

# Expressing Observations from Electronic Medical Record Flowsheets in an i2b2 based Clinical Data Repository to Support Research and Quality Improvement

Lemuel R. Waitman<sup>1</sup>, Judith J. Warren<sup>2</sup>, E. LaVerne Manos<sup>2</sup>, Daniel W. Connolly<sup>1</sup>

<sup>1</sup>Division of Medical Informatics, Department of Biostatistics, University of Kansas Medical Center, Kansas City, KS; <sup>2</sup>School of Nursing, University of Kansas Medical Center, Kansas City, KS

## Abstract

*While nursing documentation in electronic medical record (EMR) flowsheets may represent the largest investment of clinician time with information systems, organizations lack tools to visualize and repurpose this data for research and quality improvement. Incorporating flowsheet documentation into a clinical data repository and methods to reduce the flowsheet ontology's redundancy are described. 411 million flowsheet observations, derived from an EMR predominantly used in inpatient, outpatient oncology, and emergency room settings, were incorporated into a repository using the i2b2 framework. The local flowsheet ontology contained 720 "templates" employing 5,379 groups (2,678 distinct), 37,836 measures (13,659 distinct) containing 226,666 choices for a total size of 270,641. Aggressive pruning and clustering resulted in 150 templates, 743 groups (615 distinct), 6,950 measures (4,066 distinct) with 22,497 choices, and size of 30,371. Making nursing data accessible within i2b2 provides a new perspective for contributing clinical organizations and heightens collaboration between the academic and clinical activities.*

## Introduction/Background

Strategically, a clinical data repository or data warehouse integrates information "siloes" in disparate information systems. Technically, a repository allows information to be stored in database schema tailored for retrieval and visualization while the original source systems are typically optimized for efficient and reliable transaction processing, storage, and communication. The NIH has funded the Informatics for Integrating Biology and the Bedside (i2b2) National Center for Biomedical Computing to provide an open-source framework for informatics tools that facilitate clinical data reuse and integration<sup>1</sup>. A central focus is integrating clinical and genomic data for personalized medicine and targeted therapies research. In 2010, The University of Kansas Medical Center (KUMC) established an i2b2 based Healthcare Enterprise Repository for Ontological Narration (HERON) after establishing a master data sharing agreement and oversight process with its partner clinical organizations. Similar to other i2b2 adopters, KUMC started by incorporating patient demographics, diagnoses, laboratory results, and medication data. The focus of this paper is to enrich the repository with detailed clinical data observed predominantly by nurses. Our motivation was two-fold. First, fundamental clinical parameters such as height, weight, and vital signs are documented in nursing flowsheets<sup>2</sup> and common inclusion criteria for research protocols. Having this information in i2b2 would lead to more realistic cohort identification for clinical trials, expose a vast data source for hypothesis generation, and provide a foundation for vendor neutral inter-institutional research<sup>3</sup>. Second, our clinical partners informed us of the challenges with maintaining the flowsheet ontology, validating proper utilization of terminology, and measuring compliance in nursing documentation. Since the KUMC School of Nursing has well established research programs and informatics expertise in terminologies, this was a natural point of collaboration between the academic and service organizations.

KUMC's affiliated clinical organizations adopted the Epic Systems Corporation ("Epic") as its electronic medical record. The deployment of Epic began with in the hospital in 2007 and has expanded over time to encompass the emergency department, outpatient cancer center treatment, and is currently being deployed across all ambulatory clinics. In a sample flowsheet (Figure 1) for a hypothetical medical-surgical unit, flowsheet data entry is organized

into a set of *templates* for documenting “Vital Signs”, “Assessment”, “Intake/Output”, “IV Assessment”, and “Daily Cares/Safety”. Within the “Vital Signs” template there are several *groups* for documenting related observations: Vital Signs”, “Pain Assessment”, “Oxygen Therapy”, “Height/Weight”, “Patient Observation” and “OTHER”. Each group contains individual *measures* that are observed and documented such as “BP” (patient’s blood pressure), and “Heart Rate”. The flowsheet is analogous to a spreadsheet with observation time along the horizontal axis and the individual measures listed vertically for the chosen template and organized by group. While many measures are numeric (e.g. heart rate, respiratory rate), the flowsheet user interface allows documentation of discrete observations via lists (temp source, heart rate source) and also entry of free text. The terminal finding may be a numeric value associated with a measure (e.g. the patient’s temperature was 38.2 degrees Celsius) or selection(s) from a list of *choices* (e.g. the temperature source was tympanic versus oral or rectal).

