complexity to capture essential patterns, inputs and associations without being so complex as to be impractical to fit

and compute and useless for helping us understand how the biological systems actually work. After discussing the importance of models, he devoted much of the talk to providing a specific example of a systems biology model of the lifecycle of circulating red blood cells that was derived and validated with existing clinical laboratory data. The model relies on single-red blood cell measurements routinely collected on a hematology analyzer and enables the inference of rates of blood cell maturation and clearance in the peripheral circulation. The model allows us to infer and quantify aspects of pathophysiology that cannot be measured directly and this additional information may enable earlier and more accurate diagnosis of conditions such as iron deficiency anemia. This model, described elsewhere in detail,^[19] provides an important example, illustrating how through substantive modeling, we can extract substantially enhanced information from the data we already generate.

Dr. Higgins concluded his talk with a quote from Nobel Laureate, Sydney Brenner that captured many of the arguments from the talk: "The orgy of fact extraction in which everybody is currently engaged has, like most consumer economies, accumulated a vast debt. This is a debt of theory and some of us are soon going to have an exciting time paying it back-with interest, I hope."^[20]

Long Phi Le, MD, PhD

Dr. Long Le, Assistant in Pathology in the Department of Pathology at the MGH and Assistant Professor of Pathology at Harvard Medical School, described, in his talk, "Decision Support for Clinical Genomics," the approach that the MGH is taking to develop a platform for tumor genotyping by next generation sequencing. Dr. Le argued that tumor therapeutics will be increasingly targeted to the individual mutational profile of the patient's tumor. Like many of the genomics related talks, Dr. Le's talk described the shift from the targeted detection of medically actionable mutations of known significance to next generation sequencing, a technique that will lead to the detection of many novel or rare variants with unknown diagnostic, prognostic and therapeutic implications.

The current tumor sequencing goal at MGH encompasses ">1000 genes (~3.6 Mb), ×100 minimum coverage 10 bp into introns, 6-8 Gb of data/tumor-normal pair, 5-10% analytical sensitivity (with regard to tumor cellularity) and 3-4 week turnaround time." To help reach this goal Dr. Le *et al.* have developed their own molecular laboratory information management system termed "Center for Integrated Diagnostics, Wiki, Laboratory Information Management System". Based on the Semantic Mediawiki semantic web technology, this system supports computerized provider order entry (CPOE) to allow clinicians to directly input orders, a key to achieving an effective upfront workflow

and optimal, accurate test selection in the molecular diagnostics laboratory. It also supports the full breadth of the molecular laboratory internal workflow, including asset tracking (e.g. slides, blocks, nucleic acid extractions and reagents), slide labeling, assay worksheets, resulting, reporting, case tracking and document management. The system is “semi-integrated” with the main anatomic pathology LIS and has full query functionality of the stored structured data. Its scalability and capability for deployment outside of MGH is currently under investigation. Work to integrate a sequence interpretation pipeline into this setup is underway with collaboration from Dr. Gad Getz.

Finally, Dr. Le described a possible variant categorization scheme for NGS cancer genotyping reporting. Based on available evidence, variants with actionable treatment would be classified in the therapeutic category and further differentiated as “consensus” (based on documented standard of care consensus guidelines from the Food and Drug Administration [FDA], College of American Pathologists, or National Comprehensive Cancer Network) or “emerging” (based on availability of experimental drugs reported to be effective in late trials, early trials, case reports, or preclinical studies). Two other categories of action ability include “diagnostic” and “prognostic” markers with strong evidence from the literature. Variants which do not fall into the above categories but have some association with cancer pathogenesis would be reported under mutations with “other relevance.” Finally variants of clear unknown significance would be grouped by their functional pathways and reported at the end. Currently, the laboratory’s Semantic Mediawiki information management system is being expanded to support the curation of the variants and also structured reporting for the MGH cancer genotyping assay.

