Journal of
Health & Medical Informatics

Zhao et al., J Health Med Inform 2016, 7:5
DOI: 10.4172/2157-7420.1000244

Research Article Open Access

An Enterprise Approach for Pathology Reporting Utilizing PDF Functionality for the Electronic Medical Record

Zhao T[1], Stamelos EMS[2], Tsividakis G[2], Gurman J[3], Mitchell MJ[4], Klimstra DS[2], Reuter VE[2] and Sirintrapun SJ[2*]

[1]*Department of Pathology, Wake Forest Baptist Health, Medical Center Boulevard, Winston-Salem, North Carolina, USA*

[2]*Department of Pathology, Memorial Sloan Kettering Cancer Center, York Avenue, New York, USA*

[3]*Clinical Systems Administration, Memorial Sloan Kettering Cancer Center, New York, USA*

[4]*Senior Project Manager, Clinical Systems and Hospital Administration, Memorial Sloan Kettering Cancer Center, New York, USA*

***Corresponding author:** Sirintrapun SJ, Department of Pathology, Memorial Sloan Kettering Cancer Center, York Avenue, New York, USA, Tel: 212-639-2169; E-mail: sirintrs@mskcc.org

Received date: September 01, 2016; **Accepted date:** October 19, 2016; **Published date:** October 26, 2016

Copyright: © 2016 Zhao T et al. 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: Shortcomings of plain text renderings of pathology reports in the electronic medical records (EMRs) are widely known. As the pathology reports grow in complexity, Portable Document Format (PDF) renderings of pathology reports are seen as a flexible solution. Since 3/2014, our group has successfully implemented architecture for PDF functionality for pathology report integration into the EMR.

Design: We outlined our architectural framework. Within our architectural framework, pathology reports are first generated as MS Word documents. They are then exported into both HL7 and PDF formats in the EMR. The clinicians may then choose to view either format. The HL7 interface message contains the file location path of the PDF, and the PDF reports are stored on the hospital servers for display retrieval within the EMR. As the HL7 feed is maintained, data can still be extracted into the institutional data warehouse (IDW). The retained HL7 rendering of the report allows copy/pasting of the pathology report into clinical documents. A reconciliation report is generated to match the transmissions sent with those received.

Results: The number of PDF reports sent across the interface from 3/2014 to 5/2015 total 191,868. The monthly reports sent range from 11,449 to 14,422 per month, with a median & average of 12,419 and 12,791, respectively.

Conclusions: PDF renderings of pathology reports have been successfully integrated into the EMR system. PDF reports bypass errors that may occur when HL7 strips the formatting from MS Word documents, such as blood counts or gene mutation lists. Use of visual analytics is also possible with PDF reports. Most commercially available EMRs have the capability of displaying PDF documents. We hope that our experience can serve as a framework for PDF functionality in EMRs.

Keywords: Electronic medical records; Pathology; Electronic health systems; Hospital information systems

Abbreviations

AP-LIS: Anatomic Pathology Laboratory Information System; CDA: Clinical Document Architecture; EMR: Electronic Medical Record; HITECH: Health Information Technology for Economic and Clinical Health; HISs: Hospital Information Systems; HL7: Health Level-7; IS: Information Systems; IDW: Institutional Data Warehouse; MS Word: Microsoft Word; NGS: Next Generation Sequencing, PDF: Portable Document Format

Introduction

The fundamental role of pathologists is to examine tissues, analyze lab tests, and interpret the results to arrive at a diagnosis. Pathologists are considered "gatekeepers" of the tissues, and now pathologists must acknowledge their role as "stewards" of information [1]. The amount of data available to clinicians and researchers in the healthcare system continues to grow at an unprecedented rate [2]. Pathology results currently not only encompass the pathologic diagnostic information, but also quantitative laboratory results such as bone marrow counts and molecular information such as next generation sequencing (NGS) results. The rapid and increasingly wide-spread adoption of NGS in the clinical arena is expected to bring a paradigm shift in the practice of pathology [3]. As such, the incorporation of such results in a readable form for clinicians is essential. The pathology report is a representation of the patient's disease process. Through the pathology report, the pathologist not only provides the diagnosis and appropriate documentation, but also determines how that information is presented and communicated to the clinicians. How the anatomic pathology report is transformed and integrated in the electronic medical record (EMR) can critically affect decision making [1].

Many labs still transmit pathology results in plain text due to the constraints of conventional messaging standards used within laboratory information systems (LIS) and EMRs. A pathology report generated in Microsoft Word, stripped of all its formatting, cannot communicate the results as effectively, e.g., numbers that were in tables are no longer aligned. Use of Portable Document Format (PDF) based

reporting facilitates more effective communication of pathologic diagnoses.

