Original Article

Pathology informatics fellowship training: Focus on molecular pathology

Diana Mandelker, Roy E. Lee[1], Mia Y. Platt[2], Gregory Riedlinger[2], Andrew Quinn, Luigi. K. F. Rao[2], Veronica E. Klepeis[2], Michael Mahowald[2], William J. Lane, Bruce A. Beckwith[3], Jason M. Baron[2], David S. McClintock[4], Frank C. Kuo, Matthew S. Lebo, John R. Gilbertson[2]

Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA, [1]Robert J. Tomsich Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH 44195, [2]Pathology, Massachusetts General Hospital, Boston, MA 02114, [3]Pathology, North Shore Medical Center Salem Hospital, Salem, MA 01970, [4]Pathology, University of Chicago Medicine, Chicago, IL 60637, USA

E-mail: * John R. Gilbertson - jrglibertson@partners.org
*Corresponding author

Abstract

Background: Pathology informatics is both emerging as a distinct subspecialty and simultaneously becoming deeply integrated within the breadth of pathology practice. As specialists, pathology informaticians need a broad skill set, including aptitude with information fundamentals, information systems, workflow and process, and governance and management. Currently, many of those seeking training in pathology informatics additionally choose training in a second subspecialty. Combining pathology informatics training with molecular pathology is a natural extension, as molecular pathology is a subspecialty with high potential for application of modern biomedical informatics techniques. **Methods and Results:** Pathology informatics and molecular pathology fellows and faculty evaluated the current fellowship program's core curriculum topics and subtopics for relevance to molecular pathology. By focusing on the overlap between the two disciplines, a structured curriculum consisting of didactics, operational rotations, and research projects was developed for those fellows interested in both pathology informatics and molecular pathology. **Conclusions:** The scope of molecular diagnostics is expanding dramatically as technology advances and our understanding of disease extends to the genetic level. Here, we highlight many of the informatics challenges facing molecular pathology today, and outline specific informatics principles necessary for the training of future molecular pathologists.

Key words: Clinical informatics, informatics fellowship training, molecular pathology informatics, molecular pathology training, molecular pathology, pathology informatics fellowship, pathology informatics training, pathology informatics

BACKGROUND

As a medical specialty, pathology generates and interprets laboratory data on fluid and tissue specimens. The data generated and interpretations rendered must be accurate, reproducible, and presented in a clearly understandable format, as clinical decisions will be made based on this information. To manage an ever-increasing volume and complexity of data, informatics solutions have been sought to more effectively analyze, track,

integrate, and communicate this important clinical information.

Pathology informatics focuses on pathology information, analysis tools, and processes. Due to the recent exponential growth in medical data, discoveries, and diagnostic technologies; specialized skillsets have become required to manage this information. Some pathology informatics applications include the storage of intralaboratory data, test triage and utilization, electronic communication of test orders and results between locations, digital image libraries, search engines for biorepositories, and bioinformatics, which itself is a subset discipline of biomedical informatics that seeks to analyze large biologic datasets and develop computational algorithms. Despite the view of informatics as a distinct subspecialty, its roots are embedded within all pathology disciplines, as the entire specialty seeks to optimize its ability to manage information.

Molecular pathology is one such subspecialty in which informatics is essential. Here, pathologists utilize nucleic acid-based techniques and clinical correlation to diagnose, determine prognosis, predict response to therapy, and manage family care and decisions. The rapid growth of molecular pathology has left the field vulnerable to potential errors in communication, disorganized data storage, and inefficient workflows. Moreover, molecular pathology laboratories must perform comprehensive validation and quality assurance/quality control (QA/QC) of both the "wet-lab" laboratory developed tests and the "dry lab" bioinformatics pipelines. Despite these challenges, in many molecular laboratories, the steps of accessioning, intralaboratory workflow, and interpretation and resulting remain a largely paper-driven (or Excel spreadsheet-driven) process. With the high complexity of both the testing performed and the data generated, these manual error-prone processes must be addressed by informatics solutions across the entire "testing cycle" for molecular diagnostics.

Current molecular assay techniques such as quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) and fragment analysis by capillary electrophoresis require software and computational algorithms to accurately determine the results of these technical assays. The ongoing clinical implementation of next generation sequencing (NGS) will result in dramatic increases of data generated by molecular pathology laboratories. Current manual curating methods for novel genetic variants are extremely time-consuming, cannot scale with the increasing volume and complexity of cases, and will require automated and intelligent computational methods to streamline variant analysis. The speed of molecular innovation and the need to integrate genetic data with other pathology and clinical data pose informatics challenges unique to the discipline. Furthermore, in light of recent federal mandates and upcoming changes in reimbursements, molecular pathologists will need to leverage informatics to adapt the testing environment of a molecular laboratory to these new regulations.

In this paper, we will discuss informatics training in molecular pathology based on our Clinical Fellowship in Pathology Informatics experience over the past 5 years. Key aspects of the program include.

The History and Design of Our Fellowship Program

Our Pathology Informatics Clinical Fellowship Program is based in a large diversified healthcare system and operates in two large academic medical centers and a community hospital with a strong outreach program. It has approximately 20 active faculty members, has graduated 12 fellows since 2009, and has seven active clinical fellows. The details of the program have been published.[1]

The Educational Structure

The fellowship has both a customizable arm and a required arm. The customizable arm (which represents approximately two-thirds of a fellow's time) involves rotations, mentorships, research projects, and didactic activities aligned to the fellow's interests and career goals. The required arm involves largely didactic activities across the entire breath of pathology informatics domains. The goal of the required arm is to expose the fellow to all major components of informatics, while the customized area allows deep specialization.