Brian Shirts, MD, PhD

Brian Shirts, Assistant Professor in the Department of Laboratory Medicine at the University of Washington, offered a complementary perspective on the challenges of providing decision support for genomic testing in his talk: “Complexity, Uncertainty and Constant Change: Decision Support for Clinical Next-Gen Sequencing.” Dr. Shirts began by highlighting an ironic tendency in laboratory testing: The more stable that a raw laboratory results is for an individual patient over time, the less likely it is that the general interpretation guidelines for that analyte will remain stable. For example, electrolytes can change rapidly and thus remain relevant for only a short period of time, but the interpretation guidelines for electrolytes have remained constant for decades. By contrast, germline genomic data will not change much over a patient’s life; however, given the rapidly evolving state of medical knowledge, the interpretation of a genetic variant may only remain current for a relatively short period of time. This irony highlights one of the

key challenges in genomic testing: Data may need to be reinterpreted regularly throughout a patient’s life. The current processes, infrastructure and systems used in clinical and laboratory diagnosis are not well-adapted to a future where, even in the absence of new patient data or clinical changes, a laboratory result may require regular reinterpretation.

Dr. Shirts noted that the current practice in genomic data interpretation involves substantial non-automated interpretation. Genetic counselors are highly valuable in this regard and can “pay for themselves” by promoting appropriate and cost effective test utilization. In fact, a study from ARUP Laboratories demonstrated a very substantial cost saving attributable to genetic counselors actively managing genetic test ordering.

Moving forward, the desire will increasingly be to develop algorithmic and computational approaches to facilitate the interpretation and application of genomic data. However, there are inherent statistical limitations that may prevent automated genetic informatics implementation. Dr. Shirts used the example of partitioning test reference ranges based upon genomic information, with a strategy for doing this previously described in detail,^[21] to illustrate one situation where limitations to integration of genomic and clinical information can be clearly illustrated. Another challenge is that many variants are rare, occurring in only a small portion of the population. It may be very difficult to determine the significance of these rare variants, even when they have large effects, due to the statistical limitations similar to those associated with finding patterns in multidimensional data when only a small number of data points are available (see discussion of overfitting in the talk by Jason Baron).

The field of genomic testing needs automated platforms for genomic data that integrate structured clinical data with genomic data that allow for facilitation of correlation and interpretation. Dr. Shirts stated that “personalized medicine is about the relationship between signal and noise,” and it is important for pathologists to help develop systems that filter the signal (clinically actionable data) from the noise (inter-individual genetic variability of no clinical significance).

Craig Mermel, MD, PhD

Craig Mermel, a resident in Clinical Pathology at the MGH with extensive experience in cancer genomics research, described some of the infrastructural challenges that will need to be addressed to facilitate routine tumor genotyping by next generation sequencing at the whole-exome or whole-genome levels. In delivering his talk, “Infrastructure and Reporting Challenges to Clinical Next-Generation Sequencing Programs,” Dr. Mermel drew upon and contrasted his experience in tumor genomics at the Broad Institute, to his experience helping to establish a next generation sequencing pipeline in the clinical setting.

In evaluating some of the infrastructure challenges, Dr. Mermel argued that we need to be aware of several trends. First, sequencing capacity is expanding much more rapidly than computational capacity. Sequencing costs have declined by more than four orders of magnitude over the past decade to < 10 cents/megabase, while computational capacity follows roughly “Moore’s Law” doubling only about every 18-24 months. Currently, the bottleneck in many aspects of sequence analysis is computation and the imbalance between ability to produce sequence data and ability to analyze it is likely to further exacerbate moving forward. Paralleling this expansion of sequencing capacity, the number and types of sequencing platforms is expanding and shifting (Available from: <http://www.omicsmaps.com> for more up to date information) and new algorithmic techniques are rapidly being developed. Finally, the knowledge base is rapidly expanding; new mutations of functional significance are curated every day.

While all these aforementioned trends challenge both research and clinical applications of genomic sequencing, clinical applications have some additional considerations that must be addressed. These include requirements to comply with complex regulations (e.g. Health Insurance Portability and Accountability Act (HIPAA), Health Information Technology for Economic and Clinical Health Act (HITECH), Clinical Laboratory Improvement Amendments (CLIA), FDA), to interface with clinical information systems and workflows, to meet clinical standards with regard to information system stability, validity and consistency and to provide clinically acceptable turn-around times.

While the field has not solved all of these challenges, several developments may be inevitable. Foremost, most of the analysis will need to be automated; manual evaluation of variants beyond a final report would be prohibitively expensive. In fact, Dr. Mermel estimated that a single whole exome sequence would on average have approximately 500 novel variants, costing at least \$2,000 in pathologist time to review (assuming each variant could be reviewed in 1 min). In addition, it is not feasible for all institutions and research environments to develop their own analysis tools. We will need to work together as community to share tools and even analysis platforms.