Since February of 2014, our group has successfully implemented PDF functionality for pathology report integration into the EMR, which allows for more intuitive layouts of pathology results. We describe our experience in implementation of this PDF functionality for pathology reporting. Here, we review the modular components of an interfaced system architecture in the hope that our system can serve as a near term model solution for back end clinical decision support.

Technical Background

There are many intricacies to consider in interactions between electronic health systems, as nearly every medical center is set up differently. Some centers have integrated architectures whereby the LIS is embedded in the EMR. Most other centers use interfaced architectures [1]. Within an interfaced architecture, there are modular components of the system for which each component is briefly reviewed.

System Components and Functionality for an Interfaced Architecture

AP-LIS

The Anatomic Pathology Laboratory Information System (AP-LIS) is an information system which serves many functions for an anatomic pathology laboratory including being the workspace environment/application for operational recording, tracking procedures, specimens, slides, and reports [1]. Our version of the AP-LIS is Cerner CoPath version 2013. The AP-LIS provides functionality for results reporting in which diagnoses are documented and communicated for clinical decision making. Much reporting in anatomic pathology is text based and at times narrative (e.g., microscopic descriptions), despite the widespread use of synoptic reporting for cancer specimens. Molecular reports are generated within the AP-LIS while the analysis and interpretation are done in separate applications. These molecular results are usually reported as addendums to existing pathology reports, which may be buried in the merged, single, plain text report displayed in EMRs. There is often a disjoint of the pathology report generated in a rich text editor and what is displayed in the electronic medical record (EMR) system. As a result, a report's accuracy and points of emphasis may be lost, and clinicians may misinterpret what's in the report.

Electronic Medical Record (EMR)

An electronic medical record (EMR) is the main hospital information system, systematically collecting medical information [1]. The EMR serves as the workspace environment for clinicians by which all information is collected, aggregated, and presented in a patient-centric manner [4]. Though the EMR is thought of as the digitized version of the paper based medical chart [5]; the EMR goes beyond that since results are updated, distributable, and readily available [1]. Results come from various information systems separate from the EMR under an interfaced architecture. What the pathologist sees in the AP-LIS rich text document editor (e.g., MS Word) is different from what displayed in the EMR (i.e., plain text).

The version of our EMR is Allscripts 6.1. This EMR serves as the main clinician workspace environment and handles all care which occurs within our institution. Display of information was mainly in plain text until our recent implementation of PDF functionality. Outside hospital records of new patients are not viewed through this EMR system. Thus our institution maintains a legacy EMR system used for displaying scanned external documents for new patient visits. This is an important consideration which will be further elaborated on in the approach and discussion sections.

Report MS Word

Microsoft Word (MS Word) is the rich text editing program for most AP-LIS vendors. Our current version of MS Word is 2007. After an accession associated with a patient's specimen is generated, and the pathologist is ready to enter results, the AP-LIS opens a linked MS Word document. The pathologist can then enter diagnoses, comments, any references, or hospital lab disclaimers. Formatting is allowed such as different font size, bolding, and italicizing to produce a visually intuitive report and place emphasis where appropriate. Some pathology practices use graphical layouts with pictures of slides/stains and tables. The pathology report generated as a MS Word document was initially intended to be placed in a paper based medical chart either as a faxed or paper copy. However the "back end" translation of the generated pathology report involves Health Level-7 (HL7) messaging which lacks the flexibility to accommodate such items. Thus, all the formatting, attached pictures, table formatting, etc. are stripped from the report shown in the EMR. HL7 will be discussed in the following section.

Health Level-7 (HL7) and the Interface Engine

Both the AP-LIS and the EMR are categories of hospital information systems. With the advent of widespread implementation of EMR systems, communication between the AP-LIS and the EMR became necessary under an interfaced architecture. The two information systems are separate applications, and their data are housed in separate respective databases. Under an interfaced architecture, anatomic pathology reports are rendered in the EMR via communication of data between the two systems using Health Level-7 (HL7) standards. HL7 refers to a set of international standards for exchange, integration, sharing, and retrieval of electronic clinical data and administrative data between hospital information systems [6]. These standards define how information is packaged and communicated from one system to another, allowing for seamless integration between systems [7]. The interface engine is an application that facilitates this messaging "talk." Interface engines work like interpreters who translate different languages and enable hospitals and healthcare organizations to communicate with each other. With HL7, existing information systems can still be utilized rather than purchasing one enterprise wide information system in order to transfer clinical and administrative information [6]. HL7 messaging has its limitations; one being that graphical layouts and formatting are not relayed [1]. It is due to this limitation of HL7 that all graphics and formatting are lost when a pathology report is transmitted and reconstructed in the EMR.

Institutional Data Warehouse (IDW)

The institutional data warehouse (IDW) is a collection of institutional data and data processing tools that are constructed for querying and are accessed by authorized personnel. For many institutions, their IDW is constructed to support not only the current data sharing among healthcare givers, administrative personnel, and patients but also future initiatives (8). The IDW can be used to access data in a timely manner and can also be used to conduct data mining

and data analysis to support decision making, research, and even transfer of information to health information exchanges to support "meaningful use" initiatives.