The Core Didactic Course

The foundation of the required arm of the fellowship is a weekly session of readings, lectures, and faculty-led discussions termed the "core didactic". The core didactic attempts to present informatics as the study and management of information, information systems, processes, and governance. The details of the core didactic course and other components of the educational structure have been published previously.[1,2]

The Career Paths of Our Fellows and the Fellowship Tracks

Initially the fellowship was designed as a 1 or 2-year experience for pathologists interested in becoming Director of Pathology Informatics in a pathology department or healthcare system. Very quickly, however, it became clear that there was another, larger group of pathologists with a different career goal. These trainees were interested in becoming a director other pathology laboratories, such as surgical pathology, microbiology, clinical chemistry, etc., and were interested in learning how to apply informatics principles to those more traditional disciplines. Today, this type of trainee represents over two-thirds of our fellows.

To support this demand, we have added 'tracks' into the fellowship structure. The two main tracks are a 2 year "Director of Pathology Informatics" track (for those

interested in becoming directors) and a "1 + 1" track in which a trainee does two coordinated fellowships: 1 year of clinical informatics and a 1- or 2-year fellowship in another pathology domain.[2]

The Molecular Pathology/Pathology Informatics Combination

Of the 17 fellows graduated or active in our program, five have done or are currently doing fellowships in both informatics and molecular pathology, making molecular pathology the most popular "1 + 1" combination. Those five pathologists/trainees are authors on this paper (DM, RL, MP, GR, and WL).

For teaching purposes, our fellowship program divides pathology informatics into four main divisions: Information Fundamentals, Information Systems, Workflow and Process, and Governance and Management. Within this framework, we will describe some of the pathology informatics challenges that are important to molecular pathology, highlight informatics disciplines of particular relevance to molecular pathology [Tables 1-4], propose possible rotations and/or research projects to provide informatics fellows with vital practical experience in molecular pathology [Table 5], and outline a suggested curriculum for training pathology informatics fellows with a career interest in molecular pathology.

METHODS

The Partners Pathology Informatics fellowship has the unique advantage of having five former or current "1 + 1" fellows in both Pathology Informatics and Molecular Pathology. All five of these pathologists (DM, RL, MP, GR, and WL) worked to evaluate the current program's core curriculum topics and subtopics for relevance to molecular pathology. Topics pertinent to the molecular pathologist, but currently not well covered by the curriculum were also identified. Focusing on this overlap between pathology informatics and molecular pathology, we developed a draft curriculum for those with interest in both fields. Informatics and molecular pathology faculty with diverse backgrounds at Brigham and Women's Hospital, Massachusetts General Hospital, and North Shore Medical Center then reviewed and edited the proposed curriculum.

RESULTS

Information Fundamentals

Information Fundamentals include the subdomains of Information Architecture, Information Quality, Information Manipulation, Cognition and Human-Computer Interaction, Software Design Principles, and Special Domains [Table 1]. Importantly, Information Fundamentals focuses mostly on the underlying concepts;

for example, Information Manipulation focuses mostly on algorithms and algorithm design rather than actual programming. Of particular significance to a molecular pathologist is a fundamental understanding of database design, information quality and manipulation, and algorithm basics. Some of the fundamentals that need to be the focus of informatics training for molecular pathologists are highlighted by specific challenges facing molecular diagnostics described below.

A basic understanding and appreciation of computational algorithms is becoming increasingly critical for a practicing molecular pathologist. These techniques are currently used in a number of contexts, ranging from molecular test utilization to the management and analysis of large datasets. For instance, with the sheer volume of data generated by NGS technology, we rely on computationally intensive bioinformatics algorithms to apply quality scores and filters to the sequence data, align individual sequences to a reference genome, and identify variants.[3] All these processes occur before the molecular pathologist gets a list of variants to interpret for a given patient. Understanding the limits of the technology and associated informatics will inform the pathologist about

Table 1: Information fundamentals

Topic	Subtopic
1. Healthcare Informatics: History and Concepts	1. Medical and Pathology Informatics History and Foundations
	2. Fundamentals of Healthcare Information
2. Information Architecture	**1. Data Architecture and Modeling**
	2. Metadata in Healthcare Information
	3. Database Design and Architecture
	4. XML, the Semantic Web, and Ontologies
	5. Healthcare Messaging Models
	6. Healthcare Content Vocabularies
3. Information Quality	1. Information Theory
	2. Information Quality Principles
	3. Healthcare Information Quality
4. Information Manipulation	**1. Data Manipulation**
	2. Programming Principles
	3. Information Retrieval
	4. Data Analysis Principles
	5. Decision Support Principles
5. Cognition and Human Computer Interaction	1. Cognition
	2. Human Computer Interaction
	3. User Interfaces
	4. Information Display
6. Design Principles	1. Universal Design Principles
	2. Software Engineering and Development
7. Special Information Domains	1. Image Information Principles
	2. Molecular and Genomic Information Principles (Bioinformatics)
	3. Textual Information Principles
	4. Cancer Registries and Public Health Information Principles

*Subtopics of particular relevance to molecular pathology are illustrated in bold type

