Composite Patient Reports: A Laboratory Informatics Perspective and Pilot Project for Personalized Medicine and Translational Research

Adi V. Gundlapalli, MD, PhD,MS[1,2,3,4], Julio C. Delgado, MD, MS[2,4],
Brian R. Jackson, MD, MS[2,3,4], Guido J. Tricot, MD, PhD[1], Harry R. Hill, MD[1,2,4]
[1]Departments of Internal Medicine, [2]Pathology and [3]Biomedical Informatics,
University of Utah School of Medicine,
[4]ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, Utah

Clinical laboratories are a strong and integral partner in personalized health care. Laboratory information systems hold a vast amount of data representing human phenotypes, genotypes, biomarkers, progression of disease and response to therapy. These structured and unstructured free text data are critical for patient care and a resource for personalized medicine and translational research. Laboratory data are integrated into many electronic medical records that provide "summary reports" and "trending" to visualize longitudinal patient data. However, these generic reports are not sufficient to manage complex sub-specialty patients. There is an urgent need for end-user driven composite reports for the care of such patients. Using multiple myeloma as a model, this pilot was performed to assess the needs of stakeholders and create a customized report. This laboratory informatics solution is delivered at the point of care through the hospital EMR. Future work will involve further integration with hospital systems to promote clinical decision support and translational research.

Introduction

The goal of personalized health care is to make patient data available at the right time, in the right format and within the normal workflow [1]. The objective is to support providers, decrease time spent in gathering data and improve the quality of care for patients. While this is important for all patients, it is especially critical for sub-specialty patients with complex medical conditions and longitudinal data. The availability and easy access to these data are important for the sub-specialty provider for routine patient care and for others called upon to provide cross-coverage or emergency care [2].

Clinical laboratories generate and archive patient data in laboratory information management systems (LIMS). These data include biochemical manifestations of organisms (phenotypes), genotypes and biomarkers that are essential to the diagnosis, determination of progression and response to therapy of a disease. These data form the foundation for clinical care, personalized medicine and translational research from an informatics standpoint of data capture, organization, integrity and flow [3]. Where available, laboratory data are integrated into electronic medical records and are accessed routinely by medical personnel. Longitudinal or serial laboratory tests are usually viewed graphically or in tabular form via summary reports or trends [4]. These reports are useful for visualizing structured data in the form of numeric values with flags for abnormal values. They are less optimal for data that are available only in unstructured free text in written reports or pictorially represented in graphs and gels from molecular genetics, anatomic pathology and other specialized testing such as protein immunology.

The medical care of patients in fields such as transplant medicine, cancer, HIV/AIDS, etc. generates large volumes of data including critically important serial laboratory data. Currently, there are no readily available composite laboratory data reports that incorporate both structured and unstructured elements for use in the care of such complex medical patients. Prior informatics work [5-9] and our clinical partners have indicated that such automated reports would be useful in improving care while reducing time in gathering and collating these results manually.

Using multiple myeloma as a model disease, this pilot project addresses the hypotheses that: (1) Clinical providers perceive composite laboratory reports to be important for the care of complex patients and (2) Such reports can be generated using laboratory informatics methods. Multiple myeloma (MM) is the second most common cancer of the blood in which antibody producing plasma cells become malignant [10]. Routine clinical care of these patients involves a time-consuming data gathering phase where the results of a battery of immunology tests, biomarkers and more recently, gene expression

data are accessed and collated to assess progression of disease and response to therapy [11-13].

Setting

This pilot project was carried out at the University of Utah School of Medicine and ARUP Laboratories (ARUP). The University Myeloma Program consists of a dedicated staff of physicians, mid-level practitioners, pharmacists, nurses and other support personnel. ARUP is a national clinical and anatomic pathology reference laboratory that is owned by the University and performs all the testing for MM patients from the University. ARUP performs nearly 1500 protein immunology tests per week from the University and other clients. The LIMS at ARUP consists of a commercial system with a direct interface to the University of Utah electronic medical record (EMR). Results of tests are available at the University immediately upon finalizing and approval in the ARUP system.

Methods

A semi-structured survey and interview was conducted among the clinical staff of the University Myeloma Program, that included: (1) Questions regarding their role in the team, years of experience and hospital EMR access patterns; (2) Their time spent in gathering and correlating laboratory data for patients with special emphasis on protein immunology tests and (3) Their opinions on a composite report and choice of protein immunology tests if a such a report were to be available for myeloma patients. The existing flow of data for laboratory tests between the University Hospital and ARUP was mapped. A rules-based inference engine was developed to extract specific test results identified by the end-users. These data provided the design rationale of a sample composite report for MM patients.