Finally, Dr. Mermel argued that we should strongly consider using cloud architecture for our main sequence analysis needs. Among the advantages of the “cloud” are that rather than needing computer clusters capable of meeting peak needs and thus often leaving excess, unused capacity, cloud computing more efficiently distributes computing resources and only requires institutions to pay for the computing used. In addition, cloud architectures are potentially more nimble and adaptable to shifting analysis tools. Cloud based platforms may also facilitate

inter-institutional collaboration, including sharing of analytic pipelines.

Alexis Carter, MD

Alexis Carter, Director of Pathology Informatics and Assistant Professor in the Department of Pathology and Laboratory Medicine and Department of Biomedical Informatics at Emory University School of Medicine, offered insight on transfusion clinical decision support. In her talk, “transfusion guidelines versus practice: The impact of clinical decision support tools on transfusion behavior,” she described work that she and colleagues did at Emory to improve the utilization of blood products. Blood products are commonly misused, most often in the direction of transfusing a patient when not indicated. At the very least, this puts patients at risk for harm while engendering unnecessary cost for the institution. Dr. Carter *et al.* undertook an effort to understand why physicians transfuse patients in an ethnographic study and this is being followed by a study looking at decision support systems in the EHR that will help encourage physicians to transfuse according to national^[22-24] and institutional guidelines. Dr. Carter devoted the majority of her talk to discussing their specific findings which are unpublished data. Therefore, these findings are reserved for subsequent publication.

Block 3 Presentations: Technical Strategies and Methods for Implementing Data Integration and Decision Support

Ulysses Balis, MD

Ulysses Balis, Associate Professor and Informatics Director in the Division of Pathology Informatics in the Department of Pathology in the University of Michigan Health System, introduced the emerging area of multi-analyte assays with algorithmic analysis (MAAAs), in his talk, “Machine-learning-based thiopurine monitoring and decision support as an exemplar of encoded data use in clinical settings.”

MAAAs work by applying machine learning or other algorithms to a panel of individual laboratory test results from a patient to generate diagnostic information that the individual results, interpreted in isolation, could not provide. Among the reasons why MAAAs may be useful is that the human brain is limited in its ability to discern complex patterns from high-dimensional data, such as a panel of results (see description of the talk by Brian Jackson). In contrast, however, these patterns may be discernable to computers using machine learning algorithms. Two widely known examples of MAAAs include liver fibrosis algorithms^[25] and tests for fetal abnormalities (e.g., the “quad screen”).

Dr. Balis devoted a substantial portion of his talk to describing a particular MAAA developed by colleagues at the University of Michigan to monitor therapy with

thiopurine medications. Thiopurines, commonly used to treat inflammatory bowel disease, require monitoring because they have narrow therapeutic indices and not all patients respond. Metabolite levels are sometimes assayed to for use in monitoring but this metabolite testing is expensive and not optimally predictive of response. A MAAA developed at the University of Michigan used a random forest classifier (an established machine learning technique) to predict patients' thiopurine responses based upon a panel of routine hematology and chemistry analytes. Not only is this MAAA much less expensive, but actually more accurate in monitoring response, as compared to the metabolite test. This MAAA has been described in detail in Waljee *et al.*, 2010.^[26]

One challenge to implementing MAAAs and related approaches is that traditional LISs are incapable of executing the complex algorithms needed to integrate the individual results. The thiopurine MAAA was implemented using an external, distributed architecture (the LIDDEx System) interfacing with, but existing outside the LIS. However, to optimally apply such approaches going forward, LIS vendors will need to improve their systems' functionality to facilitate data extraction and interaction with external systems and improve their ability to implement advanced data processing algorithms.

Jason Baron, MD

Jason Baron, Assistant in Pathology at the MGH and Instructor of Pathology at Harvard Medical School, presented "Pathology Decision Support Meets Big Data." In this talk, he discussed emerging opportunities to identify novel patterns in clinical and laboratory data that may enhance decision support and increase the diagnostic information generated through laboratory testing. He also noted many of the accompanying challenges.