Table 2: Information systems

Topic	Subtopic
1. Infrastructure Fundamentals	1. Hardware
	2. Software
	3. Data Storage Principles
	4. Networking and the Internet
	5. Security
2. Laboratory Information Systems	**1. LIS General Concepts**
	2. Anatomic Pathology LIS
	3. Clinical Pathology LIS
	4. Transfusion Medicine LIS
	5. Specimen Identification Systems
	6. Point-of-care Testing
	7. Outreach Systems
3. Interfaces	**1. Instrument Interfaces**
	2. LIS to HIS Interfaces
	3. Middleware
4. System Life-cycle	1. System Needs Analysis and Selection
	2. System Implementation
	3. System Validation Practices
	4. System Maintenance and Disaster Planning
	5. System Retirement and Data Conversion
5. Health Information Systems	1. Healthcare Information Systems Overview
	2. Electronic Medical Record
	3. Patient Registration and Identification Systems
	4. Computerized Provider Order Entry
	5. Results Reporting Principles
	6. Billing and Coding Systems
	7. Enterprise Models and Health Information Exchange
6. Imaging Systems	1. Picture Archival and Communications Systems
	2. Pathology Imaging Systems

*Subtopics of particular relevance to molecular pathology are illustrated in bold type.
LIS: Laboratory information systems, HIS: Hospital information system

Table 3: Workflow and process

Topic	Subtopic
1. Process and Quality Improvement	**1. Process Improvement Methodologies**
	2. Software
	3. Data Storage Principles
2. Process Management	1. Principles of Process Management
	2. Fundamentals of Process Design, Mapping, and Modeling
	3. Process Modeling and Analysis
3. Workflow Analysis Methods	**1. Process Redesign/Reengineering**
	2. Workflow Redesign/Reengineering in Pathology
	3. HL7 v3 RIM and Constrained Information Models
	4. Modeling Data Flow in Health Systems
4. Automation	1. Principles of Automation
	2. Automation in the Clinical Laboratories
5. Decision Support in Pathology	**1. Business Intelligence and Decision Support**
	2. Decision Support Systems and Analysis in Pathology
6. Special Pathology Process Domains	1. Digital Pathology Workflow
	2. Molecular/Genomics Workflow

*Subtopics of particular relevance to molecular pathology are illustrated in bold type

Table 4: Governance and management

Topic	Subtopic
1. Leadership	**1. Leadership Principles, Models, and Practices**
	2. Organizational Change and Change Management
	3. Effective Communication Practices
	4. Role of the Pathology Informatician
	5. Career Planning in Pathology Informatics
2. Management	**1. Governance of Information Technology Services**
	2. Project Management-General Principles
	3. Project Management in Healthcare
	4. Healthcare Finance Principles
	5. Capital Finance Principles
	6. Operations Finance Principles
	7. Human Resource Practices
	8. Staff Hiring, Training, and Evaluation
3. Regulation	**1. Clinical Laboratory Improvement Amendments**
	2. Health Insurance Portability and Accountability Act
	3. Accreditation of Clinical Laboratories and Hospitals
	4. Transfusion Medicine Regulations
	5. Ethics and Legal Issues
	6. Informatics in Clinical and Translational Research-Principles
	7. Informatics in Clinical and Translational Research-Practice
	8. Current Topics in Healthcare Reform

*Subtopics of particular relevance to molecular pathology are illustrated in bold type

the accuracy of the variant calls and the limitations of the clinical test (i.e. capacity to detect insertions/deletions or translocations and difficulty with genes with high homology) and translate into a higher quality, actionable pathology report for the clinician.

A predicament facing genetic analysis is the variable quality of the currently available genetic/genomic databases. Large databases such as Human Gene Mutation Database (HGMD), Online Mendelian Inheritance in Man (OMIM), and Catalogue of Somatic Mutations in Cancer (COSMIC) attempt to establish disease-gene associations in a categorical fashion. As the variants found in a patient's sample are cross-referenced against these databases, the quality, and breadth of the database entry and metadata are vital to a molecular pathologist's interpretation of their given case. However, there is currently a lack of standards in information content for genetic databases and the existing data is only as good as the original studies and any subsequent curating. Often, the underlying research data may be poor or inadequate and not interpreted to a high

Table 5: Suggested curriculum areas and fellow rotations for fellowship training in pathology informatics and molecular pathology

