Development of a Comprehensive Medical Error Ontology

Pallavi Mokkarala, MS; Julie Brixey, RN, PhD; Todd R. Johnson, PhD; Vimla L. Patel, PhD; Jiajie Zhang, PhD; James P. Turley, RN, PhD

Abstract

A critical step towards reducing errors in health care is the collection and assessment of medical error data so that potential harms to patients can be identified and steps taken to prevent or mitigate them. However, no standardized framework for classifying and evaluating such data currently exists. This paper describes our efforts in developing a comprehensive medical error ontology to serve as a standard representation for medical error concepts from various existing published taxonomies. Eight candidate taxonomies were selected from the published literature and merged to create a reference ontology consisting of 12 multidimensional axes that encompass all the aspects of a medical error event. The ultimate goal of the project is to use the medical error ontology to identify strategies for preventing future adverse events in health care.

Introduction

The study and reduction of medical errors has become a major concern in health care today. In its report, *To Err Is Human: Building a Safer Health System*, the Institute of Medicine (IOM) estimated that between 44,000 and 98,000 Americans die each year because of preventable medical errors, making hospital errors between the fifth and eighth leading causes of death.[1] In order to identify medical errors and develop strategies to prevent or mitigate them, it is essential to develop reporting systems for the collection, analysis, interpretation, and sharing of medical error data.

A number of proprietary reporting systems are currently available or under development to collect and evaluate medical error data.[2, 3, 4, 5, 6, 7, 8, 9, 10] However, no current single standardized nomenclature or universal classification system is broad enough in coverage yet sufficiently detailed to encode medical error data from all health care delivery settings and application areas.[11, 12] While most of the existing classification schemes may be well suited to the particular clinical setting or safety applications for which they were developed (e.g., medication error reporting,[10] primary care error taxonomy[8]), such application-specific and often organization-specific representations limit the reuse of medical error data.[11] It is seldom possible to map terminologies of the different classification systems to each other because of differences in granularity (e.g., the NCC-MERP Taxonomy of Medication Errors has a very detailed classification of product labeling issues as a cause of error that is not matched by the codes in DoctorQuality Inc.'s RPM system) or asymmetries in classification (i.e., assigning terms under different categories in different terminologies). Additionally, many of the existing terminologies have been developed using ad hoc approaches. More complex semantic relations among terms often are lacking. Finally, such ad hoc development can also lead to incomplete, inconsistent, and/or redundant terminological representations.[11] It is clear that the lack of a framework and agreement about how to define and classify adverse patient safety events, medical errors, and systems failures is a major barrier to understanding where, how, and why problems occur. A major goal of this project was to find a standard, generalizable mechanism to bring consistency to the published literature addressing medical errors and near-miss accounts.

An ontology of medical errors is one approach to solving the problem. An ontology can be defined as a specification of a conceptualization.[13] An ontology defines a common and controlled vocabulary that enables knowledge sharing and reuse of information.[14] Many disciplines now develop standardized ontologies so that domain experts can share and annotate information in their fields. Medicine, for example, has produced SNOMED[15] and UMLS.[16] In the context of medical errors, a medical error ontology can provide formal definitions and coverage of various concepts relevant to the domain of medical errors; enable building a knowledge base through unification of information from data across various organizations, practice domains, and applications; and offer sufficient detail to be of practical use.

As part of a larger research project designed to develop a cognitive framework of medical errors, we have begun to develop a comprehensive medical error ontology to serve as a standard representation of medical error-related concepts and relations. The first goal is to create the ontology for a standardized mechanism for coding the published literature on medical errors and near-miss accounts. We believe that the ontology would also be useful in error reporting systems and medical error and near-miss classification systems. However, the original goal is focused on the codification of the published literature. Once the ontology development is complete, it will be used in a parent project for enabling collection, aggregation, comparison, and analysis of medical error events across varied health care settings and medical error domains. The ultimate goal of the project is to identify strategies for improvement to prevent future adverse events.

This paper reports the initial progress in developing the medical error ontology. In the sections that follow, the paper presents: (1) the scope and requirements for the ontology, (2) the development approach, (3) the steps in the development process we have completed, and (4) plans for future work. In particular, we focus on the integration-based development approach we adopted, which involves reusing knowledge from existing relevant taxonomies in published literature by merging individual taxonomies to create a reference ontology, followed by mapping the reference ontology with the original taxonomies to produce the end product, a shared medical error ontology.

For the purpose of this paper, our discussions are restricted to the steps addressing the creation of the reference ontology. Other aspects of the project will be described in future articles. This will include the relations among the merged taxonomies, extensions to the ontology as it is adapted for other purposes, and the use of the ontology for a real-time medical error reporting system.