Dr. Baron began by explaining that in many cases, groups of patients exhibit wide variability in responses to treatment or ultimate outcomes, despite presenting similarly in terms of diagnoses, comorbidities and other currently known predictors of clinical course. His hypothesis is that by mining large sets of clinical and laboratory data, we may be able to identify subtle patterns that can predict prognosis, response to therapy or ideal clinical management in ways that standard manual interpretation of data cannot. Among the philosophical rationales Dr. Baron provided for this hypothesis is the reality that the human brain is incapable of optimally analyzing high-dimensionality data, leaving many "unexplored" opportunities for refinement in diagnosis. However, perhaps even more significant is that identification of subtle patterns requires large sets of data (see discussion of overfitting below) and even the busiest clinicians can only see a relatively small number of patients during the course of their careers. In contrast,

a computer can "learn" from mining millions of patient records from a large health system or inter-institutional data sharing program.

After noting that large data sets will be necessary to identify complex patterns, Dr. Baron devoted the majority of his talk to answering the question, "Why 'big' data" and the related issue of the potential tradeoff between data size and data quality. To answer this question, he provided a very brief conceptualization of supervised machine learning using a project to identify spurious glucose results as an example.^[27] In particular, he discussed the idea of overfitting. Overfitting occurs when a machine learning model fits to random patterns within a set of "training" data that do not generalize. Overfit models "mistake noise for a real pattern" and will perform better in classifying training data than an independent set of test data. Overfitting will tend to increase with the complexity of the model being fit and decrease with the size of the training data used to fit the model. Model complexity parallels degrees of freedom and thus models incorporating many parameters (fit to high-dimensional data) will tend to be complex and prone to overfitting. Because these are the types of models we may wish to fit using pathology data, we will need large datasets from large health systems and eventually, large inter-institutional data exchange networks.

Another consideration Dr. Baron discussed is the potential tradeoff between data size and data quality. Data quality may be limited by factors including completeness, accuracy, accessibility and structure. In the case of structure, there may be a real tradeoff between size and quality, because although unstructured data can be manually encoded to a more structured format, the resources required to do so are roughly proportional to the size of the dataset. Although high-quality, high-quantity data would be optimal from a data mining point of view, in some cases, it is possible that a very large dataset may be useful even if relatively low quality. Google Flu Trends^[28] is an example in which high-quality information (influenza trends) is derived from very high quantity but, in many regards, low quality data (unstructured flu-related Google internet searches).

Ramy Arnaout, MD, DPhil

Dr. Ramy Arnaout, Assistant Professor of Pathology and Associate Director of the Clinical Microbiology Laboratory at the Beth Israel Deaconess Medical Center (BIDMC) and faculty in Clinical Informatics at BIDMC and the Systems Biology PhD program at Harvard Medical School, discussed the landscape of inappropriate laboratory testing in medicine in his talk "Clinical Laboratory Data by the Numbers." As pathologists look to update the specialty's value proposition for the age of big data, a fitting question is how well clinicians interact with the

single largest source of data in pathology: Laboratory testing.

Dr. Arnaout pointed out that laboratory testing is the single highest-volume medical activity, with over six billion tests performed each year in the United States alone. He also noted that the prevailing narrative at hospitals in regard to laboratory testing is that unnecessary repeat testing is the crux of the problem and is bankrupting medicine. He then proceeded to describe results from a 15-year meta-analysis, in which his group reviewed over 1.5 million orders covering 46 of the 50 most frequently ordered tests, that suggested that this narrative is categorically wrong. “Inappropriate laboratory utilization is widespread, but it’s not where we think,” he said. Understanding when and how inappropriate utilization occurs, he argued, gives pathologists a tremendous opportunity to reshape information flows to make this high volume medical activity better serve patients and clinicians. The data presented in Dr. Arnaout’s talk is available in a recent publication.^[29]

Block 4 Presentations: Practical Consideration and Strategies to Bring Big Picture Ideas to Implementable Endeavors

JiYeon Kim, MD, MPH

JiYeon Kim, Physician-in-Charge of Chemistry and Laboratory Informatics at the Regional Reference Laboratories of Southern California Permanente Medical Group, delivered a talk, “Lab Data and Patient Outcomes.” She began her talk by comparing the implications of evidence-based decision-making in medicine, such as treatment options based on lab data, to driver-less cars. She explained that Google’s tests with self-driven cars provide evidence that they are far safer than cars driven by the typical human driver. In fact, it should be noted that the Google car’s only reported collision to date was the fault of a human driver who rear-ended the car driven by the Google computer. However, people may be much more accepting of errors made by humans (from whom errors are expected) than by machines and may be willing to accept a more error-prone human performing a task over a less-error prone machine. Similarly, for laboratory clinical decision support, an algorithm providing clinical advice (or even an automated diagnosis or treatment) may be held to a much higher standard than would a person providing the same function. A consequence of this philosophy is that health care systems may choose not to implement decision support systems that fail to reach a standard of perfection for every patient, even when such algorithms could clearly improve diagnostic quality or safety for the population as a whole.