Curriculum topic	Subtopic	Examples of fellow rotations/projects
Information Fundamentals		
Information Architecture	Data Architecture and Modeling	Validation of predictive algorithms for sequence variants
	Metadata in healthcare information	Classification of sequence variants and associated supporting evidence
	Database design and architecture	Developing or improving a lab's gene and variant curation database
Information Quality	Information quality principles	Participation in working groups that develop standards for gene and variant annotations
Information Manipulation	Information retrieval	Automating information gathering for novel variant assessment
	Data analysis principles	Understanding metrics used to evaluate analytical performance of sequencing runs and the statistics and limitations of variant calling
	Decision support principles	Automating methods for identifying inappropriate molecular test ordering
Information Systems		
Laboratory Information Systems	LIS general concepts	Designing an automated method for identifying inappropriate molecular test ordering
	Specimen identification systems	Create barcode system for sample accessioning and tracking through molecular workflow
	Outreach systems	Work with information technology department to develop LIS-HIS interface at remote sites ordering molecular testing
Interfaces	Instrument interfaces	Develop/validate interface of molecular data from instruments to LIS
	LIS to hospital information system interfaces	Work with IT department to develop bidirectional LIS-HIS interface to order tests and report results
	Middleware	Identify, implement, and/or validate appropriate middleware that may facilitate interfaces between instruments, LIS, and HIS
Health Information Systems	Electronic health record	Ensure proper display of molecular reports in structured format in EHR
	Computerized provider order entry	Work with clinicians to optimize online ordering for molecular testing
	Results reporting principles	Development of enhanced reporting templates (i.e., pathology reports with graphical elements)
Workflow and Process		
Process management and quality improvement	Process improvement methodologies	Redesign of sample accessioning and tracking, material generation, and testing to improve molecular workflow
	Software	Implement software to improve molecular workflow
	Data storage principles	Participate in the development of institutional polices for data storage, use, and sharing
Workflow analysis methods	Process redesign/reengineering	Process improvements such as introducing electronic sign-out
	Workflow redesign/reengineering	Workflow analysis for specific test workflows (e.g., bottlenecks)
	Modeling data flow in health systems	Investigation of metrics and digital dashboards useful to a molecular pathologist (i.e., displaying turnaround time for different tests)
Pathology decision support	Business intelligence and decision support	Introduce commercial intelligence solutions into molecular laboratory
	Decision support systems and analysis	Use analytic approaches to design optimal genetic test ordering and reflex protocols
Governance and Management		
Leadership	Leadership principles, models, and practice	Participate in the molecular laboratory's administrative meetings
	Effective communication practices	Examination of clinician needs and satisfaction with the molecular lab's services

Table 5: Contd...

Curriculum topic	Subtopic	Examples of fellow rotations/projects
	Role of the Pathology Informatician	Apply pathology informatics fundamentals to improve molecular laboratory performance
Management	Governance of information technology services	Facilitate interaction between IT and molecular groups to ensure proper molecular IT development
	Project management	Involvement in practical decision making process for introducing a new assay/technology
	Finance principles	Implement business intelligence software into the molecular lab aimed at reducing costs and maximizing profitability
Regulation	Accreditation of clinical laboratories and hospitals	Participation in College of American Pathologist proficiency testing and/or CAP/The Joint Commission external and internal inspections
	Ethical and legal issues	Work with legal department to develop practical informed consent documents for genomic testing
	Informatics in clinical and translational research	Develop policies for use of sequence data generated in research settings

LIS: Laboratory information systems, HIS: Hospital information system, IT: Information technology, EHR/EMR: Electronic health record, CAP: College of American pathologist

clinical standard, though large consortium efforts such as the ClinVar database are being developed to address this issue.[4] In addition to public databases, many molecular labs will have local databases, including those containing sequence variants and annotations or images of fluorescence *in situ* hybridization FISH) cases and tumor slides used for molecular testing. A molecular pathologist must be able to critically analyze the quality of the archived data before applying it to his or her case interpretation. Moreover, this interpretation would be assisted tremendously by an informatics solution that can gather all relevant data from published literature, genetic databases, protein computational analysis programs, etc., and present it to the molecular pathologist in a format conducive to case sign-out and subsequent integrated reporting.

Sample archiving in molecular pathology is becoming more important as biobanking of deoxyribonucleic acid (DNA) is becoming increasingly commonplace at many academic medical centers for future clinical and research uses. These biorepositories require a strong understanding of specimen inventory control, consent tracking database design, and basic information retrieval concepts (such as ontologies and natural language processing). Information retrieval from these biobanks is complicated by a lack of consensus medical terminology, a problem that affects all healthcare informatics. For example, a query for the terms "colon cancer," "colorectal carcinoma," and "bowel cancer," all must be reconciled and return the same search results. The uniform use of medical ontologies such as SNOMED-CT or the National Library of Medicine's Unified Medical Language System (UMLS) is therefore essential for applications where the quality of the search engine is crucial.

For training purposes, some examples of projects that fall under the "Information Fundamentals" umbrella [Table 5]

may include work with laboratory database developers or administrators. For a fellow, understanding the structure of a molecular lab's database of genetic information, the metadata captured, and the ability to query, is invaluable.

Information Systems

Our fellowship divides Information Systems into Infrastructure, Laboratory Information Systems (LISs), Interfaces, System Life Cycle, Health Information systems (Electronic Health Records (EHRs), Billing, Admission Transfers Discharges Systems (ADT), etc.), and Imaging Information Systems [Table 2]. These information systems are interwoven into major healthcare systems today, with pathologists utilizing both the LIS as well as the EHRs. Currently in the vendor market, there is a shortage of LIS and EHR systems with the ability to appropriately track and store genetic data in a satisfactory manner for laboratory professionals. A fundamental understanding of the LIS and the EHR, the interfaces between them, and their dynamic capabilities will help molecular pathologists manage the unique aspects of the data generated by the molecular laboratory.

One pressing issue facing molecular pathology is the relationship between genomic data and the patient's EHR.[5] Germline genomic data is unique in that it is a static block of information that needs to be dynamically accessed and bioinformatically reanalyzed as clinical scenarios change and genetic data evolves. For example, some models of the future of genetic testing predict that all newborn babies will have their genomes sequenced and stored, perhaps at an annual fee. In this scenario, data storage issues will become paramount. Should we store the raw sequencing data for all 3 billion base pairs of the genome, or just the single nucleotide, copy number, and structural variant calls (approximately 3-4 million variants)? Storing simply the variant calls

requires a confidence in today's reference sequences and aligners, while storing the raw data allows for realignment as knowledge and informatics evolve. The LIS/EHR could update the annotation of an individual's variants as more knowledge becomes available.

