Journal of Molecular Diagnostics, Vol. 12, No. 1, January 2010

DOI: 10.2353/jmoldx.2010.090082

Modeling Complex Workflow in Molecular Diagnostics

Design Specifications of Laboratory Software for Support of Personalized Medicine

Mohamed E. Gomah,*† James P. Turley,† Huimin Lu,* and Dan Jones*

From the Division of Pathology and Laboratory Medicine, The University of Texas M.D. Anderson Cancer Center, Houston, and the School of Health Information Sciences at Houston,† The University of Texas Health Science Center, Houston, Texas*

One of the hurdles to achieving personalized medicine has been implementing the laboratory processes for performing and reporting complex molecular tests. The rapidly changing test rosters and complex analysis platforms in molecular diagnostics have meant that many clinical laboratories still use labor-intensive manual processing and testing without the level of automation seen in high-volume chemistry and hematology testing. We provide here a discussion of design requirements and the results of implementation of a suite of lab management tools that incorporate the many elements required for use of molecular diagnostics in personalized medicine, particularly in cancer. These applications provide the functionality required for sample accessioning and tracking, material generation, and testing that are particular to the evolving needs of individualized molecular diagnostics. On implementation, the applications described here resulted in improvements in the turn-around time for reporting of more complex molecular test sets, and significant changes in the workflow. Therefore, careful mapping of workflow can permit design of software applications that simplify even the complex demands of specialized molecular testing. By incorporating design features for order review, software tools can permit a more personalized approach to sample handling and test selection without compromising efficiency. ***(J Mol Diagn 2010, 12:51–57; DOI: 10.2353/jmoldx.2010.090082)***

Laboratory processes in molecular diagnostics are currently highly complex and have proven difficult to standardize and automate. Contributors to this complexity include the wide variety of tissue sample types that can be tested, the numerous steps involved in material preparation, frequent implementation of novel technical platforms leading to variable and unpredictable assay performance, and complex post-testing analysis methods. Another component of the complexity, especially in molecular oncology, is that different tumor samples require different order sets including multistep and reflex testing. Finally, the high cost of some molecular testing often requires prescreening or prioritization of limited samples. Therefore, new sample handling models that are applicable to the clinical laboratory are needed but not yet fully developed.[1]

We present here a discussion of a software design considerations related to the personalized molecular testing, and the results of implementation of applications that address some of these laboratory issues.

Materials and Methods

Project Specifications and Prior Laboratory Workflow Systems

This project was conducted over the years 2006 to 2008 in the Molecular Diagnostics Laboratory at The University of Texas M.D. Anderson Cancer Center, which performs RNA and DNA-based molecular testing for cancer and immunological applications (currently approximately 20,000 samples comprising 30,000 tests per year from cancer patients). The laboratory reports results into two different laboratory information systems (LIS), namely Cerner Classic (Cerner, Kansas City, MO) or Powerpath (IMPAC, Sunnyvale, CA), for blood and bone marrow samples and tissues and body fluids, respectively. The applications described were designed to provide internal laboratory management and workflow and to interface with the LIS reporting systems.

Before the development of these applications, the laboratory workflow included a manual system for tracking

Supported in part by a developmental grant from the Leukemia SPORE (1P50CA100632) awarded by the National Cancer Institute, Department of Health and Human Service. M.G. and D.J. have intellectual property interests in some of the data models underlying the applications described.

Accepted for publication July 14, 2009.

Address reprint requests to Dan Jones, M.D., Ph.D., Department of Hematopathology, Box 72, 1515 Holcombe Blvd, Houston TX 77030. E-mail: dajones@mdanderson.org.

and processing of samples, processed materials and products, use of numerous different spreadsheets for run worksheets, and a referential database for recording testing results. The manual system for sample/specimen routing in our laboratory involved the generation of differently colored cards for each processed material type (eg, RNA, DNA, T-cell sort, etc). These paper cards then tracked along with each of the laboratory workstations until testing was completed and the report was generated.

Definitions of Data Elements

The initial biological materials received in the laboratory, including blood, bone marrow aspirate, other body fluids or paraffin-embedded tissues blocks/slides, are referred to herein as *samples*. The processed DNA, RNA, or protein lysates are referred to as *processed materials* and are sometimes obtained following further purification, microdissection, or cell separation of the samples. These are then used as the substrate(s) for testing, often following PCR amplification, with fragment analysis by capillary electrophoresis, DNA sequencing, microarrays, or other molecular methodologies.