Dr. Kim also argued that laboratories should move beyond just presenting test results alone, as there are many instances where the lab data can be prone to

over-interpretation, potentially leading to inappropriate treatment or adverse outcomes. For example, studies have shown some physicians order troponins on patients indiscriminately and then admit patients with low-level troponin elevations, even when there is no clinical basis to suspect cardiac ischemia. In some of these cases, the clinicians may be failing to interpret the troponins in a Bayesian context; given a very low pretest likelihood of cardiac ischemia, even a positive troponin result may likely be a false positive. Dr. Kim implies that laboratories could help improve this situation by taking a greater role in ensuring that troponins be ordered only when clinically indicated and that they be reported in a way that incorporates clinical context. Troponins are just one of many examples fitting this paradigm.

Another point made by Dr. Kim related to how laboratories have large quantities of structured data at their disposal; mining this data has the potential to better identify patterns that may inform and refine appropriate treatment. Finally, Dr. Kim argues that given the laboratory’s key role in diagnosing a wide-range of diseases, the laboratory needs to develop partnerships through the health care system and particularly with patients themselves. This may include expanding the provision of actionable information directly to patients.

Walter Henricks, MD

Walter Henricks, Medical Director of the Center for Pathology Informatics and Staff Pathologist at the Pathology and Laboratory Medicine Institute of the Cleveland Clinic, offered a number of practical considerations and decision support opportunities in his talk, “Pathology’s Role in Implementation of Laboratory Test Order Management in the EHR.” Many of the opportunities he described are not only possible in the present, but might also facilitate some of the longer term goals described in some of the other talks.

Dr. Henricks began by noting that many of the informatics and decision support challenges we face are “more than IT” and really involve “people” and “processes” as well. He next noted some of the strategies he and colleagues employ at the Cleveland Clinic within their Epic EHR and its CPOE capabilities. Among these are “hard stops” implemented in the CPOE systems to prevent clinicians from placing duplicate orders on the same day for certain laboratory tests; overriding these hard stops requires placing a call to the laboratory. “Soft stops” are a utilization tool related to hard stops, in which the ordering provider has the opportunity to override the alert and proceed with ordering the test. Because of initially uneven adoption of CPOE and the availability of process workarounds at community hospitals in its health system, Cleveland Clinic deployed soft stops as a lower risk and a more politically viable first step in such mixed-provider environments.

Cleveland Clinic has also implemented decision support rules that restrict the ordering of certain complex and typically expensive genetic and genomic tests to clinicians who are “deemed users.” Deemed users mostly consist of specialists in the disorder(s) for which the test is appropriately used. Non-deemed users must get approval from the laboratory, medical genetics, or a deemed user prior to ordering. For requests referred to the laboratory, a laboratory-based genetics counselor assists in the review and a molecular genetic pathologist is the approver. In addition, the Cleveland Clinic further restricts use of these tests specifically on inpatients by allowing the test only if recommended by a medical genetics consultation. The CPOE system also displays guidance for pharmacogenomic testing and relevant previous pharmacogenomic results to providers when they order certain medications.

Because “people” and “governance” are key considerations in implementing these types of decision support, Dr. Henricks noted the importance of having a test utilization committee to determine the decision support rules and alert criteria used in the CPOE system. A pathologist chairs this committee, which has “multidisciplinary” membership including members of other clinical departments (all department heads are invited) and IT leadership. The committee provides recommendations to hospital leadership, including the Chief Executive Officer, Chief of Staff and Chief of Medical Operations for support and approval.

Anand Dighe, MD, PhD

Anand Dighe, Associate Pathologist and Director of the Core Laboratory in the Department of Pathology at the MGH and Associate Professor of Pathology at Harvard Medical School, delivered the final talk of the symposium, “Barriers and Opportunities for Pathologist-Driven Decision Support.” In it he argued that expectations for clinical laboratories are too low; the current bar which simply requires laboratories to report basic observations and data needs to be raised. Laboratories should increasingly be looked upon help guide test selection and provide diagnostic information, rather than simple observations. While meeting these higher expectations could drastically enhance the value of laboratory testing, there are numerous barriers to doing so. Given the sheer volume of laboratory tests, manual pathologist involvement must be limited to select cases; in most instances, electronic decision support systems need to be developed and deployed to provide this enhanced diagnostic value.