Results

End-user survey: All ten members of the Myeloma program that routinely access patient labs participated in this pilot project. Their experience in this field was on average 9 years (range 1-30 years). All accessed the EMR multiple times per day and the single most accessed tab in the EMR was the laboratory results screen. The team members spent an average of 18 minutes *per patient* gathering all laboratory data and an average of 4 minutes per patient on protein immunology labs. Only 6 of the 10 indicated that they were either familiar with or used the "trend" or "graph" feature of the EMR to view serial labs with numeric results. All providers accessed the free text interpretation of the serum protein electrophoresis (SPEP) and immune fixation electrophoresis (IFE) as that was the only way to learn about the presence of a myeloma protein, its quantitation and type. Only 7 of the 10 indicated that they accessed and viewed a pdf file of the actual gels offered via a secure website (ARUP Enhanced Reports [14]). All of them indicated a desire to see multiple labs on a single report with the ability to view serial changes in key myeloma biomarkers. Nearly all (8 of 10) expressed a willingness to collaborate with laboratory informatics teams to brainstorm the ideal composite report. They were also willing to participate in a validation study of the benefits of such a report for clinical care. All were in favor of providing this composite report directly to patients. Two key elements were highlighted in discussions with the team: (1) the access to and downloading of disparate protein immunology lab data along with the free text interpretation of SPEP and IFE was challenging and time-consuming, often requiring the simultaneous use of two computer screens and (2) A composite report with oft-used results would benefit patient care and improve the work flow.

Data flow of laboratory orders and results: At the University of Utah, laboratory orders are currently initiated on paper, and then entered into the laboratory information system by clerks and laboratory staff. Orders flow from the hospital information system (the "Olympus" system, powered by Cerner Millenium PowerChart; Cerner Corporation, Kansas City, MO) over an HL7 interface to ARUP's laboratory information system (Cerner Millenium PathNet). As shown in Figure 1, ARUP maintains a separate information technology organization from the hospital, including staff and infrastructure, and so Olympus runs on a separate instance of Cerner Millenium from the LIMS. Once laboratory testing is complete, test results are returned via HL7 interface to Olympus. Results are simultaneously copied to both the hospital clinical data warehouse and ARUP's own long term repository. Although ARUP's LIMS remains the primary source of test results (data is retained approximately 90 days online, and in archived form for 7 years), order and result data are replicated in a SQL Server database (Microsoft, Redmond, WA) called the ANSR system (ARUP's Networked System Repository) in order to support a variety of in-house developed software programs. It is from this

latter database that ARUP's composite and enhanced laboratory reports are generated. A knowledge- and rules-based inference engine was developed that stores the logic to determine which laboratory data to present in enhanced form and on which patients, based on the data present in the ARUP data repository. The rules are simple if-then-else business rules that identify patients with immunology tests and their results. When specified lab data enter the data repository, the engine triggers the generation of the actual composite report based on an end-user defined information layout, including numeric data, graphs, gels and text.

During the calendar year 2007, ARUP performed a total of 4699 protein immunology tests on 1450 unique patients from the University. As noted in the table, these tests are performed multiple times on myeloma patients as serial results are important for the determination of progression of disease and response to therapy. Accessing and correlating even the last 3 results of tests that are reported in free text such as SPEP/IFE poses a challenge to providers using the existing reporting format. Often, results of two or more different tests are used in the assessment of the patient's disease status.

Sample report: A sample composite report has been created (Figure 2) that captures the most often used biomarkers for myeloma patients. The design rationale for this report is based on the survey results and semi-structured interviews with providers who care for these patients. The free-text interpretations of the SPEP and IFE results are provided by the ARUP immunology lab and transmitted to the University EMR as part of the report. These results provide key information regarding the presence, type and quantitation of a myeloma protein in the patient's blood. Users are able to view and download the report as a pdf file via a secure website. This feature has been successfully implemented for providing other enhanced reports. ARUP provides secure access to the website with a computer generated log-in and password available only to authorized viewers of the original report [14].

Table. Protein immunology testing at ARUP for University patients, January 1 – December 31, 2007.