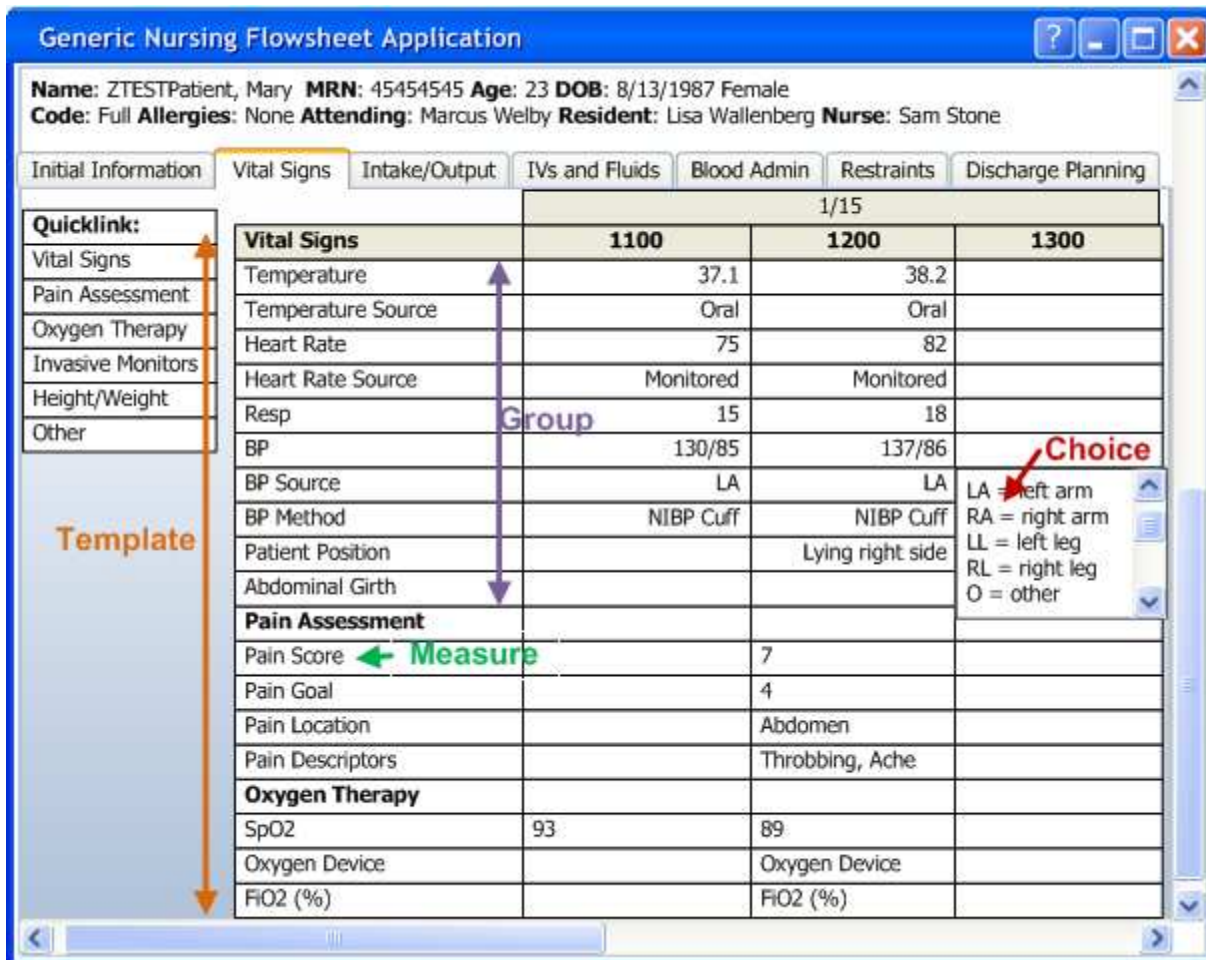


Figure 1: Vital Signs Template of a Flowsheet within the Electronic Medical Record: a. denotes the template (includes vital signs, pain assessment, etc.), b. group (includes temp, temp source, etc), and c. an individual flowsheet measure (BP – blood pressure).

As shown below (Figure 2), we use this hierarchical relationship to organize the local ontology in the i2b2 framework. The flowsheet ontology appears after peer ontologies for Demographics and Diagnoses in the upper left. We have assigned a higher level categorization based on the name of the templates. Within the “KU IP” (inpatient templates) category, the “Airways” *template* (internal template ID #171 is shown after then name) has been expanded. List choices 1 through 10 (e.g. “Dry and Intact” through “Other (comment)”) are included in the current flowsheet content ontology within Epic and the two terms, “~C” and “~DRYAND INTACT” are residual from an initial content build that was briefly used but retired.