Dr. Dighe specifically described six barriers along with possible solutions. The first of these barriers was “Someone else is already doing it.” By this he meant that at many institutions, the CPOE and other electronic systems are controlled by non-pathologists; however,

pathologists need to be involved in or ideally control the clinical content of the CPOE system for laboratory tests to enable them to manage test utilization and appropriate test selection. A related barrier is “we (pathologists) don’t have the tools to do the job.” With this second barrier, Dr. Dighe was pointing out that even when pathologists want to be involved in the content of a CPOE system and have the administrative authority to do so, their ability may be technically limited by a need to request IT resources to make CPOE updates. One solution is for pathologists to develop middleware applications, such as the MGH Path Connect System,^[30] that allow them to update the clinical content of the CPOE systems without involvement of IT staff.

The third barrier discussed is that most laboratories and health systems now use primarily commercially developed systems that often leave limited opportunities for customization. Nonetheless, Dr. Dighe argued that many of these systems offer “untapped capabilities” meaning that some apparent limitations may be in a health system’s understanding or application of its systems and not in the systems themselves. In addition, when a new system is being selected or deployed, Dr. Dighe argued that it is imperative that pathologists be involved early and throughout the process. The fourth barrier Dr. Dighe mentioned is “Why can’t our providers just use the system properly?” The key here is that laboratories should not assume that clinicians will properly use whatever systems they are initially given. It is important to collect data on how providers interact with the system to make adjustments and provide additional user education as indicated.

The fifth barrier is “I can’t get useful reports out of the system.” By this, Dr. Dighe is cautioning against being reliant on central hospital or enterprise IT services to provide a report every time important data is needed from the LIS or CPOE systems. It is important for pathologists to have a way to access this data in real time to make operational improvements and optimally fulfill their responsibilities. Having to wait in queues (potentially months long) to get needed data can severely impair laboratory operations. One solution is for pathologists to create data marts that store mirrors of data in the LIS and CPOE systems that pathologists can query as needed to generate key reports and metrics.

The sixth barrier is the historic perception in some laboratories that “Once the result is sent to the EHR, I’m done.” In contrast to this view, Dr. Dighe argued that laboratories could greatly enhance value by taking steps to ensure that “actionable” laboratory results are properly acted upon. “Dropped balls” (laboratory results not properly follow-up or acted upon) are of particular concern with information that is time-sensitive, but is not critical, since critical results are directly called

to a responsible person. Results viewers that require responsible clinicians to acknowledge actionable results with processes to follow-up if acknowledgment does not occur within a specified period may be useful. Laboratories need to consider processes and strategies to monitor and avoid result communication and action failures.

DISCUSSION

A number of recurrent themes emerged from the symposium. Perhaps the most pervasive of these related to the dichotomy between data and diagnostic information^[2] with regard to laboratory testing. In particular, laboratories currently produce primarily observational data and leave it to clinicians to integrate these data and convert data into true diagnostic information. However, the human brain is not well-equipped to interpret multidimensional data,^[17,18] and laboratory and clinical data typically exist as large numbers of discrete observations. This highlights the need for computational systems to convert laboratory data to information, effectively reducing the dimensionality of the information that clinicians must process.

Likewise, there are differences between human talents and computer abilities; well-designed symbioses between humans and computers may ultimately optimize diagnosis. Furthermore, unsupervised statistical analysis of data alone may not be sufficient and we may instead need to constrain our analysis through use of mechanistic physiologic models. We will also need to incorporate human intuition and domain expertise into our models.