While germline assays need only be performed once in a patient's lifetime, microbiome, and cancer assays likely require acquisition of new data in the form of retesting with a more current specimen. For example, at the point of disease recurrence or metastasis, sequencing the emerging tumor clone would be preferable to relying on the sequence of the primary tumor. Tracking this data and being able to assess temporal changes becomes critical. In addition, any genetic data should not be stored in isolation, but must be in a system that supports integration for pharmacogenetic support, and integration with other laboratory data. Finally, security of genomic data is imperative, since unlike other laboratories test results such as a white blood cell count, a patient's genetic sequence contains inherent sample identity.

Two-way decision support that can guide both clinician test ordering and pathologist interpretation is a key component of today's LIS optimized for genetic testing. Molecular tests are expensive relative to traditional laboratory tests and inappropriate orders by clinicians can potentially cost a laboratory hundreds of thousands of dollars per year. For instance, one recent report suggests that approximately 30% of complex genetic tests are ordered in error,[6] adding unnecessary costs and delaying proper diagnosis. While many laboratories employ genetic counselors to review test orders, manual test order review is a time-consuming and expensive process. Furthermore, with the upcoming changes in Medicare reimbursement and federally mandated value-based healthcare on the horizon, cost containment will become critical to a molecular laboratory's financial sustainability. Many payors (insurance companies) have developed and implemented preapproval requirements for expensive genetic tests. These preapproval processes can be complicated and confusing, involving online questionnaires and embedded algorithms to triage orders and enable insurance company approval. Much of this confusion and uncertainty could be alleviated by implementing electronic decision support directly in the EHR for computerized provider order entry (CPOE). This embedded decision support would allow a clinician to enter a patient phenotype or condition, and then suggest an appropriate genetic test (i.e. single gene test vs gene panel vs exome vs genome). For the pathologist, once a test is technically completed, the patient phenotype could guide analysis by selecting candidate genes and variants in Mendelian disease cases, driver mutations in tumor cases, or relevant bacterial sequences in microbiome cases. A molecular pathologist with a keen understanding of their LIS could guide effective development of a system tailored for such a workflow, and assist clinical colleagues

in implementing test selection and algorithmic/tiered testing directly in the EHR. Such a system would optimize molecular test utilization and facilitate timely and accurate test interpretations.

Due to the rapid pace of medical discoveries and technological advances, new molecular assays are developed at a much faster rate than other pathology tests. This accelerated pace of test development magnifies several challenges faced by the LIS. For healthcare systems that have a nonintegrated LIS and EHR, adding new test definitions to the LIS that cross over to the EHR is a time-consuming process that could strongly benefit from workflow optimization. Moreover, for molecular reference labs, electronic test order codes that differ between hospital locations need to be consolidated. The processes surrounding the LIS must evolve to keep pace with an ever-changing molecular test menu. Although the difficulties stemming from the rapid growth of molecular tests and their subsequent test definitions in a LIS/EHR could be alleviated with integrated systems that contain both Anatomic and Clinical Pathology databases, many healthcare systems today rely on the "best of breed" approach, consisting of multiple, nonintegrated systems from different vendors. Additionally, different hospitals within the same healthcare system may use a different LIS/EHR system or run on separate installations/deployments of the same software, resulting in differing medical record number schemes and different test codes for the same test. This rapid growth in molecular tests poses a unique problem for healthcare systems.

For training considerations, experience with information systems can be obtained through operational rotations with the LIS team, where the fellow can gain practical experience with the capabilities and limitations of an institution's LIS and EHR. Specific projects will be influenced by an institution's information technology (IT) priorities, but one area where a pathology informatics fellow could lend key clinical expertise is building genetic testing decision support tools and reporting templates into the LIS. These test utilization projects are exceptionally timely given recent reports that many complex genetic tests are ordered in error,[6] and the consequences of genetic test misinterpretation by clinicians.[7]