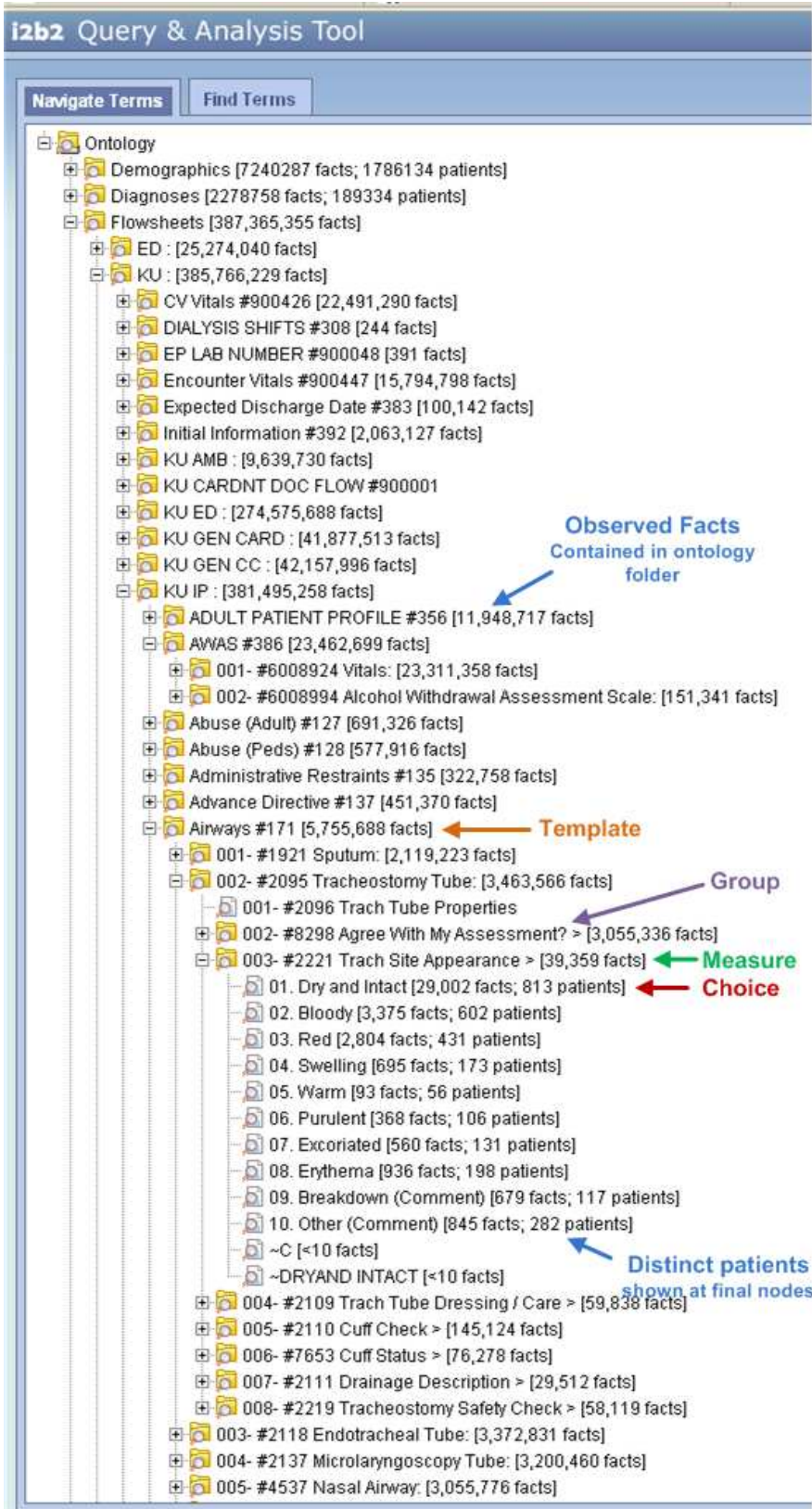


Figure 2. Flowsheet data represented in i2b2 illustrating an “Airway” *Template*, “Tracheostomy Tube” *Group*, “Trach Site Appearance” Flowsheet *Measure*, and the *Choices* for the Measure (“Dry and Intact”, “Bloody”, etc.)

There are challenges with directly translating flowsheet documentation into i2b2 relative to other clinical observations. Most hospitals record diagnoses in adherence to national or international standard terminologies such as International Statistical Classification of Diseases and Related Health Problems (ICD9, ICD10) or Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT®). While clinicians may express dissatisfaction with the loss in expressivity, ICD9 drives reimbursement and as a byproduct, research obtains a common ontology, visualizations, and ultimately information exchange<sup>3</sup>. Similarly, organizations and their ancillary functions are adopting terminologies for results such as Logical Observation Identifiers Names and Codes (LOINC®)<sup>4</sup> due to Meaningful Use requirements<sup>5,6</sup>. Research can benefit from the required mapping investment made by the clinical organizations. In the case of medications, terminologies are often adopted based for pharmacy formulary management requirements or to support clinical decision support (drug-drug, drug-allergy interactions) which are commonly provided by commercial content vendors such as First DataBank's© National Drug Data File (NDDF™), Wolters Kluwer's Medi-Span®, and Cerner® Multum. Research can derive organizing hierarchies from these commercial frameworks or map them to national standards such as RxNorm<sup>7</sup> when data sharing is required or valuable visualizations or other informatics tools are provided based upon the national terminology. All three examples share two key common characteristics: they are driven by external requirements that encourage standardizing information and they largely position a concept once in the ontology. For example, a patient with retinal damage caused by complications of diabetes mellitus, is diagnosed as having “diabetes with ophthalmic magnifications” (ICD9 code =250.6), a specialization of “diabetes mellitus” (250), a disease of other endocrine glands (249-259), and in the highest classification “Endocrine, nutritional and metabolic diseases, and immunity disorders” (240–279).