Ontology Scope and Requirements

The scope and requirements for the medical error reference ontology resulted from a series of discussions among the project members and the goals of the greater project. The first goal of the ontology was to code and represent the reported literature of medical errors and near-miss events. In order to be useful for data analysis and for learning from errors and to promote development of interventions for patient safety, it was determined that the ontology should have the following attributes:

- Cover the full range of settings in which health care can take place (e.g., not limited to hospital care).
- Capture the richness of the domain of errors and adverse events.
- Enable the capture of data from all sources (including event reports and sources, such as drug use data).
- Permit identification and analysis of medical error events.
- Enable analyses that support identification of strategies for improvement by all users (including health care providers, policymakers, and others).

Apart from the above requirements, the project team agreed that our ontology should incorporate the following suggestions from Cimino's desiderata for controlled medical vocabularies:[17]

1. **Concept orientation.** Each concept in the ontology should have a single coherent meaning; i.e., avoid redundancy, ambiguity, and vagueness in a concept.
2. **Formal definition of concepts in the ontology.** Expressed as a collection of relations to other concepts.
3. **Polyhierarchy.** A concept can belong to more than one location in the hierarchy.
4. **Multiple granularities of concepts.** Expression of concepts at different levels of detail.
5. **Graceful evolution.** To enable integration of new concepts from other, independently developed ontologies of medical errors.

Development Approach

A primary decision was that this ontology should "reuse" knowledge from existing, published medical error classifications. Reuse of independently developed, heterogeneous taxonomies together in a single system requires an integration-based ontology development approach. There are three main approaches for ontology integration:

1. **Single ontology approach.** Merging existing ontologies to create a single coherent ontology.[18, 19, 20] Often, the original ontologies cover similar or overlapping domains.

2. **Multiple ontologies approach.** Aligning existing ontologies and establishing links (mappings) amongst them for exchange of information, while preserving their original states.[18, 19, 20]

3. **Hybrid approach.**[21] Incorporating features of both the above approaches. Like the single ontology approach, a single, common, standard ontology is created. It is then used as a basis for reconciling the original ontologies through mappings to each individual ontology.[22] Like the multiple ontologies approach, the individual ontologies are left unchanged, but mappings among them are indirect; i.e., to map information and knowledge from one individual ontology (O1) to another individual ontology (O2), two steps are required: first, map from O1 to the common ontology, then from the common ontology to O2.

While the single ontology approach makes alignment of the ontologies easier, the resulting ontology is often too rigid and does not scale well.[21] On the other hand, while the multiple ontologies approach preserves the integrity of the source ontologies, it is much harder to align them, and it necessitates many sets of mappings when many ontologies need to be reconciled (*O(n2)* sets of bidirectional mappings for *n* individual ontologies).[22]

The methodology we chose for developing the medical error ontology was the hybrid approach—i.e., merge independently developed medical error-related ontologies to build a reference meta-model ontology first, and then use the vocabulary, definitions, relationships, and constraints specified in the reference ontology to implement actual mappings to the source ontologies (Figure 1).

We chose this approach because it incorporates advantages of both other approaches.[22] New ontologies can easily be added without the need of modification in the mappings, thus encouraging scalability and evolution of the medical error ontology. The use of a reference ontology makes the source ontologies comparable, but at the same time, avoids the disadvantages of multiple ontology approaches, since it cuts down the number of sets of mappings to just *n* (bidirectional) mappings for *n* ontologies.

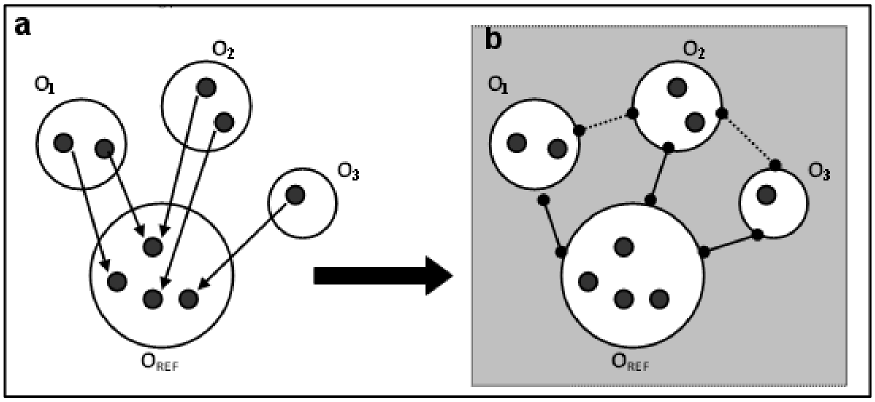


Figure 1. Medical error ontology development approach. (a) The source ontologies (O_1, O_2, O_3) are merged to create a single reference ontology (O_{REF}) Arrows indicate incorporation of concepts from each ontology in the reference ontology. See text for the detailed merging process. (b) The mapped ontologies. Solid links indicate direct mappings; dotted links indicate mappings that did not exist before the merging of the ontologies. They are derived from the mapping process and show the value of the integrated ontology approach.

