SIIM 2017 Scientific Session
Posters & Demonstrations

Four Real-World Implementations of Imaging Informatics in an Academic Radiology Medical Center

Eric Mischo, MS; Daryn Belden, BS; Elizabeth Simcock, BS, MS; Andy Craven, BS, MBA; Robert Bour, MD; Richard Bruce, MD; Charles Mistretta, PhD; Scott Reeder, MD, PhD; Charles Strother, MD; Michael Tuite, MD; Gary Wendt, MD, MBA

Background

The imaging informatics component of the Department of Radiology at the University of Wisconsin School of Medicine and Public Health (“UW–Madison Radiology”) is relatively new but rapidly maturing. UW–Madison Radiology has enthusiastically adopted the tenet that developing and sustaining a strong informatics program is essential for the department to fulfill its academic, clinical, outreach, and research missions. Our informatics team benefits from its close collaborations with leadership, faculty, trainees, and staff representing the department’s missions. We are also indebted to our partners at UW Health and relationships with other UW–Madison departments for their critical roles in our shared endeavors. We would like to take this opportunity to present some examples of how these relationships result in practical utilization of imaging informatics at UW–Madison Radiology.

Four Case Presentations

Our faculty leadership regularly ask us to create new dashboards. We will discuss three representative dashboard requests and also present a final case study on automated de-identification of data. The TimePoints dashboard must track metrics that are critical to the radiology patient experience. These include Scheduled, Arrived, Exam Begin, Exam End, Preliminary Report, and Final Report as well as derivative durations (e.g., Exam End – Exam Begin = Exam Length). The Productivity dashboard must track Volume, Professional Charges, and RVUs. The MRI Volume dashboard must track volume by Site, Patient Type, and Body Region. Because data sources for the requested dashboards are quite varied, flexible techniques of acquisition are required. While the majority of our data comes from UW Health, approximately 40% of our business is done at other sites. Thus, we deal with multiple EHRs. We also query data from our PACS and speech recognition systems as well as an HL7 repository.

Case A: TimePoints - Because leadership wants current-day metrics, we cannot base the data on overnight extracts. Further, while our HL7 repository collects up-to-the-minute HL7 data from all our sites, the database itself is not optimized for persistent querying throughout the day. Utilizing stored procedures, we query a subset of fields from the HL7 repository. This extracted data, now stored on our local database server, is combined through a series of views and pulled into a dashboard. To make the queries run more efficiently, all locally stored data older than 365 days is truncated once per day. Because the views are able to complete their work relatively quickly, our dashboard administrators have allowed us to forego the 3-tier architecture that is typical of its solutions. Database and dashboard servers coordinate their data pulls in a "leap-frog" manner so as not to interfere with one another. The TimePoints dashboard has been in operation for many months. In addition to providing leadership with an overview of patient experience metrics, it has helped us better understand our HL7 dataflow. The combination of multi-flavored SQL and dashboard servers is powerful and stable. As we encounter unforeseen anomalies in the data (e.g., studies that change modality midstream) the solution is also proving to be flexible. Rather than updating the dashboard itself, we are able to accommodate new logic directly within the SQL views.
**Figure 1**

![Figure 1](image1.png)

**Case B: Productivity** – This data is available via flat-file, but the current download process is completely manual. We are developing an automated package to pull that data into our SQL server. UW Health is also working to replace their flat-files with a direct connection to their database. In addition to automating the process of extraction, this will help us transform the data using SQL rather than a dashboard script. If our automation solution works as expected, it will pave the way for various other applications in the future.

**Figure 2**

![Figure 2](image2.png)

**Case C: MRI Volume** - The requested data is restricted to only one site and does not need to be up-to-the-minute. It is thus available through our nightly EHR extract. Two views have been providing data for a number of months. Because this data is theoretically also available through other data sources, we are taking this as an opportunity to validate queries and sources around this common request. Co-validation of our various data sources is ongoing.
Figure 3

UW Health MRI Volume Report

Monthly Volume: by Site
Apr 2016 - Mar 2017

Monthly Volume: Total
Apr 2016 - Mar 2017

Figure 3. Volume 1
Figure 4: Annual Volumes, 2012

By Patient Type

By Site


* 2017P values are found by multiplying the
Case D: Automated De-identification of Protected Health Information

One of the common challenges experienced by imaging informaticists is determining how to effectively distribute data and analyses among personnel and information systems external to the imaging department. Much of the data we work with naturally contains at least some components of Protected Health Information (PHI) at origin. We have business associate agreements and similar mechanisms in place, but wherever possible we prefer to remove PHI entirely from data we share outside the department.