In contrast, flowsheet documentation is relatively unconstrained and costly manual abstraction from the EMR or paper chart is the norm for fulfilling external requirements. Additionally, local ontologies are customized and replicated for specific patient populations to speed data entry and assist clinical cognition. For example, head circumference would be displayed in a neonatal intensive care unit but inappropriate in an adult medical-surgical unit. However, both flowsheets would record weight, heart rate, and urine output. Figure 3 below provides a visual comparison of the similarity between vital sign templates within the outpatient cancer center flowsheet and obstetrics. These unit level redundancies and customizations increase complexity, which hinders clinical organizations using the data for quality improvement, and the terminology managers charged with harmonizing concepts. If methods could group similar measures together it would aid researchers and providers unfamiliar with each individual unit's practice. Unfortunately, the ontology's large size and redundancy of similar measures made manual grouping cost prohibitive and motivated a technique for automated ontology pruning and clustering. This method provides complimentary data preprocessing for subsequent expert guided methods for achieving organizational consensus and aligning concepts with national standards<sup>8,9,10</sup>. Our methods will describe the process and limitations of loading flowsheet documentation within the i2b2 framework, adding descriptive statistics within i2b2 for preliminary validation of flowsheet data, and develop the pruning/clustering method.

## Methods

*ETL Process:* Data from the Epic EMR were obtained by exporting the Clarity database (Spring 2008 release). The Clarity module transforms data from Epic's operational database (Intersystems Caché®) into a relational form (Oracle® 10g Release 2) for reporting. Clarity stores patient and system configuration information in over 7,000 tables with over 60,000 columns. Extracted data is transformed into an i2b2 compatible star schema, de-identified, and loaded on a separate database server to be accessed by the i2b2 application. These Extract, Transform, and Load (ETL) processes are written in Structured Query Language (SQL) statements and the Python programming language. The resulting data used by investigators and analyzed in this study is deemed non-human subjects research by the KUMC institutional review board but data use requires approval by the HERON Data Request Oversight Committee composed of representatives from KUMC and participating clinical organizations.

