1. Title: TruData: The Facilitator of Clear Communication at Weill Cornell

Weill Cornell Medical College

Primary Contact: Victor Brodsky, MD
Assistant Medical Director of Informatics
Assistant Professor of Pathology, Laboratory Medicine, and Public Health
victor.brodsky@med.cornell.edu
(646) 962-2997

Tru V. Tran, B.S
Data Dictionary Architect
Author of TruData
trt2001@med.cornell.edu

Adam D. Cheriff, MD
Chief Medical Information Officer of Weill Cornell Physician Organization
Associate Professor of Clinical Medicine
212-746-4721
adc9001@med.cornell.edu

J. Travis Gossey, M.D., M.S.
Medical Director of Physician Organization Information Services
Assistant Professor of Internal Medicine
(212) 746-0417
jtg9003@med.cornell.edu

Curtis L. Cole, MD
Chief Information Officer
Associate Professor of Clinical Medicine and Public Health
(212) 746-0483
ccole@med.cornell.edu

4.1 2011 Hospital National Patient Safety Goals
Improve staff communication “X”

4.2 2010 Hospital National Patient Safety Goals
Improve the effectiveness of communication among caregivers. “X”

4.3 National Priorities Partnership
Improve the Safety of America’s Healthcare System “X”
Eliminate Waste While Ensuring the Delivery of Appropriate Care “X”

5. Health Information Technology
Electronic Health Record (EHR) “X”

6. Did you attempt to meet any Meaningful Use Goals?
Improve quality, safety, efficiency, and reduce health disparities “X”
When the over 800 faculty clinicians at Weill Cornell Medical College initially installed their EMR system over nine years ago, the clinicians were excited that patients’ lab and study results were going to be available to them electronically in one place. The ability to trend labs over time and compare historical results would allow them to spot disease trends more easily and reduce the need to repeat tests. That dream was shattered when they discovered that different resulting agencies called identical tests different names. Their EMR receives data from 19 different resulting agencies, so the possible names for a test got impossible to remember. Additionally, the system could not automatically trend results generated from the different systems. Given this problem, we searched for a program or standard vocabulary to try to match the results names to consistent ones, however none were available. So, we created a system called TruData that now automatically maps incoming test results to a standard name. The physicians are happy because they can see all of the results of a particular test without having to remember the different names that it could be called. Additionally, they can trend the data electronically regardless of resulting agency. It’s been a big win because it helps our physicians use all of the available data to its fullest. Patients love it because we can share with them trends in their results. For example, a graph of a patient’s cholesterol tells the complete story of how they’re doing at a glance. Even patients who cannot read, understand what that graph means. TruData has transformed over 12 million results and has brought order out of what once was chaos. In addition, it now automatically suggests mappings for new codes as they come into the system. Our submission details how TruData was created and how it has transformed results review at our organization.
2. Background Knowledge

Weill Medical College of Cornell University was founded in 1898, and has been affiliated with what is now New York-Presbyterian Hospital since 1927. It is among the top-ranked clinical and medical research centers in the country. Cornell offers degrees in medicine and Ph.D. programs in biomedical research and education at the Weill Graduate School of Medical Sciences. Together with Rockefeller University and the Sloan-Kettering Institute, Cornell has a joint MD-PhD program as part of its mission of education, research, and patient care.

Approximately 800 Weill Cornell faculty members serve as attendings at New York Presbyterian-Weill Cornell Medical Center (WCMC) and supervise over 600 residents in addition to numerous fellows in a variety of subspecialties. Nearly 100 different practice groups within Weill Cornell Physician Organization are supported by approximately 1500 total employees, with the majority of the practices located on the main campus on the Upper East Side of Manhattan. There are several additional off site practices around Manhattan and some of our physicians also have Cornell practices at other affiliated institutions in the outer boroughs of New York City. The faculty members see more than 780,000 outpatient visits per year, with only 15% of the visit volume made up of primary care due to the highly specialized nature of our quaternary care center. Over half of the physicians are expected to directly use our Epic Electronic Medical Record system, which is mainly focused on ambulatory services.

If left unaddressed, the problem of consolidation and alignment of results received by the EMR from multiple disparate sources compromises safety and quality in many hospitals since lack of automated display of previous result values not only adversely impacts patient care and physician experience, but may also result in unnecessary repeat orders and expenditures.

3. Local problem

At WCMC clinicians send patients to have studies at over 19 different resulting agencies (Figure 1). Unfortunately, the results return to the EMR labeled by the respective laboratories with their own, proprietary codes. Allowing the EMR blindly accept test results makes it impossible to display them in a categorized, logically grouped fashion and prevents the system to display trends. Over 97% of the result components in the EMR are resulted by one or more agency, so there is significant overlap.
Over 11,000 procedures pertaining to laboratory tests and radiology studies are available for the physicians to order across the various sources, with the number of lab tests offered by each laboratory usually numbering between 1600 and 2500. Among the 11,000 procedures nearly 2500 are distinct from each other, while the rest, although coming from a different source, map to distinct concepts. Asking the physician to sift through such tremendous variety without logically consolidating the results on the screen would present not only time management issues, but would also be detrimental to patient care quality as previous results from a different source lab could be left unnoticed.