Data elements from the paper test requisition captured at the time of sample accessioning include ordering physician, medical record number, sample ID registered in the LIS, sample type (eg, blood), and test(s) ordered. Data elements produced by the accessioning personnel include sample cell count and sample condition (eg, lysed or clotted blood). The initial status of any test (ie, "active," "screen," "hold," or "cancel") is determined by the accessioning personnel based on the sample type, quality or quantity, previous testing results and the laboratory's internal ordering rules/protocols.

Data elements captured during the sample processing steps are different for DNA, RNA, and protein lysates, as well as for cell sorting and plasma preparation. Minimum data elements collected for most processed materials included the technician(s) doing the processing, quantity and concentration (usually determined by spectrophotometry), as well as quality and purity.

Software Design Process

To produce design specifications for an integrated computerized workflow solution, laboratory personnel were sequentially interviewed, including managers, supervisors, and bench technicians. This allowed us to develop, from the start, a parallel set of workflow diagrams from the perspective of the laboratory managers (Figure 1A), or the bench technicians (Figure 1B). The most experienced technicians in the processing and testing areas of the laboratory were interviewed to obtain information on the granular components of their activities. Finally, all of the findings from the interviews and observations were taken back to the laboratory director for final discussion and refinement before work on the system began. These diagrams were then translated into tools for each of these personnel roles.

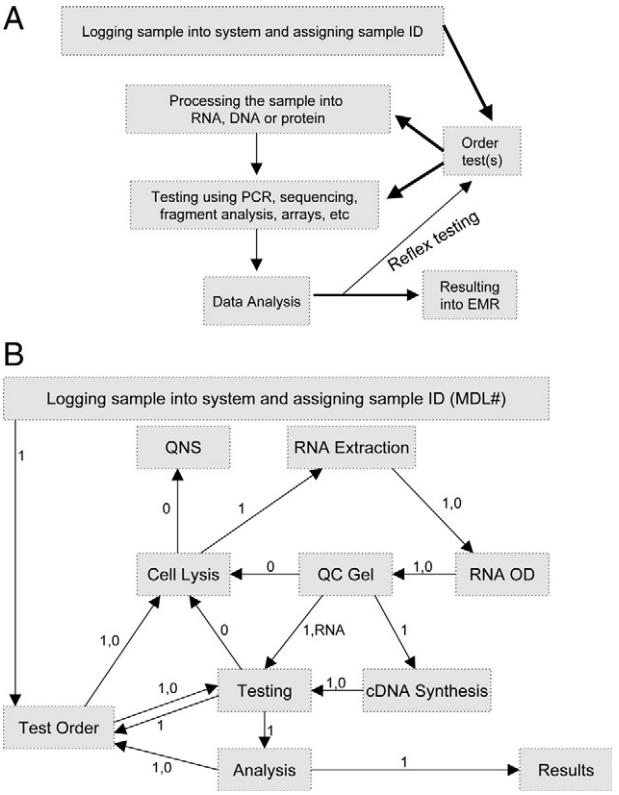

Figure 1. Overview of laboratory workflow. **A:** Design improvements were separately implemented for sample login, processing, and testing. **B:** Laboratory workflow for RNA from the perspective of the bench technician. Workflow at this level is similar to tiered structure necessary for programming, and includes workflow irregularities. Abbreviations: OD: optical density from spectrophotometry reading, QC: quality control, QNS: quantity not sufficient (cancel testing). In flow diagram, 1 indicates "proceed with next step" and 0 indicates "problem" with possible actions of trouble shoot (supervisor review), retest, or cancel.

Programming Tools and Network Architecture

The computer hardware included one database server and a separate application web server. We developed an overall container application composed of modular tools programmed using the .NET 2.0 platform (Microsoft, Bellevue, WA) that was accessed by end-users via a web server. Client machines across the laboratory received software updates automatically as changes were made and uploaded to the application server. These applications connected to a consolidated database hosted on a SQL Server 2000 platform (Microsoft). The database was designed to support complex testing algorithms and dynamic building of the testing library by a nested table structure. The format of results, accessed by external systems, was managed through stored procedure-based queries or by the open database connectivity interface.

Results

Redesign of the Sample Login Process: Implementing Screening Tools to Aid in Accessioning

The core functionality of the accessioning workstation in any molecular laboratory involves receipt of samples with