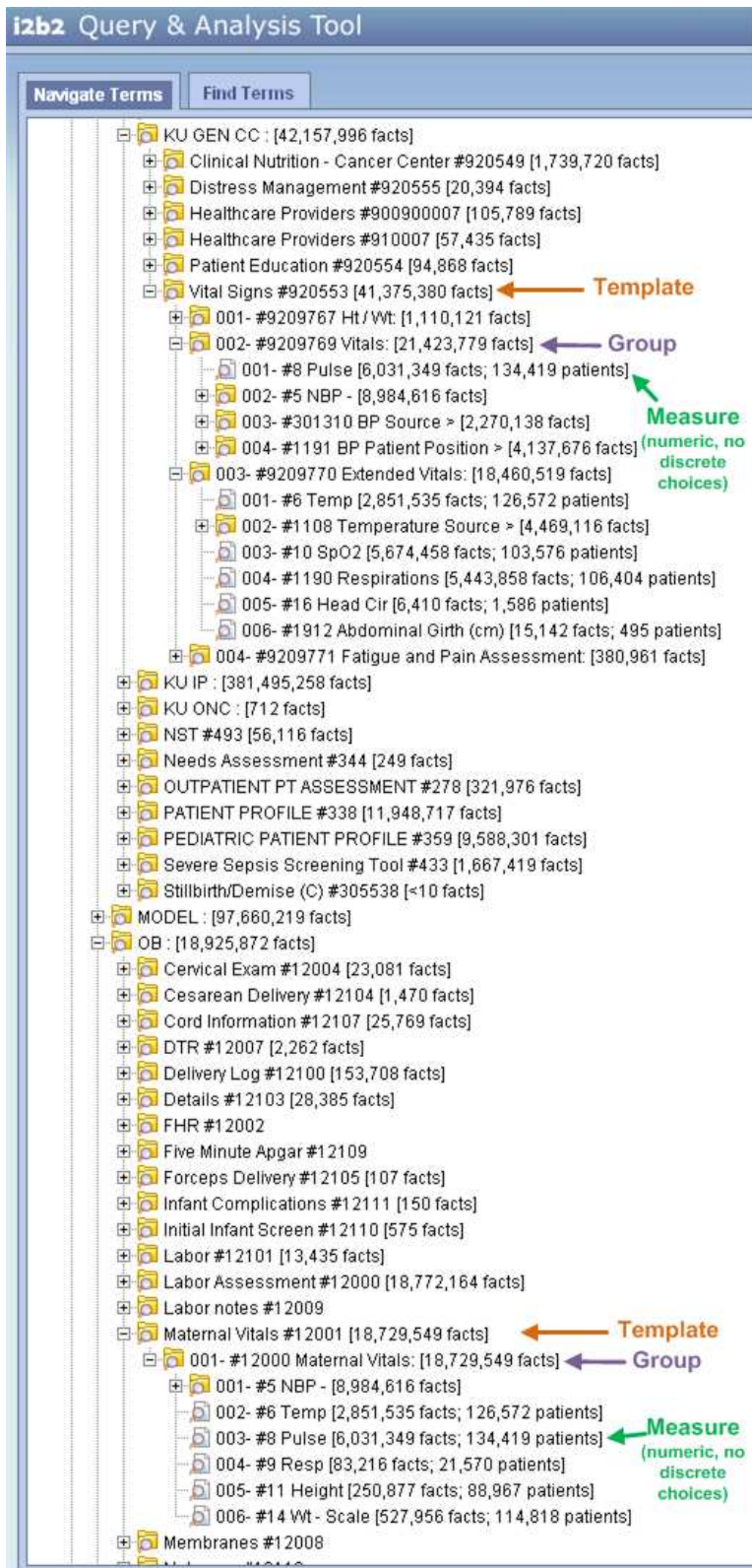


Figure 3. Similarity of flowsheet measures within the Cancer Center’s Vital Signs Template (#920553), Vitals and Extended Vitals Groups (#9209769, #9209770) and the Obstetrics Maternal Vitals Template (#12001), Maternal Vitals Group (#12000).

While similar to ETL processes for business domains, there are unique characteristics which may be analogous with other vendor's EMRs. To preserve privacy, string data types were only stored as a binary observations, not the actual content of the field (e.g. "Emergency Contact" may reveal the identity of a patient's family member). Several data types units of measure are implicit (weight in ounces, height in inches, temperature in Fahrenheit, blood pressure stored as systolic/diastolic in a single measure). Measures documented as structured choices allowing multiple responses (e.g. "Dressing Status" is "clean dry and intact" and "reinforced") are stored as a concatenated string that must be separated into individual observations within i2b2. Local system configurations contained in Clarity also include sample, model, and test content not routinely used in clinical practice.

A limitation with the Spring 2008 Clarity release is attributing a measurement to its parent group and template. The stored flowsheet measure database record within Clarity does not include fields for the group and template used to create the observation though that knowledge is preserved in Epic's operational database. For example, groups are defined for different types of catheters (e.g. single, dual, and triple lumen catheters, peripheral iv, central line). The same measure concepts are used to document "Line Status", "Site Assessment", "Phlebitis Scale", "Infiltration Scale", "Dressing Status", and "Line Care" within these groups. As a result, analysis can be done regarding all lines but not to specifically compare one type of catheter versus another. Similarly, the measure "#8298 Agree With My Assessment?", a response which indicates the nurse agrees with the previously charted values for a group, occurs across 136 disparate assessment groups, reducing its analytical utility.

*Descriptive Statistics:* Using the de-identified database, the number of patients, observations, and frequency of observations are calculated to validate the data against current practice. Additionally (Figure 2, annotations in blue), statistics are added to each node in the i2b2 metadata. This allows researchers, informatics, and clinical organizations' personnel to see how many observations have been made at each level of the ontology. Due to the large size of the flowsheet ontology and data, observed "facts" are computed at each node but distinct "patients" are only computed at the terminal measure or choice. Finally, the size of the local ontology is calculated and compared with the ICD9 ontology provided by the i2b2 National Center for Biomedical Computing.

*Pruning and clustering method:* We applied six steps to reduce and consolidate the local flowsheet ontology. Since we were not classifying the local measures into an existing ontology and had no assumptions regarding the appropriate number of groups or structure, cluster analysis was appropriate<sup>11</sup>. Groups and templates could be clustered based upon whether they shared the same measures or groups respectively. Applying an agglomerative hierarchical clustering methods (versus divisive or nonhierarchical e.g. k-means) seemed most appropriate for creating hierarchical ontologies typically provided within the i2b2 environment. The pruning and similarity scoring are written in SQL with Python for the iterative clustering algorithm. First, examples provided by the vendor or which clearly represent spurious test builds were removed. Second, we applied an arbitrary threshold that a measure or choice must occur at least once a month to be included (35 months were contained in our data: November 2007 to September 2010). Working up the ontology, any choice, measure, group or template that contained fewer than 35 observations was pruned. Third, we used the following equation to calculate a similarity score<sup>11</sup>,  $s(G_i, G_j)$ , between two groups using binary variables to indicate the presence or absence of measures:

$$s(G_i, G_j) = \frac{2a}{2a + b + c}$$

where  $a$  is the number of measures present in both groups,  $b$  is the number of measures contained in the first group ( $G_i$ ) but not the second group ( $G_j$ ), and  $c$  is the number of measures contained in the second group but not the first.

Fourth, groups more similar than the threshold are iteratively merged into new generalized groups using an agglomerative hierarchical clustering algorithm<sup>12</sup> of the following form:

START: Groups  $G_1, G_2, \dots, G_n$  each containing a set of items  $I_{11}, I_{12}, I_{21}, I_{22}, \dots, I_{nm}$ .  
 And a threshold for similarity,  $t$ , where  $0 \leq t \leq 1$ .