Procedure (System Design and Information Flow)

Figure 1 outlines the system design and information flow for pathology reporting under our interfaced architecture. The pathologist uses the AP-LIS/MS Word to generate and edit pathology reports. Once the report is finalized and signed out in the AP-LIS, the report (i.e., data) is packaged via HL7 standards and pushed to the interface engine. The interface engine then routes the HL7 data stream to various information systems such as the EMR and IDW. The reconstruction of the HL7 data stream in the EMR is displayed in plain text which can diverge from the original formatting. Such a system design represents the majority of interfaced information systems architectures currently implemented. From here we describe differences in our system design from the majority of the currently implemented interfaced information systems architectures.

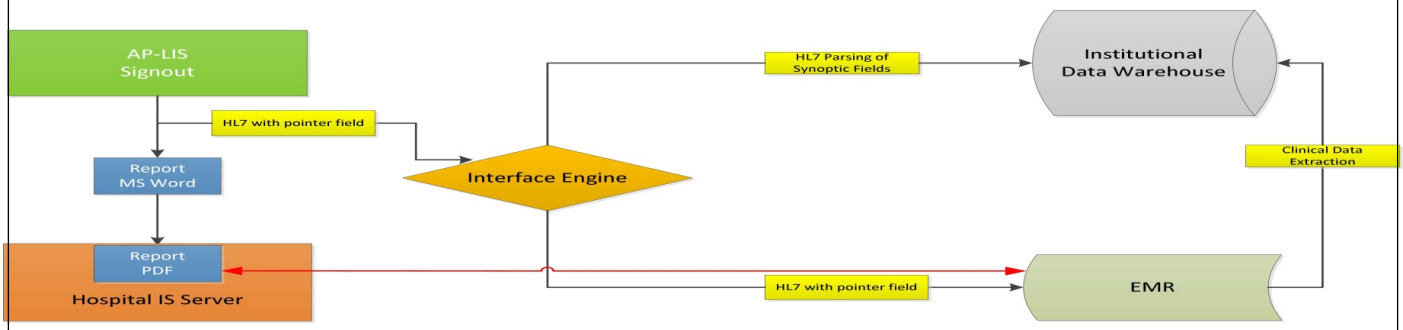


Figure 1: System design showing flow of information under our interfaced architecture.

One modification is that once the report is finalized and the HL7 data stream is generated, this triggers the creation of a PDF rendering of the pathology report by the AP-LIS. This generation of a PDF version of the MS Word report is a customization available under our AP-LIS (Cerner CoPath version 2013). The PDF files are then placed on an accessible hospital server. These PDF versions of pathology reports are virtual prints of the reports that pathologists edit and sign out under MS Word. Additionally, the HL7 data stream that is initially pushed to the interface engine contains a pointer field within the message that provides the link for the EMR to pull the corresponding PDF file from the hospital server for viewing. Thus, both the HL7 reconstructed plain text rendering and the PDF version are available for view in the EMR. The reconstructed plain text rendering is available for copying and pasting into clinical notes since directly copying and pasting from the PDF may run into technical performance issues.

Because our pathology department utilizes standardized reporting, the final diagnosis sections of our pathology reports are structured and consistent, making the reports amenable to parsing and data extraction. Our IDW has a computational team that creates natural language parsing algorithms to parse the final diagnosis field of the HL7 data stream for extractions into IDW. Therefore, the maintenance of the HL7 data stream not only enables triggers for the EMR to pull the PDF reports for display in the EMR but also makes the final diagnosis field available to IDW.

Figure 2 illustrates the number of PDF reports sent across the interface from March 2014 to May 2015. There are a total of 191,868 reports, ranging from 11,449 to 14,422 reports per month with a median and average of 12,419 and 12,791 reports respectively. Moreover, the PDF format has been very useful for reports like hematopathology and molecular pathology in which the complete blood counts and gene mutation lists were improperly laid out under old HL7 plain text renderings in the EMR.

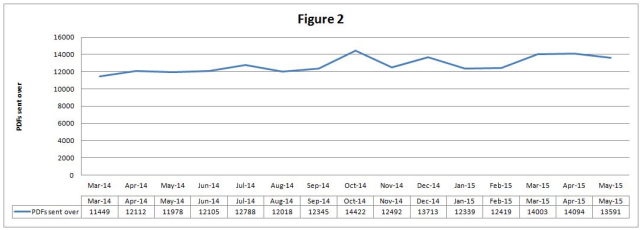


Figure 2: Total number of PDF reports sent across the interface March 2014 to May 2015, numbers 191,868 reports. The reports range from 11,449 to 14,422 reports per month with a median and average of 12,419 and 12,791 reports respectively.