Ontology Development Process

Our development process was divided into seven major steps:

1. **Identify source ontologies.** Literature search and identification of candidate ontologies.
2. **Align source ontologies (ontology comparison).** Resolve semantic heterogeneities among the individual ontologies to bring them to mutual agreement.
3. **Merge source ontologies (reference ontology).** Merge and rationalize concepts from the aligned ontologies into a reference ontology.
4. **Validate and refine reference ontology.** Test (and refine based on feedback) the ontology for logical consistency, programmatic accuracy, and completeness.
5. **Map source ontologies.** Map original ontologies via linkages between reference ontology and each individual ontology to create a shared medical error ontology.
6. **Validate medical error ontology.** Repeat step 4 to verify and extend defined mappings in order to detect inconsistencies and implied mappings.

7. **Build knowledge base.** Instantiate the ontology with medical error data from various data sources.

Our research is currently at the end of step 4. Therefore, for the rest of this section, we devote our discussions to steps 1 through 4.

Identifying Source Ontologies

Identification of source ontologies started with a literature review using MEDLINE and other Web resources, to search for relevant articles containing medical error taxonomies. We used terms such as taxonomy, categorization, classification, and documentation in conjunction with terms, such as medical error and adverse event. The criterion was to include any taxonomy that would assist us in describing medical errors in one or more clinical domains and/or settings of care and had the potential to be made open source. The taxonomies obtained from the literature review were evaluated for model coverage, utility, and validity in an ad hoc fashion. At the end, the following eight taxonomies were selected for supplying the sum of domains and semantic relations needed for building the reference medical error ontology: NCC MERP Taxonomy of Medication Errors[10] (NCCMERP); Joint Commission Patient Safety Event Taxonomy[7] (PSET); Joint Commission Sentinel Events Reporting[9] (JSER); Taxonomy of Nursing Errors[6] (TNE); A Preliminary Taxonomy of Medical Errors in Family Practice[3] (PTFP); Cognitive Taxonomy of Medical Errors[23] (COG); Taxonomy of Medical Errors for Neonatal Intensive Care[24] (NIC); and MedWatch Index[25] (MEDWATCH). Table 1 shows the high level concepts in the selected taxonomies.

All of these taxonomies were structured as hierarchical taxonomies. We are not aware of and could not find any medical error taxonomies that have been implemented as formal ontologies. In order to integrate these classifications, it was necessary to create an ontology corresponding to each source taxonomy first. While time consuming, this also guaranteed that all the taxonomies and the reference ontology were represented in the same language, thus avoiding the problem of language-level heterogeneity described in the next section.

The source ontologies were implemented in OWL-DL using the Protége OWL Plugin[26] (protege.stanford.edu/plugins/owl/index.html). OWL-DL is based on description logics DL, dl.kr.org/) that make it possible for concepts to be defined and described. Complex concepts can therefore be built up in definitions out of simpler concepts. Furthermore, the logical model allows the use of DL reasoners that can help build and maintain sharable ontologies by revealing inconsistencies, hidden discrepancies, redundancies, and misclassifications.

Table 1. Top level concepts in the selected eight taxonomies

NCC MERP Taxonomy of Medication Errors[10] (NCCMERP)	Preliminary Taxonomy of Medical Errors in Family Practice[3] (PTFP)	Joint Commission Sentinel Events Reporting[9] (JSER)	Joint Commission Patient Safety Event Taxonomy[7] (PSET)
• Patient information • The event • Patient outcome • Product information • Type • Causes • Contributing factors	• Process errors • Knowledge and skill errors	• Type • Settings • Cause • Outcomes • Sources for identification • Method of response	• Impact • Type • Domain • Cause • Prevention and mitigation
Taxonomy of Errors for Neonatal Intensive Care[24] (NIC)	**Taxonomy of Cognitive Errors[23] (COG)**	**Taxonomy of Nursing Errors[6] (TNE)**	**MedWatch Index[25] (MEDWATCH)**
• Actual or potential harm • Patient location • Time since event • Categories of error • Contributing factors	• Cognitive error • Cognitive mechanism • Potential solution	• Lack of attentiveness • Lack of agency/ fiduciary concern • Inappropriate judgment • Medication errors • Lack of intervention on the patient's behalf • Lack of prevention • Missed or mistaken MD/health care provider's orders • Documentation errors	• Patient information • Adverse event or product problem • Suspect medication(s) • Suspect medical device • Reporter • Routes of administration