Another facet that emerged related to the often complex relationship between data and the information that can be derived from it.^[26,27] For example, MAAs use large sets of laboratory values to provide key information that raw data alone could not. A corollary to this idea is that much of the laboratory data currently produced is not fully converted to information. A key strategy to improving laboratory diagnosis is to better utilize available data. However, a challenge will involve identifying patterns in data without overfitting, or in other words, clearly discriminating “signal from noise”. Another major theme from the symposium involved the unique challenges posed by emerging diagnostic technologies and in particular, next generation sequencing. In genomics, the separation between data and information is far greater than in most traditional areas of diagnostic testing. While with basic laboratory testing, considerable information can be manually extracted from the relatively unprocessed data, this is impossible within the context of next generation sequencing. Raw sequence reads are useless in diagnosis without further computational processing, alignment, base calling and variant-calling with additional and potentially extensive clinical and

familial correlation for variants that are not already well defined. Furthermore, interpretation of next generation sequencing data requires a computer-usable knowledge base derived from the literature and prior experience and systems to apply this knowledge to help automate interpretation. Without this, manual review of identified variants would be cost prohibitive. Other challenges with regard to genomics include the consideration that much of the data generated is of yet to be determined significance and that the knowledge base is in constant flux. As with other areas, guiding clinicians as to the appropriate action to take in response to genomic results is critical.

The utility of automating various aspects of information processing was another recurrent theme. For example, a computational infrastructure may eventually help transform the fields of pathology and laboratory medicine into central diagnostic specialties, providing integrative reports that clearly define precise diagnoses with clear therapeutic information in many cases. Likewise, laboratory data could potentially be used to provide overarching objective evaluations of pathologist and trainee performance as well as a comprehensive understanding of the value and utility of various health care services.

A final theme that emerged related to how many of the barriers to bringing decision support toward reality are really personnel, political, infrastructural and administrative challenges rather than technological limitations. For example, training and educating pathologists is a key measure in driving this field forward. Likewise, while superficially a technological consideration, a key component to developing the “professional LIS” may lay in building co-development relationships between pathologists and LIS vendors. Moreover, the functionality of information systems may be limited by the knowledge base of the people using them. Likewise, strong governance and organizational structures are keys to the success of decision support efforts at large medical centers.

Both patients and clinicians stand to benefit from the greater diagnostic precision, optimized test utilization, improved efficiency and reduced costs that may stem from enhanced pathology decision support. For example, pathology decision support could benefit patients through more rapid or precise diagnosis, allowing them to receive treatments that are more “personalized,” timely and clinically optimal. This should in turn improve outcomes, reduce side-effects from ineffective treatments and reduce patients’ cost of care. Likewise, physicians may benefit through improved efficiency, reduced risk of error and enhanced diagnostic capacity. The improved efficiency may also provide clinicians with the opportunity to spend more time interacting with patients and performing other rewarding clinical activities.

However, the potential benefits of pathology decision support are not limited to direct patient care. Rather, this emerging field may also offer new research opportunities and improve the economics of pathology services. For example, expanded decision support should generate novel clinical research questions. In particular, clinical research will be needed to investigate the optimal strategies for management of clinical conditions in the setting of increased diagnostic precision. Likewise, health systems research will be needed to study the effects of the decision support systems themselves. Similarly, the improved ability to convert raw genomic data into clinically actionable information will facilitate translational genomic research. Basic science researchers may be called upon to investigate the biologic or physiologic mechanisms underlying newly identified patterns within clinical data. Finally, by enhancing the value of laboratory services and helping to demonstrate the contribution that laboratory testing provides to patient care, electronic decision support systems could improve the economics of pathology. Indeed, as accountable care organizations expand and provide health systems with a single limited pool of resources,^[16] it may be particularly incumbent upon pathologists to demonstrate the value of their services.

Overall, the symposium illustrated that, while pathology decision support is a field still in its infancy, it is a field with tremendous opportunity and potential that is moving quickly. There is a clear need for inter-institutional collaboration to solve the technical, infrastructural and data acquisition challenges and to support LIS and EHR vendors in developing systems that can support emerging decision support strategies. We look forward to holding a follow-up symposium on this topic to discuss interim progress.