Approach

Our procedure in system design and information flow is not the only mechanism for enabling display of PDFs in the EMR. Another procedure is encoding the PDF document using the Base64 algorithm and transmit that data into the EMR via HL7 standards directly. We selected our system design mostly because of the inability of our EMR to handle an actual PDF embedded/encoded within an HL7 data stream. The only alternative was then to have the EMR receive a pointer in the HL7 data stream to a hosted server for PDF display.

One positive feature of the Base64 mechanism is that this solution does not depend on co-location of file storage between LIS and EMR. However, this same feature does not meet with our internal compliance requirements of having a standing document available for future retrieval, as the encoded PDF is not stored. In addition our system design accommodates our older legacy EMR system, used for displaying scanned documents from outside hospitals/clinics for new patient visits. The older legacy system is unable to store or display a PDF in that is embedded in an HL7 message. It is not even able to

display any PDF documents. Rather, scanned PDF documents are converted to TIFF images, which are then stored for display in that system. This made our system design more compatible with path reports accessible at scale for the PDF to TIFF conversion process.

With the Base64 mechanism, if older legacy EMR systems are not able to store the PDF in an embedded PDF HL7 data stream, problems arise. Future retrieval of the PDF is not possible without re-transmission of the HL7 data stream. This then requires a request and re-transmission pipeline which can add burden to already taxed pathology and EMR teams.

Base64 potentially offers a better solution with EMRs that can accept and store the PDFs in PDF embedded HL7 data streams. In our system design, because the PDFs are hosted on internal servers, these files are not retrievable for clients outside of our institutional firewalls. The outside clients who have referred patients to our institution cannot view these PDF reports, and faxing remains the primary method of delivering reports. Outreach has historically not been a priority for our institution. If in the future outreach becomes more mission critical, we may turn to Base64 as a solution.

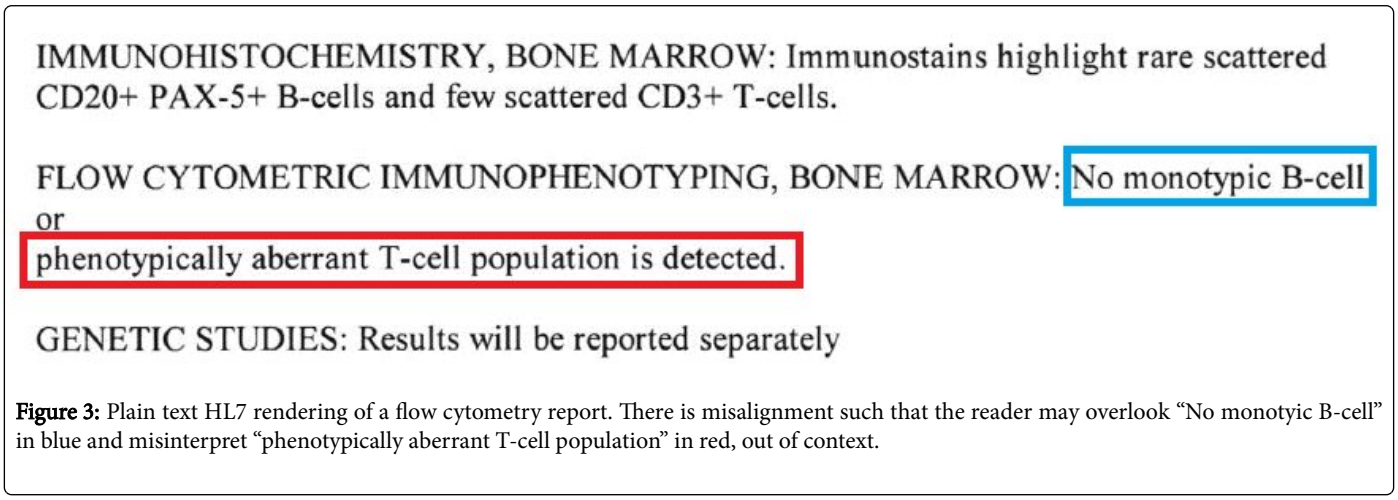
Our system is also designed to simplify adhering to internal compliance requirements with pathology report amendments. All pathology report PDFs are hosted on a server which is maintained by the EMR teams. Thus management of various versions of pathology reports falls mainly under the EMR teams, which also relieves workload on our pathology information system (IS) team. The original report along with the amended report is available and accessible to the EMR teams for comparison. Such a mechanism deters drastic revisions of reports without reference to what has changed. By having the original available, there is a baseline for comparison. This creates accountability for changes made. It is easier to spot major changes within the amended report when the report layout has headings (e.g., Final Diagnosis, Comments) that catch the clinician's attention. The changes may be buried within the sea of plain text in a standard HL7 data stream reconstructed report. The older legacy system does not provide an intuitive display of multiple versions of pathology reports. This is another advantage of PDF reporting which will be elaborated on further in the discussion section.