1. Find nearest (i.e. most similar) pair of distinct groups,  $G_i$  and  $G_j$ , merge  $G_i$  and  $G_j$  into a group containing a union of their items.

If  $s(G_i, G_j) < t$  then stop, else return to 1.

Fifth, the same score is applied to compare templates relative to the presence or absence of the generalized groups created in the fourth step. Finally, the similar templates are merged using the algorithm from the fourth step. The naming convention for the generalized groups and templates is the most frequent words included in the names of the merged groups or templates. To aid interpretation, graphs are generated that diagram the clustering process. We present the impact of this process on the ontology for similarity thresholds of 0.9, 0.5, 0.25, and 0.1 in the results.

## Results

265,406,409 flowsheet observations were loaded from Clarity. After expansion of systolic/diastolic blood pressures and concatenated choices, 411,349,435 facts for 146,970 distinct patients were available within i2b2. While the majority of the observations were for inpatients, flowsheet documentation also occurred for 56,074 patients treated in the Emergency Room, 7,661 in the outpatient cancer center, and over 46,000 in perioperative, procedural and ambulatory settings. Of the 6,742 distinct measures contained in Clarity, the most frequently observed was pulse (6,065,857 facts) while 411 measures were observed only once. The majority of clinical observations are made with a moderate number of concepts (Figure 4) and approximately 100 measures accounting for half of the observations.

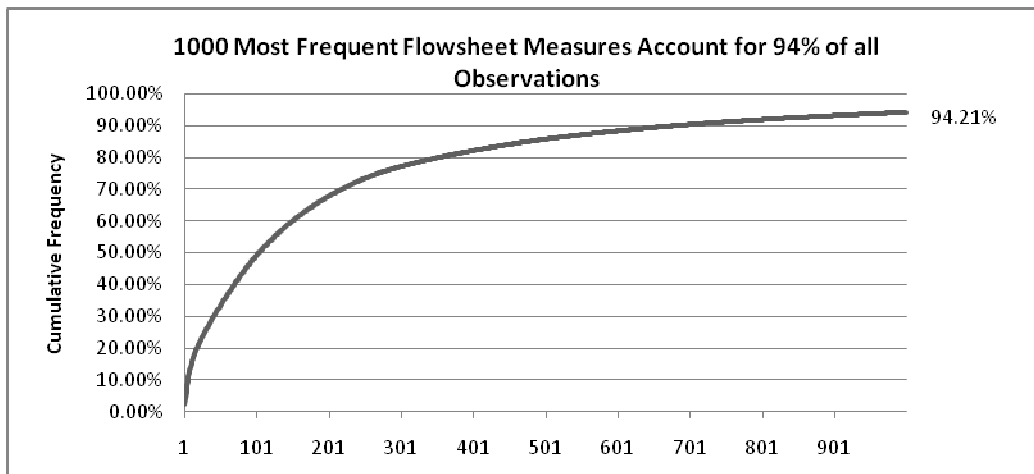


Figure 4. Cumulative frequency diagram for the 1000 most frequently documented flowsheet measures

The density of flowsheet observations per patient encounter by department was consistent with clinical practice (Surgical Intensive Care 14,150; Emergency Department 63) as was the finding that measures recorded in intensive care had a higher documentation burden (facts/encounter by individual measure concept) than general measures required once a shift or admission (NICU Isolette® temperature 224; central venous pressure 111; pain goal 8; height 1.3).

Portions of the clustering dendrograms for groups (Figure 5) and templates (Figure 6) are shown below. Nodes display the description, ‘T’ for template or ‘G’ for group, ‘#’ followed by the template or group identifier, and ‘>’ followed by the number of items contained within the group or template. Arcs indicate the similarity score between the two merged sets. If one group/template is a superset of a second, the diagram illustrates the smaller being subsumed by the larger set.

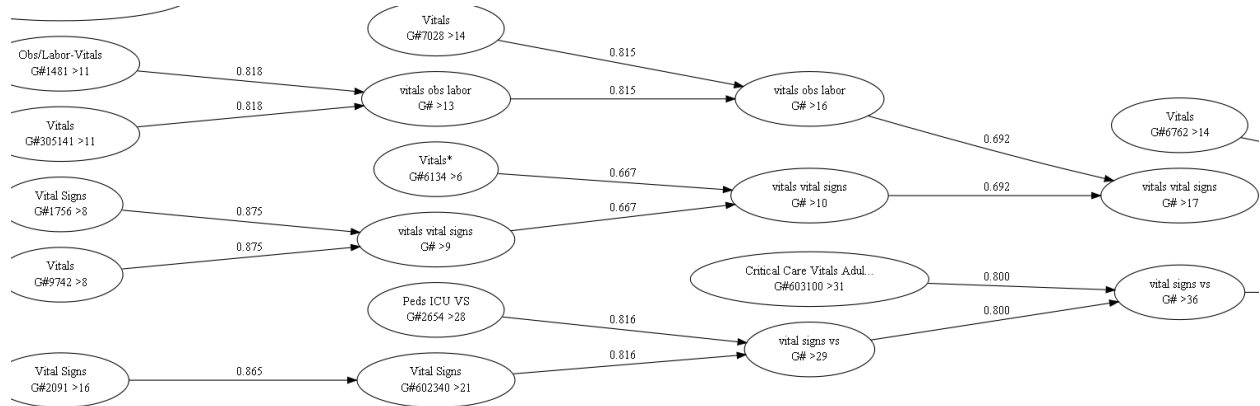


Figure 5. Portion of the clustering dendrogram at 0.5 similarity threshold illustrating group mergers of vital signs.