Workflow and Process

Workflow and process theory as applied to pathology seeks to optimize the efficiency and effectiveness of a laboratory's practices and procedures. Subdivisions of this area of study include Process and Quality Improvement, Process Management, Workflow Analysis, Automation, and Decision Support [Table 3]. While workflow is critical to the functioning of any clinical laboratory, molecular pathologists are faced with challenges that make training in workflow design and engineering particularly important [Figure 1]. The test ordering process with molecular testing can be complicated as they
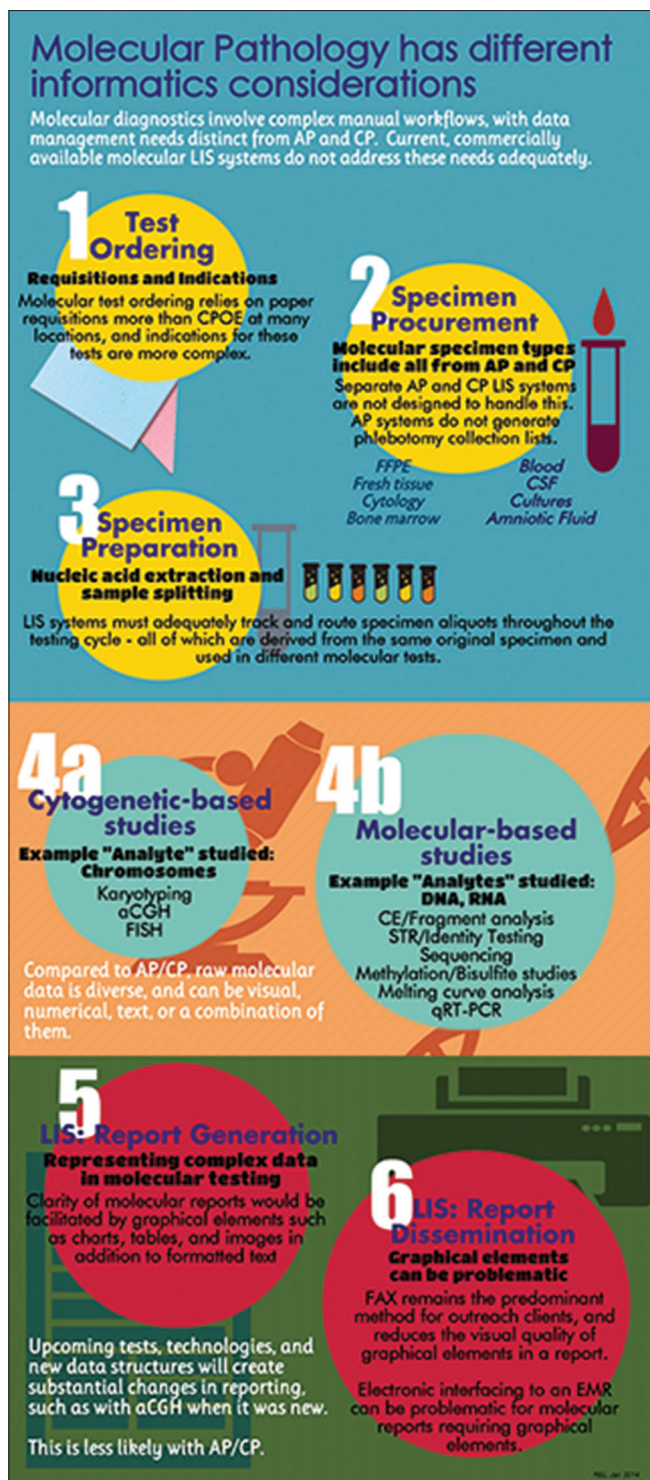
Figure 1: Informatics considerations for molecular pathology workflow. The testing cycle for a molecular diagnostic test is outlined with key informatics and workflow considerations highlighted for each step

Furthermore, a molecular laboratory must be able to collect, accession, and work with a variety of specimen types, including frozen tissue, paraffin-embedded tissue, cytology fine needle aspirate specimens, cerebral spinal fluid, bone marrow, blood, amniotic fluid, and other body fluids. Each specimen type may have different associated methodologies even for the same test (i.e. a PCR-based test requires a different DNA extraction protocol for formalin-fixed paraffin embedded tissue as compared to blood). Once a sample is in the lab, important to molecular pathology is the concept of unidirectional workflow. Given the sensitivity of nucleic acid amplification methods, samples are tracked through the laboratory with the designation of "pre" or "post" PCR to minimize contamination. Molecular specimen tracking is nontrivial because the system needs to differentiate between the original specimen and the specimen tested (post DNA extraction/purification), while also efficiently tracking shared specimens (e.g. clinical microbiology and molecular microbiology). Locating misplaced specimens, wastes personnel time and unnecessarily prolongs the testing cycle. Finally, workflows amongst molecular tests vary widely in turnaround times, ranging from hours/days (molecular microbiology) to weeks/months (array comparative genomic hybridization and sequencing panels). All test interpretations must then be reconciled to form a cohesive diagnosis.

For genomic testing involving direct sequence analysis, after the technical component is complete, hours to days of informatics processing is required. Including time for interpretation, the total testing cycle can take weeks to months to complete. Moreover, a single genomic test may require technical confirmation using an orthogonal technology (e.g. Sanger sequencing for sequence variants) or functional confirmation of a variant of unknown significance. Additionally, the data generated from NGS is often "incomplete", with poorly covered regions of the genome that cannot be evaluated unless the data gaps are filled in with an additional assay. All these steps add to the clinical sensitivity of genetic testing, but unfortunately also increase turnaround time. Furthermore, as molecular diagnostics evolves, the "total testing cycle" may actually be considered to be the lifetime of the patient, with reinterpretation being performed on a previously performed assay at appropriate intervals. With these variables, it is clear that formal workflow analysis and strong sample tracking mechanisms are critical to the optimal functioning of a molecular pathology lab.

Additionally, send-out testing poses a particular challenge for molecular pathology. Many new esoteric molecular tests are being developed and offered by private companies. At our institutions, many molecular send-out tests are ordered once by a clinician and may never be ordered again that year. There may be hundreds of these

may be placed by either the original ordering clinician or internally by the pathologist (add-on testing), and multiple tests may be ordered on the same sample both within molecular and in other pathology subspecialties.

individual tests ordered at an extremely low frequency. From the laboratory's perspective, there remains a need for optimal test ordering and utilization and the requirement to store results within the local LIS that may or may not have the particular send-out test defined. Regulatory requirements exist for handling results of send-out tests in the laboratory. Also, defining LIS test codes for these tests can improve information quality by storing the data as discrete elements as opposed to free-text elements, thus more effectively communicating with the EHR. Furthermore, the common practice of returning complex genetic data in paper form limits the capabilities of integrating this information with other systems. Although many of these issues are not unique to molecular pathology, they become exponentially more complicated in the realm of molecular testing and need to be addressed with better informatics solutions.