The RSNA Clinical Trial Processor (CTP) [1] provides an outstanding mechanism for de-identifying DICOM studies but introduces the known possibility of stripping information vital to the workflows of external partners. [2] We work extensively with those collaborators to develop what we consider relatively sophisticated mechanisms to query and move studies directly from modalities, de-identify them with CTP and store them in PACS systems like dcm4chee [3] for our associates to work with further, stripped of PHI but retaining all other necessary information such as values contained in specific, essential private tags.

One workflow example we are particularly proud to highlight is an automated process that is instrumental in forwarding 4D Digital Subtraction Angiography research, pioneered at UW–Madison Radiology by Drs. Charles Mistretta and Charles Strother [4]. The XA modality is queried every morning for new studies that meet Dr. Strother’s specific criteria and any matching studies are C-MOVEd through...
a process that includes querying a database to determine if that patient has a previous relevant study. We have configured CTP to assign the previous patient ID number to the current study if that query results in a match, otherwise a new patient ID number is created, assigned to the study and stored in the ID database. Study IDs are simply created incrementally but they are also stored in a separate table of the ID database along with the information needed to re-identify the study if necessary. That ID database is maintained by UW–Madison Radiology informaticists and the information contained within is never shared with anyone other than Dr. Strother or his research study coordinator. We provide Weasis [5] integration into dcm4chee for those interested in using the web interface exclusively, but for the most part Dr. Strother and his collaborators interact with dcm4chee through custom tools on their workstations.

**Figure 5**

![Automated De-identification of PHI](image)

**Outcome**

The *TimePoints, Productivity, and MRI Volume* dashboards demonstrate how alphanumeric data can be harnessed to improve the clinical mission of the department. They highlight the power of channeling data from a variety of sources into common tools, allowing faculty and staff to garner insights which improve processes and outcomes. But while alphanumeric data is critical, it is not the whole picture. As highlighted above, the automated 4DXA de-identification process is one of the workflows of which we are most proud, and it is also one of our earliest forays into that approach. Based on the success of that model we have been able to create and maintain a relatively large number of similar workflows for other faculty and trainees. UW–Madison Radiology is now incorporating the automated techniques we developed to provide a de-identification service for other departments that greatly reduces the need for persistent oversight by an informaticist. We have also evolved the process to incorporate parallel processing techniques for large-scale study de-identification. We were recently able to successfully de-identify and distribute more than 23,000 CT studies in short order in the interest of advancing automated segmentation techniques in virtual colonoscopy research.
Discussion

Our imaging informatics work has led to successes, setbacks, and insight into further opportunities. While our efforts continue to grow, we would like to briefly discuss a few of the lessons we have learned up to this point.

- Perhaps the most critical lesson learned is that we need to ensure that key faculty stakeholders understand they will need to devote a fairly significant amount of their time to the projects we are tasked with developing. At least half of the work described in this poster involves heavy amounts of back–end data validation. We regularly meet with faculty members to scrutinize raw data coming from DICOM and HL7 sources.
- It is often necessary to tie together several separate applications and databases from a variety of vendors and open source solutions.
- Advanced understanding of DICOM and HL7 workflows is essential.
- We have benefitted from virtualizing nearly the entirety of our workflow, making extensive use of an enterprise–class virtual machine and container infrastructure.
- We have found that vendors often need to be able to access de–identified data.
- Our clinical and research work is closely intertwined — efficient transfer and de–identification of clinical DICOM studies to researchers is vital.

Conclusion

Informatics are everywhere. Every week our faculty are asking for new views into data, new ways to see data that can answer major questions to help improve our clinical and research work. For a long time the data did not exist or was not available in any kind of readily-accessible form. We have moved into a time where we are flooded with data, but it comes from many disconnected sources, making it difficult to assess. Through a series of projects in our institution, we are finding new ways to put the data up on the light box to diagnose clinical inefficiencies, quality control issues, and operational productivity shortcomings.

Figure 6
References

3. dcm4che.org. dcm4chee. https://dcm4che.atlassian.net/wiki/display/ee2
5. dcm4che.org. Weasis. https://dcm4che.atlassian.net/wiki/display/WEA

Keywords
anonymization; common rule; cross-disciplinary research; data sharing; de-identification; DICOM; HIPAA; image archive; imaging informatics; open access; open source; PHI; radiology; supplement 142