Discussion

What is, in essence, a current day pathology report?

If we want to discuss the importance of enhanced pathology reporting, we must address what lies at the heart of a pathology report. A pathology report is a documentation tool. Pathologists document the list of pathologic findings, the workup performed, and the interpretation of what is described. The pathology report must document the pathologic findings relevant for clinical decision making (back end decision support), i.e., the main (final) diagnosis. Examples of clinically relevant pathologic findings include tumor histologic subtype, margin involvement, and pathologic stage which are critical for further management of the patient. Pathologists also may cite findings which are sometimes interesting and sometimes for academic curiosity, often in comment sections.

Another role of the pathology report is as a communication tool. The diagnosis rendered plays a critical role in clinical decision making. Any misrepresentation of the results can lead to catastrophic consequences. The misalignment of results can occur when pathology results are packaged through HL7 standards and rendered in the EMR. Tabbing and spacing may look almost identical in MS Word, but the alignment doesn't translate to plain text. Misalignments of hematopathologic, immunohistochemical, and molecular results may seem more annoying and less consequential. However, there may be instances of design flaws inherent to the information systems that generate specific hazards (9). Consider the deleterious potential for a misalignment of "no carcinoma identified" where "no" shows up at the end of one line/page and the next line/page states "carcinoma identified" (1). Figure 3 illustrates a plain text HL7 rendering of a flow cytometry report where there is a misalignment, which can easily result in a misinterpretation for presence of a critical abnormality. Clearly this would not match the intent of the pathology report.


IMMUNOHISTOCHEMISTRY, BONE MARROW: Immunostains highlight rare scattered CD20+ PAX-5+ B-cells and few scattered CD3+ T-cells.

FLOW CYTOMETRIC IMMUNOPHENOTYPING, BONE MARROW: No monotypic B-cell or
phenotypically aberrant T-cell population is detected.

GENETIC STUDIES: Results will be reported separately

Figure 3: Plain text HL7 rendering of a flow cytometry report. There is misalignment such that the reader may overlook "No monotyic B-cell" in blue and misinterpret "phenotypically aberrant T-cell population" in red, out of context.

There are other considerations beyond more accurate representation of the pathology results. We are moving increasingly towards personalized medicine, and with it, towards increasing amounts of information. As a communication tool, the pathology report needs to provide visual cues outside of the plain text to aid the reader in extracting and comprehending the information contained in the report. Simple measures such as bolded text, headings, and hierarchical use of indentations provide this facility which plain text reports lack. We will revisit this constraint later in this section when discussing next generation sequencing (NGS) results.

How PDF enables enhanced pathology reporting

Exchange of PDF documents via messaging exchanges is not necessarily an innovation, per se. It is part of HL7 messaging standards, and some pathology groups have already implemented this functionality. However, implementation of PDF reporting within EMRs has not been discussed broadly in the literature, and for this reason we decided to describe our experience.

In order to provide enhanced pathology reporting capabilities in the EMR, we have implemented a PDF functionality to overcome the limitations of HL7 reconstructed reports. We have described our framework and extensive experience for implementation of PDF functionality for pathology reporting within the EMR. Errors in conveying information which can occur through formatting issues in HL7 renderings are eliminated. What the pathologist intends for clinician view is displayed via the PDF format and viewed in the EMR. Bolded text, varying font sizes, and hierarchical use of indentations make the report more intuitive to read. The report in PDF format also creates opportunities for complex reporting since it enables the use of colors, photomicrographs, figures, graphs, tables, and charts, etc. Pathologists will be able to better support clinicians by having surgical pathology reports that are easier to read and comprehend. Less confusion by clinicians will translate to improved quality of care.

What about amended reports?

As touched upon previously in the approach section, it is difficult to compare the original and revised reports side by side in our legacy system. Our legal department has mandated that both the original and revised report must be combined into one report. Pathology reports can often be lengthy with synoptic reports and multiple addendums to report molecular results, stain results, send-out tests, etc. The final report combining the original report with all its addendums and all the revised versions makes a very large sea of plain text. The clinicians may find it difficult to dig within that sea of text for the pieces of information they need. After the implementation of PDF functionality, our IS teams are evaluating two solutions. Under our system design, because PDFs are housed on a hosted server, the prior PDFs can be pulled back to our AP-LIS and incorporated into the current report version as an image using the Base64 mechanism as described previously. Unlike our EMR systems, our AP-LIS does have certain Base64 capabilities. Another mechanism is to have all strings of text from the prior report pulled into the current report version. The strategy is to then utilize fonts, bolds, colour schemes, etc. to emphasize the relevant current information and/or de-emphasize prior information. Again, the goal is to generate a PDF report that is intuitive to read without worrying about the formatting under HL7 constraints.