SPEP = serum protein electrophoresis
IFE = immune fixation electrophoresis
Q FLC = quantitative free light chains
Ig = immunoglobulin

Protein immunology test	Total number of patients receiving this test	Average number of orders per patient	Max. number of orders per patient
SPEP with reflex to IFE	**18**	**1**	**1**
SPEP alone	**741**	**1.99**	**34**
IFE, Ig A/G/M	**746**	**1.32**	**12**
IFE, Ig D/E	**19**	**1.05**	**2**
Urine Q FLC	**473**	**1.72**	**19**
Serum Q FLC	**150**	**5.15**	**40**
Serum Ig	**143**	**3.84**	**31**

Limitations

This is a pilot project that was implemented with only one disease and one set of laboratory data. The unstructured free text interpretations of test results such as SPEP and IFE were not converted to "structured" data. The existing interface to the hospital EMR was used to provide the report in the EMR along with other protein immunology results. At this time, no information extraction is being performed on the free text sections of the report using text processing methods that have been applied to pathology and radiology reports [15, 16]. Similarly, there are no analytics performed on the composite report such as temporal abstraction for trends, patterns and inferences to the data [7, 17].

Future Work

An important next step is to design and perform validation studies to document the utilization, usability and benefit of these composite reports for patient care. We are planning a validation with the Myeloma team.

There exist opportunities for us to mine the composite report and extract information using text processing [15, 16] and temporal abstraction methods [7, 17]. Another opportunity is to be able to use a simple ontology that links patients, their diseases and specific batteries of tests, including serial ones (possibly with LOINC and

SNOMED as vocabularies). These relationships would provide an advanced knowledge base for the inference engine that drives the generation of the composite report.

Conclusions

There are several key steps in realizing the goal of personalized medicine and having the ability to review integrated laboratory data in the electronic medical record that is both available at the point of care and be useful for clinical care. The modern clinical laboratory has moved beyond reporting results that consisted of only of a numeric value with reference ranges and has to innovate and adapt to be a key partner in personalized medicine and translational research.

This novel descriptive project brings together multi- and inter-disciplinary stakeholders to understand the needs of the users and provide comprehensive patient reports that contain both structured and unstructured data. The intent is to decrease the information gathering burden on providers and support clinical care. This pilot laboratory informatics perspective to personalized medicine seeks to inform future collaborative work in this area. The data flow and informatics solutions identified here will be directly applicable to developing complex reports involving molecular genetics, gene expression, micro-array and biomarker data.

The key issues identified by this project are (1) Composite reports for the care of complex medical patients are disease- and end-user specific and thus calls for an inter-disciplinary approach; (2) Unstructured free text interpretations of tests are important elements and so further work needs to be done to provide "structure" to these data at their origin in the clinical laboratory; (3) Integrating these reports with the EMR poses a challenge that needs to be addressed before clinical decision support can be offered based on these reports and (4) There needs to be a mechanism for providing these reports directly and securely to patients.

Acknowledgements

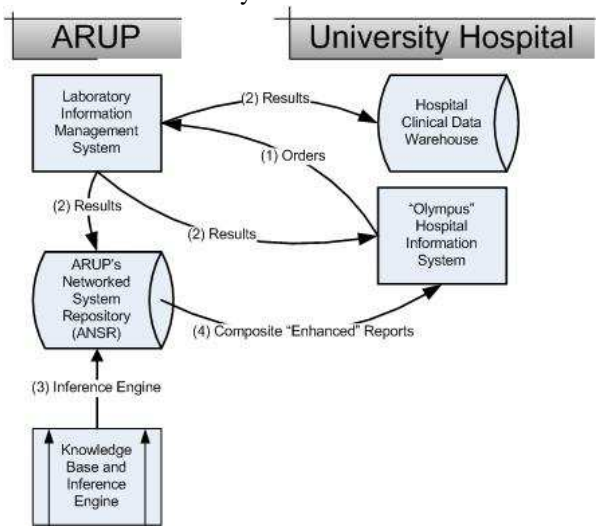
The authors gratefully acknowledge the professionals who work in the Myeloma program at the University of Utah School of Medicine/Huntsman Cancer Institute including Ashlie Finlayson; the protein immunology laboratory at ARUP including Nancy Pitstick, Robbie Myatt and Gregory Philipps; laboratory informatics at ARUP including Andrew Wilson and David Davis; and Brett South (Internal Medicine and Biomedical Informatics). We also acknowledge funding and support from the Department of Pathology and the ARUP Institute.