Formal workflow and process training may begin by working with a laboratory's director of operations. Ideally, a fellow could be involved in developing the workflow for a new assay being implemented in the lab, or a process improvement for a current suboptimal workflow, such as the introduction of electronic sign-out. Analysis of a particular test workflow to identify turnaround time bottlenecks is also a potentially valuable learning experience. Moreover, participation in College of American Pathologist (CAP) proficiency testing and/or laboratory inspections would be highly beneficial.

Governance and Management

Our fellowship's Governance and Management section focuses on the principles of leadership, management, and regulation. These skills are difficult to teach, yet particularly vital for a molecular pathologist who oversees a diverse clinical, technical, IT, and bioinformatics staff [Table 4]. All clinical laboratory systems must comply with regulatory requirements dictated by the Centers for Medicare and Medicaid Services (CMS), with regulatory inspections carried out by agencies such as the Joint Commission (TJC), the CAP, or state-run agencies. The regulatory environment for molecular laboratories is rapidly evolving and requires knowledge of ethical, legal, and regulatory statutes.

One pressing ethical and legal issue facing molecular pathology is the need to develop policies regarding consent and return of incidental findings during genomic testing. Incidental findings are genetic findings unrelated to the patient's testing indication. For example, a patient may undergo genomic testing for cardiomyopathy and be found to have a cancer susceptibility variant. This past March, the American College of Medical Genetics and Genomics (ACMG) released a list of 57 genes that should be interrogated for incidental findings.[8] However, policies regarding reporting these incidental findings vary between molecular laboratories, and pathologists and

geneticists must develop a cohesive strategy for both obtaining informed consent and reporting these findings. In addition, when developing an informed consent form, an institution's legal department and the molecular laboratory must be in agreement, and state laws vary regarding informed consent regarding genomic testing.

Unlike many clinical laboratories, molecular pathology tests are largely laboratory developed tests requiring diligent validation and QA/QC measures.[9] While many labs are accustomed to performing technical QA/QC, the bioinformatics pipeline must also be validated, versioned, and maintained. Likewise, formal documentation not only of the "wet-lab" validation, but also of the code underlying the informatics analysis, is mandatory. Bioinformatics proficiency testing is beginning to emerge and ever-changing algorithms may split the revalidation of technical components from the revalidation of wet-lab components. We may soon see proficiency testing samples in the form of emailed digital files instead of traditional vials filled with a biological analyte.

Bioinformaticians are increasingly becoming a key part of clinical molecular pathology labs. This presents an added management difficulty because most bioinformaticians are academically trained and less familiar with IT industry standards and practices, such as strict software versioning control platforms (such as Git and Mercurial) and documentation practices, both important in clinical production environments. Additionally, bioinformaticians possess a different skillset than traditional informatics analysts such as a LIS manager or a project manager. Therefore, even if a molecular pathology laboratory director is not formally trained in pathology informatics or bioinformatics, a basic understanding of the two disciplines will be helpful in guiding and managing specialized teams with diverse backgrounds.

Finally, Governance and Management training is best achieved through practical experience. Giving the fellow a position of responsibility is the ideal training method. For example, a fellow could be charged with overseeing all technical, informatics, and personnel components associated with a specific laboratory assay. We also continue to participate in focused case-based retreats using simulated business school style case studies to expose exposing trainees to real-life practical scenarios.[10]

DISCUSSION

Informatics is an increasingly important component of pathology practice and several programs, including our own, have developed fellowship training to support that need. In our experience, however, while some informatics fellows envision a career as a full time informatics specialist, the majority do a second fellowship in a diagnostic pathology specialty in addition to their clinical

informatics fellowship. The career goal of these fellows is to become a subspecialty pathologist with the ability to use informatics to enhance that subspecialty. Because these trainees typically do 1 year of clinical informatics fellowship and 1 year of a diagnostic specialty, our program designates this model the "1 + 1" tract.

The rise of the 1 + 1 fellow raises an important issue for pathology informatics training programs. It is important that all fellows learn a core set of informatics knowledge (otherwise there is no real definition of the pathology informatics domain); however, different parts of the core informatics curriculum (and different parts of the "informatics specialty") will have different degrees of relevance to different fellows depending on their long-term career goals. The fellowship structure must be both flexible enough to accommodate the career goals of each trainee and structured enough to be meaningful.

Our fellowship balances the competing interests of a large, defined domain, and individual trainee interests through a required, didactic common core and customizable mentorships, projects, and research experiences. In this context, this paper addresses the training needs of the pathology informatics fellow who is headed for a career as a molecular pathologist. We have defined the most relevant parts of an informatics curriculum and have suggested specific areas for projects and research. We felt this was an important topic as this combination is common in our program (five out of 17 fellows since 2009). We also feel that our work could inform molecular pathology programs wanting to introduce their trainees to pathology informatics.

An examination of our current educational structure finds that it generally works well for informatics fellows who wish to subspecialize in molecular pathology. While a broad informatics curriculum remains essential and didactics as a teaching modality for pathology informatics have been discussed and published previously,[11-14] we recognize that the fellowship structure must be flexible enough to accommodate the varied career goals of each fellow. Therefore, those interested in molecular pathology may choose to focus more deeply on the most relevant aspects of our curriculum, outlined in this paper. While each fellow's individual experience will vary based upon the institutions active projects, every fellow should seek to leverage informatics approaches any time an issue involving data inherent to molecular testing arises. Such reciprocal molecular and informatics training would be enhanced by the continuity provided by pursuing both fellowships at the same institution.