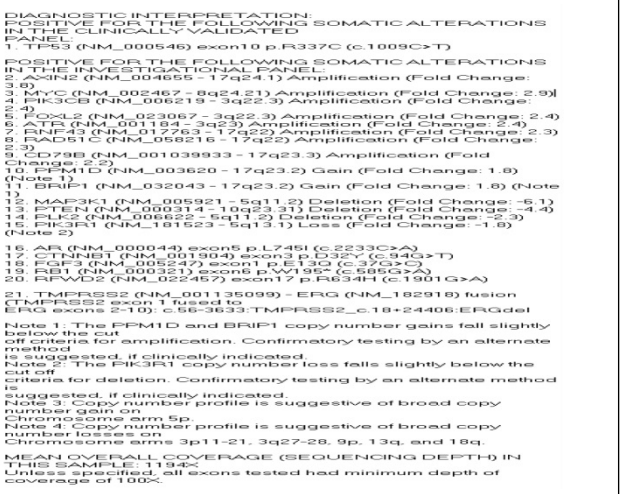
Expected barriers and resistance from institutional information system teams

Pushback may be present if only a PDF document is provided in the EMR in many academic medical centers with their own IDW. Most pathology information housed in IDWs does not come from scalable automated data queries and extractions from the AP-LIS. Most AP-LIS vendors do not have databases which are robustly constructed to handle both large data queries and transactions at the same time. Thus data queries and extractions from the AP-LIS usually occur during off hours and at a smaller scale. Our institution, like a few others, uses a scaled automated natural-language-based process to extract data from the main (final) diagnosis section of the HL7 data stream. The prerequisite is consistent structure of the pathology report across the department, such as using standardized cancer protocols. If only a PDF is provided to the EMR, then the scalable pipeline for pathology data going into IDW is lost. The advantage of our architecture is that it maintains the HL7 data stream, which can be parsed and information extracted into the IDW.

The change to the HL7 message is minimal. A pointer field is added, which houses a link to the corresponding PDF on the hospital IS server. Despite this simple change, interface testing efforts were not trivial. There were costs associated with time and human resources and coordination with project implementation timelines. Another issue was the engineering of a reconciliation dashboard to ensure that the HL7 message feeds corresponded and were correctly routed. A mismatch between the PDF and the HL7 message would be disastrous. The dashboard ensured a 1:1 relationship between HL7 messages and PDF and also provided a platform for troubleshooting issues of reports not crossing over the interface.

Enhanced reporting with NGS

Molecular ancillary studies have become an integral component of anatomic pathology, and there is a need to integrate such complex results into the anatomic pathology report. NGS results serve as prime examples of the potential dangers of information overload. Without visual cues, cascading lists of molecular results may blend together. Important information can also be buried inside the walls of narrative text. Figure 4 illustrates an example of a plain text HL7 rendering of a NGS report as displayed in an EMR. In addition to misalignments, there is considerable difficulty in being able to parse through the NGS report without the help of visual cues.


DIAGNOSTIC INTERPRETATION:
POSITIVE FOR THE FOLLOWING SOMATIC ALTERATIONS IN THE CLINICALLY VALIDATED PANEL:
1. TP53 (NM_000546) exon10 p.R337C (c.1009C>T)

POSITIVE FOR THE FOLLOWING SOMATIC ALTERATIONS IN THE INVESTIGATIONAL PANEL:
2. AXIN2 (NM_0004655 - 17q24.1) Amplification (Fold Change: 3.6)
3. MYC (NM_002467 - 8q24.21) Amplification (Fold Change: 2.9)
4. PIK3CB (NM_006219 - 3q22.3) Amplification (Fold Change: 2.4)
5. FOXL2 (NM_023067 - 3q22.3) Amplification (Fold Change: 2.4)
6. ATR (NM_001184 - 3q23) Amplification (Fold Change: 2.4)
7. RNF43 (NM_017763 - 17q22) Amplification (Fold Change: 2.3)
8. RAD51C (NM_058216 - 17q22) Amplification (Fold Change: 2.3)
9. CD79B (NM_001039933 - 17q23.3) Amplification (Fold Change: 2.2)
10. PPM1D (NM_003620 - 17q23.2) Gain (Fold Change: 1.8) (Note 1)
11. BRIP1 (NM_032043 - 17q23.2) Gain (Fold Change: 1.8) (Note 1)
12. MAP3K1 (NM_005921 - 5q11.2) Deletion (Fold Change: -6.1)
13. PTEN (NM_000314 - 10q23.31) Deletion (Fold Change: -4.4)
14. PLK2 (NM_006622 - 5q11.2) Deletion (Fold Change: -2.3)
15. PIK3R1 (NM_181523 - 5q13.1) Loss (Fold Change: -1.8) (Note 2)