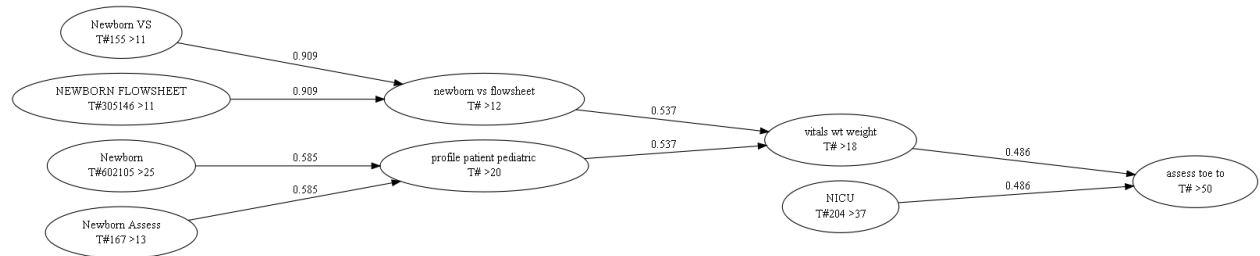


Figure 6: Portion of the clustering dendrogram at 0.5 similarity threshold illustrating merger of pediatric templates.

All phases of the pruning and clustering algorithm had consistent reductions in the size of the local ontology as the similarity score threshold was lowered (Table 1). Removing spurious builds and undocumented/infrequent measures reduced the number of distinct measures used by the ontology from 13,659 to 4,066 which dramatically reduced measure choices. While not directly comparable because of variable depth, ICD9-based diagnosis ontology provided by the i2b2 Center contained 44,577 total concepts, 158 high level concepts, and 20,116 ICD9 codes (8,231 were used locally).

Table 1. Pruning and clustering method’s impact on ontology size for different similarity score thresholds.

Pruning/Clustering Step	Total concepts	Templates	Groups (distinct)	Measures	Measure choices
Original	270,641	720	5,379 (2,292)	37,656	226,666
Actual Build	221,531	542	4,070 (1,655)	30,604	186,151
Remove < 35	121,841	428	3,350 (1,325)	21,983	95,916
Threshold = 0.9 (G and T)	102,285	390	3,008 (1,026)	19,937	78,818
0.5 (G and T)	50,146	266	1,359 (799)	10,699	37,751
0.25 (G and T)	36,594	193	953 (687)	8,228	27,169
0.1 (G and T)	30,371	150	743 (615)	6,950	22,497



## Discussion

While flowsheets and their grouping of observations are well established constructs for acquiring clinical information, they are neglected as a source for clinical data repositories because the data until recently existed on paper flowsheets rarely repurposed beyond the bedside. Our work comprehensively incorporates such data and its accompanying ontology within a widely used clinical data repository framework. Our analysis confirmed clinical information system content management teams who are customer focused tend to create flowsheet template and grouping customizations for each clinical activity or unit. Such customizations play a critical role in decision making at the bedside but without methods to curate ontologies across organizations, this can significantly compromise reuse.

Even when flowsheet observation concepts are standardized, unit based visualizations are not tailored to general clinical research and quality improvement. Informatics methods can aid navigation within these ontologies. By developing our methods to only manipulate the i2b2 metadata schema, researchers can visually compare the original ontology against multiple pruned ontologies using different similarity measures. This approach does not require any manipulation to the underlying facts and allows iterative refinement. Anecdotally, the ICD9 ontology may be daunting for clinical researchers so it is encouraging that at moderate similarity thresholds ( $t=0.5$ ), the flowsheet ontology is smaller than the ICD9-based diagnosis ontology provided by the i2b2 National Center for Biomedical Computing. Our pruning and clustering method applies well established techniques which can be improved by altering the similarity score, clustering algorithm, and naming conventions. More importantly, technical development to reduce the tedium of merging similar groupings should be coupled with nursing informatics terminology development, feedback from quality improvement and general clinical research users, and evaluation. Expert review is likely required to validate and name the merged groups and templates. Future work will synthesize these algorithmic techniques with expert review to identify identical terms and harmonize local ontologies with national standards.

While we pruned vendor sample and model ontologies for our analysis, variability of a local ontology from the vendor's model ontology is an area for future implementation science research. Clinical information system providers might assist their clients when local ontologies require standardization in order to repurpose data for external requirements (e.g. meaningful use). The ability to support a client who has deviated from a model system designed to satisfy federal reporting requirements could be invaluable to hospitals that lack informatics expertise. A cursory review indicated that the vendor's SAMPLE and MODEL ontologies cover approximately 15 and 30 percent of the observed facts from our clinical organizations.