REFERENCES

- Baron JM, Dighe AS. The role of informatics and decision support in utilization management. *Clin Chim Acta* 2014;427:196-201.
- Shortliffe EH, Cimino JJ. *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. 3rd ed. New York: Springer; 2006.
- Greenes RA, ebrary Inc. In: *Clinical Decision Support the Road Ahead*. Amsterdam, Boston: Elsevier Academic Press; 2007.
- Singh H, Giardina TD, Meyer AN, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings. *JAMA Intern Med* 2013;173:418-25.
- Laposata M, Dighe A. "Pre-pre" and "post-post" analytical error: High-incidence patient safety hazards involving the clinical laboratory. *Clin Chem Lab Med* 2007;45:712-9.
- Plebani M, Laposata M, Lundberg GD. The brain-to-brain loop concept for laboratory testing 40 years after its introduction. *Am J Clin Pathol* 2011;136:829-33.
- Plebani M. The detection and prevention of errors in laboratory medicine. *Ann Clin Biochem* 2010;47:101-10.
- Jackson BR. Managing laboratory test use: Principles and tools. *Clin Lab Med* 2007;27:733-48, v.
- Van Cott EM. Laboratory test interpretations and algorithms in utilization management. *Clin Chim Acta* 2014;427:188-92.
- Gullapalli RR, Desai KV, Santana-Santos L, Kant JA, Becich MJ. Next generation sequencing in clinical medicine: Challenges and lessons for pathology and biomedical informatics. *J Pathol Inform* 2012;3:40.
- Bossuyt X, Verweire K, Blanckaert N. Laboratory medicine: Challenges and opportunities. *Clin Chem* 2007;53:1730-3.
- Blanckaert N. Clinical pathology services: Remapping our strategic itinerary. *Clin Chem Lab Med* 2010;48:919-25.
- Plebani M. The future of clinical laboratories: More testing or knowledge services? *Clin Chem Lab Med* 2005;43:893-6.
- Plebani M. Charting the course of medical laboratories in a changing environment. *Clin Chim Acta* 2002;319:87-100.
- Huck A, Lewandrowski K. Utilization management in the clinical laboratory: An introduction and overview of the literature. *Clin Chim Acta* 2014;427:111-7.
- Sussman I, Prystowsky MB. Pathology service line: A model for accountable care organizations at an academic medical center. *Hum Pathol* 2012;43:629-31.
- Kirwan JR, Chaput de Saintonge DM, Joyce CR, Currey HL. Clinical judgment in rheumatoid arthritis. II. Judging 'current disease activity' in clinical practice. *Ann Rheum Dis* 1983;42:648-51.
- Slovic P. In: Behavioral problems of adhering to a decision policy (unpublished manuscript) (1973). Cited through the Psychology of Intelligence Analysis by Heuer RJ Jr. Ch. 5. 1999. Available from: <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/PsychofIntelNew.pdf>. [Last accessed on 2013 Oct].
- Higgins JM, Mahadevan L. Physiological and pathological population dynamics of circulating human red blood cells. *Proc Natl Acad Sci U S A* 2010;107:20587-92.
- Brenner S. In theory. *Curr Biol* 1997;7:202.
- Shirts BH, Wilson AR, Jackson BR. Partitioning reference intervals by use of genetic information. *Clin Chem* 2011;57:475-81.
- Carson JL, Grossman BJ, Kleinman S, Tinmouth AT, Marques MB, Fung MK, et al. Red blood cell transfusion: A clinical practice guideline from the AABB*. *Ann Intern Med* 2012;157:49-58.
- Society of Thoracic Surgeons Blood Conservation Guideline Task Force, Ferraris VA, Ferraris SP, Saha SP, Hessel EA 2nd, Haan CK, et al. Perioperative blood transfusion and blood conservation in cardiac surgery: The Society of Thoracic Surgeons and The Society of Cardiovascular Anesthesiologists clinical practice guideline. *Ann Thorac Surg* 2007;83:527-86.
- Napolitano LM, Kurek S, Luchette FA, Corwin HL, Barie PS, Tisherman SA, et al. Clinical practice guideline: Red blood cell transfusion in adult trauma and critical care. *Crit Care Med* 2009;37:3124-57.
- Chou R, Wasson N. Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: A systematic review. *Ann Intern Med* 2013;158:807-20.
- Waljee AK, Joyce JC, Wang S, Saxena A, Hart M, Zhu J, et al. Algorithms outperform metabolite tests in predicting response of patients with inflammatory bowel disease to thiopurines. *Clin Gastroenterol Hepatol* 2010;8:143-50.
- Baron JM, Mermel CH, Lewandrowski KB, Dighe AS. Detection of preanalytic laboratory testing errors using a statistically guided protocol. *Am J Clin Pathol* 2012;138:406-13.
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009;457:1012-4.
- Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaut R. The landscape of inappropriate laboratory testing: A 15-year meta-analysis. *PLoS One* 2013;8:e78962.
- Grisson R, Kim JY, Brodsky V, Kamis IK, Singh B, Belkiz SM, et al. A novel class of laboratory middleware. Promoting information flow and improving computerized provider order entry. *Am J Clin Pathol* 2010;133:860-9.