16. AR (NM_000044) exon5 p.L745I (c.2233C>A)
17. CTNNB1 (NM_001904) exon3 p.D32Y (c.94G>T)
18. FGF3 (NM_005247) exon1 p.E13Q (c.37G>C)
19. RB1 (NM_000321) exon6 p.Y195* (c.585G>A)
20. RFWD2 (NM_022457) exon17 p.R634H (c.1901G>A)

21. TMPRSS2 (NM_001135099) - ERG (NM_182918) fusion
(TMPRSS2 exon 1 fused to
ERG exons 2-10): c.56-3633:TMPRSS2_c.18+24406:ERGdel

Note 1: The PPM1D and BRIP1 copy number gains fall slightly below the cut
off criteria for amplification. Confirmatory testing by an alternate method
is suggested, if clinically indicated.
Note 2: The PIK3R1 copy number loss falls slightly below the cut off
criteria for deletion. Confirmatory testing by an alternate method is
suggested, if clinically indicated.
Note 3: Copy number profile is suggestive of broad copy number gain on
Chromosome arm 5p.
Note 4: Copy number profile is suggestive of broad copy number losses on
Chromosome arms 3p11-21, 3q27-28, 9p, 13q, and 18q.

MEAN OVERALL COVERAGE (SEQUENCING DEPTH) IN THIS SAMPLE: 1194×
Unless specified, all exons tested had minimum depth of coverage of 100×.

Figure 4: Example of a plain text HL7 rendering of a NGS report as displayed in an EMR.

Clinicians may need to search for relevant information in a sea of text strings, which includes legally required documentation such disclaimers, laboratory policies, laboratory controls, and laboratory methods. This is illustrated near the bottom of Figure 4, describing laboratory controls and methods. Even if one were to get to the relevant sections of NGS data, common questions which arise include "what is the relevance?", "can the results be filtered down?", and even

"where do I start to parse the data?" Figure 5 illustrates the same NGS report in Figure 4, transformed as a PDF document. The visual cues are very appreciable with section headers and use of bolds, italics, colors, and alignments; none of which are possible through plain text HL7 renderings. We are in current efforts to enhance further our NGS reports with use of tables, charts, and images.

Summary	6 mutations, 14 copy number alterations, 1 structural variant detected
Comments	Copy number profile is suggestive of broad copy number gain on Chromosome arm 5p. Copy number profile is suggestive of broad copy number losses on Chromosome arms 3p11-21, 3q27-28, 9p, 13q, and 18q.

Somatic alterations detected in this sample:

Gene	Type	Annotation	Location	Additional Information+	
Mutations					
TP53	Missense Mutation	R337C (c.1009C>T)	exon10	MAF: 85%, COSMIC: 20	α
CTNNB1	Missense Mutation	D32Y (c.94G>T)	exon3	MAF: 40%, COSMIC: 135	
RB1	Nonsense Mutation	W195* (c.585G>A)	exon6	MAF: 70%, COSMIC: 1	
RFWD2	Missense Mutation	R634H (c.1901G>A)	exon17	MAF: 5%	
FGF3	Missense Mutation	E13Q (c.37G>C)	exon1	MAF: 18%	
AR	Missense Mutation	L745I (c.2233C>A)	exon5	MAF: 12%	
Copy Number Alterations					
MAP3K1	Whole gene	Deletion	5q11.2	FC: -6.1	
PTEN	Whole gene	Deletion	10q23.31	FC: -4.4	
PLK2	Whole gene	Deletion	5q11.2	FC: -2.3	
PIK3R1	Whole gene	Loss	5q13.1	FC: -1.8	a
PPM1D	Whole gene	Gain	17q23.2	FC: 1.8	b
BRIP1	Whole gene	Gain	17q23.2	FC: 1.8	c
CD79B	Whole gene	Amplification	17q23.3	FC: 2.2	
RNF43	Whole gene	Amplification	17q22	FC: 2.3	
RAD51C	Whole gene	Amplification	17q22	FC: 2.3	
PIK3CB	Whole gene	Amplification	3q22.3	FC: 2.4	
FOXL2	Whole gene	Amplification	3q22.3	FC: 2.4	
ATR	Whole gene	Amplification	3q23	FC: 2.4	
MYC	Whole gene	Amplification	8q24.21	FC: 2.9	

Figure 5: Example of the same NGS report in Figure 4, transformed as a PDF document.

Visual cues such as tables, charts, and images are instruments of a discipline known as visual analytics which seeks to understand how the human mind is able to manage a deluge of information [2]. Utilizing visual analytics, more intuitive layouts in pathology reports can be constructed to provide more informative and more effective reporting. Unfortunately to date, there are no established guidelines for reporting NGS information, and the idea of what determines an intuitive layout is still evolving. The application of visual analytics is necessary to cope with the constant expansion of complex "omics" data.