References

1. United States Department of Health and Human Services, *Personalized Health Care: Opportunities, Pathways, Resources*. 2007: Washington DC.
2. Hripcsak, G., et al., *Emergency department access to a longitudinal medical record.* J Am Med Inform Assoc, 2007. **14**(2): 235-8.
3. The Harvard Medical School-Partners HealthCare Center for Genetics and Genomics (HPCGG), *Improving Health and Accelerating Personalized Health Care through Health Information Technology and Genomic Information in Population- and Community-based Health Care Delivery Systems*. 2007. p. 46.
4. Denekamp, Y., et al., *Exploring methods of visualizing laboratorial data in electronic medical record systems.* AMIA Annu Symp Proc, 2006: p. 907.
5. Smellie, W.S., *What is a significant difference between sequential laboratory results?* J Clin Pathol, 2008. **61**(4): p. 419-25.
6. Klimov, D. and Y. Shahar, *Intelligent querying and exploration of multiple time-oriented medical records.* Stud Health Technol Inform, 2007. **129**(Pt 2): p. 1314-8.
7. Shahar, Y., et al., *Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions.* Artif Intell Med, 2006. **38**(2): p. 115-35.
8. Klimov, D. and Y. Shahar, *A framework for intelligent visualization of multiple time-oriented medical records.* AMIA Annu Symp Proc, 2005: p. 405-9.
9. Koski, E., et al. *Visualizing Complexity - Medical Test Reporting for the 21st Century*. in *MEDINFO 2004*. 2004. Amsterdam: IOS Press.
10. Kyle, R.A. and S.V. Rajkumar, *Multiple myeloma.* Blood, 2008. **111**(6): p. 2962-72.
11. Dispenzieri, A., et al., *Appraisal of immunoglobulin free light chain as a marker of response.* Blood, 2008. **111**(10): p. 4908-15.
12. Van Ness, B., et al., *Genomic Variation in Myeloma: Design, content, and initial application of the Bank On A Cure SNP Panel to detect associations with progression free survival.* BMC Med, 2008. **6**(1): p. 26.
13. Mulligan, G., et al., *Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib.* Blood, 2007. **109**(8): p. 3177-88.
14. ARUP Laboratories. *ARUP Connect - Enhanced Laboratory Reports*. 2008 [cited 2008 September 22, 2008]; Available from:

http://www.aruplab.com/Client-Services/arup-connect-login.jsp.
15. Friedman, C., et al., *A general natural-language text processor for clinical radiology.* J Am Med Inform Assoc, 1994. **1**(2): p. 161-174.
16. Fiszman M, C.W., Aronsky D, Evans RS, Haug PJ, *Automatic detection of acute bacterial pneumonia from chest X-ray reports.* J Am Med Inform Assoc, 2000. **7**(6): p. 593-604.
17. Shahar, Y., et al., *Interactive visualization and exploration of time-oriented clinical data using a distributed temporal-abstraction architecture.* AMIA Annu Symp Proc, 2003: p. 1004.

Figure 1. Data flow and architecture of laboratory test orders and results between ARUP and the Hospital



Data flow

(1) Laboratory orders are transmitted from the Hospital to ARUP

(2) Laboratory results are transmitted from the ARUP LIMS to the Hospital Information System, Hospital Clinical Data Warehouse and the ARUP System Repository

(3) Pre-defined rules are applied to the data in the repository using a knowledge base and inference engine

(4) On-demand and automatic generation of composite enhanced reports sent to the Hospital Information System

Figure 2. Sample composite report for multiple myeloma patients. The design rationale is based on end-user preferences in terms of tests to be displayed, longitudinal data, free-text reports and images of gels

ARUP LABORATORIES

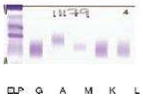
ARUP *Composite Enhanced Report*
Protein Immunology Testing

Name: John Doe Date of Birth: 01/01/1920 Account Number/Unique Identifier: 123456789

SPEP/IFE

7/1/2008

Interpretation: Normal SPEP, IFE shows no monoclonal proteins



Last 1 month:
6/1/2008
Interpretation: M-spike in the gamma region. The monoclonal protein peak accounts for 0.93 g/dL of the total 1.11 g/dL of protein in the gamma region. IFE reveals Ig G Kappa monoclonal protein.

Last 3 months

Other Protein Immunology tests:
Bence Jones Protein, Quant

Page 2 as needed

Signed
Chief Medical Officer/Laboratory Director

Quantitative Immunoglobulins

Date	Ig G 768-1632	Ig A 68-378	Ig M 60-263
7/1/2008	1100 (N)	26 (L)	12 (L)
6/1/2008	2400 (H)	88 (N)	144 (N)

Kappa-Lambda Qnt Free Light Chains (FLC)

Date	Kappa Qnt FLC 0.33-1.94	Lambda Qnt FLC 0.57-2.63	K:L FLC Ratio 026 – 1.65
7/1/2008	0.36 (N)	0.23 (L)	1.57 (N)
6/1/2008	5.89 (H)	0.44 (L)	13.39 (H)

Urine IFE

Date	Result
7/1/2008	Urine is NEGATIVE for monoclonal Free Light Chains (Bence Jones Protein).
6/1/2008	Urine is POSITIVE for monoclonal kappa light chains