Incorporating flowsheet data in the repository also provides a foundation for nursing and informatics research. With updated releases of Clarity we can evaluate clinical documentation authors versus information consumers<sup>13</sup> and how their ontologies may differ. The needs of consumers can also be used to evaluate what data is most critical for clinical decision making, regulatory compliance, and non-clinical activities. As a worst case, flowsheet templates might include concepts which are no longer needed for clinical decision making or satisfying regulatory requirements. Our descriptive statistic for documentation burden (facts/encounter by individual measure concept) is a start to understanding how documentation requirements consume clinical resources. Improved measures, (e.g. incorporating length of stay) should be developed so organizations can evaluate how documenting new practices such as intensive insulin therapy<sup>14</sup> impacts nursing workload.

*Limitations:* This work currently represents clinical activities associated with one medical center and is primarily inpatient focused. Widespread deployment of the EMR throughout ambulatory settings commenced in Fall 2010 and is scheduled to complete in 2012. The lost relationship between an observed measure concept and its parent group and template limits the analytical utility of many observations and has been observed by the authors in other commercial and homegrown documentation systems. Reconstructing the proper attribution of a measurement to its parent group and template will be evaluated in future work after incorporating a newer release of the EMR relational

database. Similarly, robust content management systems for managing and versioning clinical content builds remain an ongoing area for improvement among commercial and nonprofit EMR developers. Providing good history of when concepts such as list choices were retired or deprecated to new concepts is a significant task when developing complex content driven systems and satisfying secondary consumers of this information.

## Conclusion

While challenging, adding flowsheet documentation to a clinical data repository provides essential patient data for clinical research and new insight into care processes. Because of the large volume of data and complex flowsheet ontologies, organizations that wish to realize the full return on their investment in electronic documentation will require increasingly sophisticated informatics techniques to efficiently organize, navigate, and share information derived from flowsheets.

## References

1. Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, Kohane I. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc.* 2010 Mar-Apr;17(2):124-30.
2. Hammond J, Johnson HM, Varas R, Ward CG. A qualitative comparison of paper flowsheets vs a computer-based clinical information system. *Chest.* 1991 Jan;99(1):155-7.
3. Weber GM, Murphy SN, McMurry AJ, Macfadden D, Nigrin DJ, Churchill S, Kohane IS. The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. *J Am Med Inform Assoc.* 2009 Sep-Oct;16(5):624-30.
4. Huff SM, Rocha RA, McDonald CJ, et al. Development of the Logical Observation Identifier Names and Codes (LOINC) vocabulary. *J Am Med Inform Assoc.* 1998 May-Jun;5(3):276-92.
5. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. *N Engl J Med.* 2010 Aug 5;363(6):501-4. Epub 2010 Jul 13.
6. Medicare and Medicaid Programs; Electronic Health Record Incentive Program Final Rule. *Federal Register / Vol. 75, No. 144 / Wednesday, July 28, 2010 / Rules and Regulations; 44347.* <http://edocket.access.gpo.gov/2010/pdf/2010-17207.pdf> [Accessed on March 16, 2011]
7. Liu S, Wei M, Moore R, Ganesan VAGV, Nelson SANS. RxNorm: prescription for electronic drug information exchange. *IT Professional* 2005;7(5):17-23.
8. Ozbolt JG, Russo M, Stultz MP. Validity and reliability of standard terms and codes for patient care data. *Proc Annu Symp Comput Appl Med Care.* 1995:37-41.
9. Kim H, Harris MR, Savova GK, Chute CG. The first step toward data reuse: disambiguating concept representation of the locally developed ICU nursing flowsheets. *Comput Inform Nurs.* 2008 Sep-Oct;26(5):282-9.
10. Kim H, Harris MR, Savova G, Chute CG. Content coverage of SNOMED-CT toward the ICU nursing flowsheets and the acuity indicators. *Stud Health Technol Inform.* 2006;122:722-6.
11. Johnson RA, Wichern DW. 1992. Clustering. *In Applied Multivariate Statistical Analysis.* 3 ed. Prentice Hall, Inc., Upper Saddle River, NJ; 578-9.
12. Everitt BS. 1993. Cluster analysis. 3rd ed ed. E. Arnold, London; 56-66.
13. Champion TR Jr, Denny JC, Weinberg ST, Lorenzi NM, Waitman LR. Analysis of a computerized sign-out tool: identification of unanticipated uses and contradictory content. *AMIA Annu Symp Proc.* 2007 Oct 11:99-104.
14. Champion TR Jr, Waitman LR, May AK, Ozdas A, Lorenzi NM, Gadd CS. Social, organizational, and contextual characteristics of clinical decision support systems for intensive insulin therapy: a literature review and case study. *Int J Med Inform.* 2010 Jan;79(1):31-43. Epub 2009 Oct 7. Review.