4. Intended improvement

The intended improvement in this case was the logical organization of the incoming procedure results within the EMR across multiple source laboratories and radiology practices in order to improve patient safety and efficiency of communication. Dr. Curt Cole working closely with Tru Tran as the two champions of this project decided to tackle the problem since navigating the results within the EMR was becoming unmanageable, and finding available equivalent reference lab tests was a painstaking process. The goal was to address this specific issue with a project that would require no additional staff training, thereby improving staff communication and effectiveness of communication among caregivers who provide and receive test results later performing a qualitative study of the tangible improvement to the physician’s experience of locating and trending laboratory results. Making previous results more visible would improve the safety of America’s healthcare system, thereby reducing repeat testing and helping eliminate waste while ensuring the delivery of appropriate care.

Usually, mapping any two terminologies to each other begins with an attempt to find a common ground. For the laboratory test result components, this common ground is dominantly the LOINC standard, while SNOMED is usually used for organisms. Interface terminologies such as LOINC, SNOMED, ICD9/10, CPT and others are created by modeling truth: a group of experts periodically come together to approve and assign standard codes to new defined entities which could subsequently be referenced by others. Relying purely on such existing codes within the applicable interface terminology may appear sufficient. Such an approach however is fraught with issues. To begin with, many labs do not use standard vocabularies to label their results. Also, when compared with real world data being sent by the labs, standardized vocabularies are inevitably behind the times and do not contain codes for the newest tests. Additionally, problems arise when different source laboratories incorrectly label their test results with differing LOINC codes – after all everyone is free to assign a test result with one of nearly 61,255 LOINC codes that appears to fit best. To be fair, aside from the potential ambiguity and granularity issues, the process of determining the best fitting LOINC code is a painstaking task.
5. Planning the intervention

Tru Tran and Dr. Cole therefore chose a different path: start at the real world raw data they saw being sent in to the EMR and derive the truth from this data stream, aligning and matching the laboratory result components to each other based on the contents of the character strings and the specific ontology of each result message. Instead of monitoring the live interface, laboratory catalogue crawlers and data gatherers were built and source laboratory master data files were requested. The collected data was subsumed into a common database and the string and ontology matching engine was built: TruData was thus born.

6. HIT Dimensions Utilized

At its core function, TruData establishes and sustains a system-wide mapping space for various domains via employment of profiling engines consisting of mathematical algorithms that work in cycles to parse, iterate, infer, and process raw data string contents from disparate data sources on an ad-hoc or bulk basis. The operation is supported by a dynamic hash cloud collection within which all data fragments are atomically tagged with relevant attributes.

A laboratory result that is new to TruData is used to search for its exact or closest known matches, each ranked according to the likelihood of being equivalent to this new item. Often only one closest and correct match is identified by the system – particularly if adequate data supporting the characteristics of a test or result is given or when it is optimally inferred by the system. The mapping is then manually reviewed and approved, thereby establishing an equivalency link between result codes from two different source laboratories. By this commitment process, the system leverages the action to extend its knowledge base, essentially learning and making the system more comprehensive for subsequent tasks. LOINC is treated here just as one of the available sources to be matched. This lattice of equivalencies across codes (known as the dictionary cluster) from n different sources can be graphed in (n+1)-dimensional space, ideally producing a single point in that “DD” space, called TruPoint (Figure 2).

Figures 4 & 5 demonstrate the effect of the TruData system on result display.

7. Outcomes: (a) Nature of setting and improvement intervention

The mapped result components were then organized into a logical six level tree structure via a complex rule-based engine that has subsumed over 45,000 LOINC codes into the hierarchy to date – predominantly containing 85% of the laboratory and 97% of the radiology terms respectively. This tree
view was integrated into the Epic Electronic Medical Record to provide an optimal results review experience. The TruData database of mapping is also accessible to the interface engine which uses it to look up codes for each result coming into the EMR and if necessary replaces the codes currently unknown to the Epic EMR with the equivalent known and therefore acceptable ones.

Using the described methodology over 35,000 local result codes from various laboratories out of the known 51,000 have been assigned a LOINC code with the help of TruData and brought into the EMR as an organized result tree. Within the Epic EMR, of the approximately 9800 active result components, 87% have been assigned a LOINC code. Since selecting any of the top level branches of the tree within the EMR displays the aggregate of the available results from all of the child and grandchild branches, the results are logically grouped, aligned as per equivalency map, and are only a click away (Figure 3.)