Despite PDF exchanges being part of the HL7 standard, implementation of PDF functionality within the EMR remains rare. Most patient portals as modular components of EMRs, also do not have functionalities to display PDF reports but rather display plain text renderings from the HL7 data stream. One can imagine confusion among patients in looking at their clinical NGS reports with lists of genes, or even worse tables in a PDF report, which did not translate over to the portal.

Beyond PDF, the pathology report as a data collection tool

The PDF display for pathology reporting is a valid near term solution, particularly as a documentation and communication tool [1]. Perhaps the next evolution in reporting beyond usage of PDFs will occur under web-based platforms allowing for more degrees of freedom in presentation and visual analytics. The emergence of initial successful attempts at web-based reporting in the literature demonstrates the potential for easier report integration between

historically isolated data sets. Arnold et el. have developed an active web-based reporting method for integrating radiology and pathology data sets into intuitive reports and have shown benefits with quality assurance [8-10]. Although their experience demonstrates one specific use case in pulmonary pathology, the integrated web-based reporting method can be adopted to use with surgical pathology, molecular, and genomic datasets. With web-based reporting, there is the added capability of incorporating hyperlinks for back end decision support. This is something currently not possible with the constraints of MS Word as the document editor for most AP-LISs. Moreover most commercial AP-LISs lack the application programming interfaces (API) to pull and integrate disparate datasets from other isolated databases to make such integrated reporting possible.

Pathology reports are gold mines for data collection. This is where usage of the PDF has its shortcomings. The information contained in PDF reports is not easily accessible for data mining. Parsing the corresponding HL7 message for data mining which contains visual, tabular, chart, and/or imaging data translated from the PDF is challenging. Data mining automation, such as optical character recognition technology, has error rates that become appreciable at scale. In addition, machine learning algorithms for interpreting and analysing images in the PDF report are in their infancy.

Progress is anticipated with more advanced HL7 standards like clinical document architecture (CDA) that can go beyond PDF. CDA is intended for transmission standards to take into account the structural components of clinical documents, which may include multiple sections, such as history, physical exam, review of systems, and assessment and plan [1]. Surgical pathology reports are similar. The reports usually have components such as the final diagnosis section, gross description section, frozen section diagnosis section, addendum section, and etc. CDA provides organizational context for how reports are structured and positioned, and standardizes the appearance of documents when rendered in the EMR [1]. The advantage over a PDF interface is that with CDA capability, the data stream or report rendered in the EMR is more structured and more amenable to processing and data. CDA functionality is likely to appear in all US commercial EMRs because this functionality is included in the requirements for Meaningful Use in the United States [1].

Conclusion

We have been successful in our implementation of PDF functionality in the EMR. Pathology reports in PDF are viewed as the pathologists intended, and thus are more intuitive. Errors in conveying information that occur through formatting issues are obviated, and additional flexibility is provided through formatting options such as colours, bolds, and fonts. Use of visual analytics is also possible to enable better conveyance of information [11]. Most commercially available EMR vendors have the capability of displaying PDF documents, and thus we hope that our experience can serve as a framework for PDF functionality in EMRs in the near term for back end clinical decision support.

References

1. Sirintrapun SJ, Artz DR (2015) Health Information Systems. Surg Pathol Clin 8: 255-268.
2. Murdoch T, Detsky A (2015) The inevitable application of big data to health care. JAMA 309: 1351–1352.
3. Roy S (2015) Molecular Pathology Informatics. Surg Pathol Clin 8: 187-194.
4. William HR (1995) The Electronic Medical Record: Promises and Problems. J Am Soc Inf Sci 46: 772-776.
5. Tracy GD, Nicolas TP (2005) The Emergence of National Electronic Health Record Architectures in the United States and Australia: Models, Costs, and Questions. 7: e3.
6. Robert DH, Beebe AL, Biron C, Boyer PV, Essin SL, et al. (2001) The HL7 Clinical Document Architecture J Am Med Inform Assoc 8: 552-569.
7. http://www.hl7.org/implement/standards/
8. Berndt DJ, Hevner AR, Stunicki J, et al. (2000) Hospital Discharge Transactions: A Data Warehouse Component. Proceedings of the 33rd Hawaii International Conference on System Sciences.
9. Ash JS, Berg M, Coiera E (2004) Some Unintended Consequences of Information Technology in Health Care: The Nature of Patient Care Information System related Errors. J Am Soc Inf Sci 11: 104-112.
10. Arnold CW, Wallace WD, Chen S, Oh A, Abtin F, et al. (2015) RadPath: A Web-based System for Integrating and Correlating Radiology and Pathology Findings During Cancer Diagnosis. Acad Radiol 23: 90-100.
11. Caban JJ, Gotz D (2015) Visual Analytics in Healthcare Opportunities and Research Challenges. J Am Med Inform Assoc 22: 260-262.