A discussion of informatics in the context of molecular pathology raises questions regarding the nature and scope of pathology informatics. Are molecular centric techniques such as bioinformatics part of the study (or domain) of pathology informatics? Should all pathology informaticians be expected to be competent in bioinformatics? Or, is bioinformatics a discipline distinct from pathology informatics and the exclusive domain of the molecular lab? These are obviously important questions for a pathology informatics fellowship program and ones that we have thought extensively about.

Bioinformatics in molecular pathology is just one emerging subspecialty discipline. While pathology informatics developed in the age of traditional anatomic pathology, clinical pathology, and classical LIS systems, over the past several years we have seen a number of large, complex, quantitative domains becoming increasingly important in laboratory medicine, including bioinformatics, "big data" statistical analysis, optics, disease modeling, population modeling, image analysis, business analytics, etc., Are all of these part of pathology informatics and if so, how can one teach (or be an expert in) a field that is expanding so fast, in so many directions?

As discussed above, the educational structure of the fellowship has two main arms: One arm is fixed, required for all fellows, and includes a core informatics course. The purpose of the core is to provide exposure, for all the fellows, to the full scope and breadth of pathology informatics, which we define as the study and management of the information, information systems, workflows, and (human and machine) processes of pathology practice. The core does provide exposure to bioinformatics (we feel it is an important domain and technique in pathology), but it is only an exposure. The other arm involves rotations, projects, mentorships, clinical activities, and even courses customized to each fellow to create expertise in the fellow's area of interest and career path. Therefore, while a broad clinical informatics fellowship cannot focus heavily on bioinformatics, a more formal and focused handling of bioinformatics within this curriculum is warranted for those pathology informatics fellows with a focus in molecular pathology. We strongly recommend that these joint fellows pursue projects in collaboration with an institution's bioinformaticians or biostatisticians in order to gain valuable practical experience in the field. Note that fellows planning a career in anatomic pathology or clinical pathology would almost certainly not pursue extensive training in bioinformatics, but may choose to focus, for example, on imaging or statistical analysis. This fellowship structure is in accordance with our view that pathology informatics is both something intrinsic to all of pathology and is evolving its own "subspecialties" to serve the growing informatics needs of an increasingly subspecialized pathology practice.

For molecular pathology, as the testing performed and data generated by these labs becomes more vast and complex, informatics will naturally become interwoven with the specialty. In this paper, we have defined

areas of molecular pathology that could benefit from informatics approaches, and by extension, areas in which pathology informatics fellows could develop relevant and useful molecular pathology operational and research projects. This integrated approach is designed to prepare molecular pathology fellows for the informatics challenges that they will certainly face in their future practice.

REFERENCES

1. McClintock DS, Levy BP, Lane WJ, Lee RE, Baron JM, Klepeis VE, *et al.* A core curriculum for clinical fellowship training in pathology informatics. J Pathol Inform 2012;3:31.
2. Levy BP, McClintock DS, Lee RE, Lane WJ, Klepeis VE, Baron JM, *et al.* Different tracks for pathology informatics fellowship training: Experiences of and input from trainees in a large multisite fellowship program. J Pathol Inform 2012;3:30.
3. Dolled-Filhart MP, Lee M, Ou-Yang CW, Haraksingh RR, Lin JC. Computational and bioinformatics frameworks for next-generation whole exome and genome sequencing. ScientificWorldJournal 2013;2013:730210.
4. Baker M. One-stop shop for disease genes. Nature 2012;491:171.
5. Marsolo K, Spooner SA. Clinical genomics in the world of the electronic health record. Genet Med 2013;15:786-91.
6. Available from: http://www.aruplab.com/files/resources/genetics/White-paper-1-value-of-GCs-in-lab.pdf [Last accessed on October 14, 2013.
7. Brierley KL, Campfield D, Ducaine W, Dohany L, Donenberg T, Shannon K, *et al.* Errors in delivery of cancer genetics services: Implications for practice. Conn Med 2010;74:413-23.
8. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, *et al.* ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med 2013;15:565-74.
9. Jennings L, Van Deerlin VM, Gulley ML, College of American Pathologists Molecular Pathology Resource Committee. Recommended principles and practices for validating clinical molecular pathology tests. Arch Pathol Lab Med 2009;133:743-55.
10. Lee RE, McClintock DS, Balis UJ, Baron JM, Becich MJ, Beckwith BA, *et al.* Pathology informatics fellowship retreats: The use of interactive scenarios and case studies as pathology informatics teaching tools. J Pathol Inform 2012;3:41.
11. Henricks WH, Healy JC. Informatics training in pathology residency programs. Am J Clin Pathol 2002;118:172-8.
12. Henricks WH, Boyer PJ, Harrison JH, Tuthill JM, Healy JC. Informatics training in pathology residency programs: Proposed learning objectives and skill sets for the new millennium. Arch Pathol Lab Med 2003;127:1009-18.
13. Harrison JH Jr, Stewart J 3rd. Training in pathology informatics: Implementation at the University of Pittsburgh. Arch Pathol Lab Med 2003;127:1019-25.
14. Kang HP, Hagenkord JM, Monzon FA, Parwani AV. Residency training in pathology informatics: A virtual rotation solution. Am J Clin Pathol 2009;132:404-8.