7. Outcomes: (b) Changes in processes of care and patient outcomes associated with the intervention.

Since lab data is dynamic, TruData is able to detect attribute changes in known result ontology or string content found in the updated test or result catalogue of a given laboratory and react accordingly, remapping the equivalency connection if warranted, hence maintaining mapping integrity. A notification subsystem also delivers the announcements of the changes to subscribed downstream systems, and an acknowledgment subsystem probes the recipient systems to ensure the needed changes have taken place downstream. Over the course of 8 years of operation, TruData has detected, logged, and reacted to over 280,000 of such changes, bridging communication, and preventing tens of thousands of interface errors, phone calls and delays that would have been necessary to verify and re-file the patient test results appropriately. Clinician satisfaction was widely expressed but no questionnaires were administered.

Deploying TruData greatly improved physicians’ satisfaction level with using the EMR and removed the necessity to maintain several full time employees on staff tasked with mapping thousands of result component codes and procedure codes, which would necessarily be more prone to human error. To date it has mapped over 12 million results, and the volume of transactions continues to grow.


One of the major perils encountered during the implementation of this system was obtaining complete laboratory test and result catalogues from the involved upstream systems and making sure that the updates to these catalogues make their way into TruData. The solution ranged from careful
negotiations with the people in the organizations providing the results, to technical system workarounds by building complex Internet crawler and scraper subroutines, to analyzing the data streams themselves. Since the algorithms do not run within the EMR but instead operate as a separate standalone TruData system, only the result display within EPIC EMR needed to be adopted. Working closely with the vendor in this respect did not present any significant challenges.

10. Summary

The most important result of this project is a sustainable, automated system which algorithmically locates and suggests most likely known equivalents among the procedures and result components enabling appropriate trending (Figure 4) and classification of the patient data within the EMR, in turn improving the quality of patient care by making relevant historical results appropriately accessible to the physician without unexpected outcomes.

11. Interpretation

In order to potentially improve the system further, additional heuristic methods may be investigated to see whether it may be possible to fully automate the mapping process and eliminate the manual verification and approval step for all but the most difficult cases.

12. and 13. Conclusions and Financial Considerations

Deriving matching concepts using TruData has proven to be a robust method of mapping proprietary local terminologies without overly relying on interface terminologies as the intermediary. This in turn enabled a tremendous improvement in the presentation and accessibility of the current and historical procedure results, as seen in Figure 5.

The system reduces the need for duplicate testing, which is both a medical and financial benefit. Additionally, we were able to repurpose people to work on other projects rather than focusing full-time on result mapping. The financial considerations of such a project included internal funding, nearly full dedication of a driven experienced individual (Tru Tran), part time involvement of a physician informaticist (Curt Cole) and of several members of Tru’s team for a couple of years to create and launch the TruData system. TruData is now mostly maintained and updated by Tru himself.
Stories of Success! Leveraging HIT, Improving Quality & Safety

---CASE STUDIES PUBLICATION AUTHORIZATION---

DEADLINE DATE: Mon, April 18, 2011, 5pET
Send to: dcollins@himss.org or Fax to Pat Johnson at 312-915-9511
Note: Form must be completed by those submitting case studies
Primary author Name: ______ Victor Brodsky
Institution Name: ___________ Weill Cornell Medical College

I hereby grant to the Healthcare Information and Management Systems Society ("HIMSS") as sponsor and to the American Society for Quality ("ASQ") as co-sponsor, a non-exclusive, perpetual license to use, distribute or sell in any medium and to create derivative works from the case study in which I submit to the HIMSS Stories of Success! Leveraging HIT: Improving Quality & Safety Case Studies (known as “The Work”) including written or visual material distributed or submitted. Such use, distribution, and creation may include webinars, web broadcasting, podcasts, printed materials, and electronic/digital/computer media and other media. The Work may be edited as reasonably deemed necessary by HIMSS and ASQ in order to render suitable for use or sale. On behalf of myself, my heirs successors, and assigns, I hereby release any and all claims against HIMSS which may arise directly or indirectly from the exercise of the license granted hereunder.

Additionally, I hereby agree that if my case study is selected, I will not submit my HIMSS selected case study for a period of three months after notification of this selection.

In consideration for the above grant, HIMSS agrees to acknowledge my contribution to the Work in whatever form HIMSS uses, distributes, or creates for the Work.

I warrant the following:
1. That I hold all rights to this Work, unless I created the Work in my role as an employee of the Federal government.

2. That I have obtained all necessary clearances, have cited all sources and/or included all necessary acknowledgements. Any claims or expenses caused by failure to do so will be paid by me.

IF THIS FORM IS NOT RECEIVED AT HIMSS HEADQUARTERS BY October 22, 2010 I UNDERSTAND THAT I MAY NOT BE ABLE TO BE CONSIDERED AS A CASE STUDY CANDIDATE FOR SELECTION.

Name (please print) Stories of Success! Leveraging HIT, Improving Quality & Safety