Aligning Source Ontologies (Ontology Comparison)

In order to successfully merge or map individual ontologies, it is necessary to first align them to bring them to mutual agreement. Several semi-automatic and automatic ontology comparison and integration methods and tools are available.[19, 27, 28, 29] After reviewing several of them, the project team decided that none completely met the needs for this reference ontology. However, using insights gained from these methods, a manual approach for aligning the ontologies by identifying and resolving mismatches between them was adopted.

Two types of mismatches between ontologies have been reported to occur:[19, 29, 30]

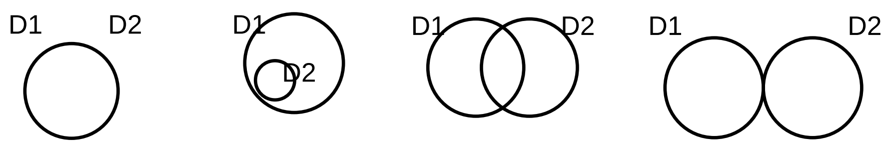
1. **Language-level heterogeneity.** Includes differences in syntax, logical representation, language expressivity, etc., and is a result of different representational languages used for defining classes and relations. As mentioned earlier, since we were creating source and

reference ontologies in the same language, we were able to avoid having to translate between different representational languages.

2. **Ontology-level (semantic) heterogeneity.** Includes differences in conceptualization and explication of the same real world entities and may be caused by differences in scope, model coverage, granularity, concept description, and terminology mismatches (synonyms and homonyms).

Differences in model coverage and granularity do not induce conflicting views of the same concept and therefore do not obstruct the integration process. On the other hand, semantic differences due to differences in conceptualizations can be identified by means of semantic relations between terms and definitions in the individual ontologies.[19, 29, 30, 31] If we assume that the conceptualization of a real-world entity consists of a term *T* and a definition *D* (e.g., natural language definition), then a concept *C* is represented by *C* = (*T*, *D*).[29] Different combinations of these two elements result in a set of possible semantic relations between similar concepts as shown in Table 2.

Table 2. Semantic relations between similar concepts resulting from different combinations of term (T) and definition (D) cases



	D1 = D2	D1 > D2	D1 * D2	D1 ≠ D2
T1 = T2	Equivalency	Additional	Overlap	Homonymy
T1 ≠ T2	Synonymy	Is-A	Overlap	Disjoint

Using these similarity relations, the project team completed the task of aligning the source ontologies manually by finding places where the ontologies were equivalent, disjointed, overlapped, or filled gaps.

Merging Source Ontologies (Reference Ontology)

Once aligned, the source ontologies were merged into the reference ontology, resulting in the creation of the following 12 multidimensional axes in the reference ontology model:

1. Practitioners involved.
2. Patient profile.
3. Health care service.
4. Error location.
5. Contributing factors.
6. Professional activity.
7. Time of error.
8. System factors.

9. Patient outcome.
10. Human factors.
11. Interface design factors.
12. Medical product involved.

The creation of the axes was the equivalent of a metastructure, representing the sum of domains from the source ontologies. While they do not necessarily represent all the individual concepts available in the source ontologies, they represent an aggregation of the existing concepts and their relationships. The ontology was designed in such a way that a given medical error event could be represented as an intersection of the above-mentioned multidimensional axes.

Below are a few examples of how concepts from source ontologies were incorporated into the reference ontology:

- If two concepts had similar definitions (equivalency, synonyms), the result of merging was a single conceptual definition, to be referred to by both original terms during mapping. For example, “Personnel Involved” (NCCMERP) and “Staff” (PSET) are both defined as the health care personnel involved in a medical error event. In the reference ontology, we have created the “Practitioners Involved” class corresponding to both terms.
- If the definition of one concept specialized the definition of another concept—i.e., the latter implied the former—they were in an additional relationship. The result of merging in this case was two concepts related through a concept-subconcept relationship (i.e., every instance of the subconcept also belongs to the parent concept). For example, “Documentation Error” (NIC) was added as a subconcept of “Process Error” (PTFP).
- In some cases, concepts overlapped with each other or existed under a single category in a force-flattened, redundant structure (i.e., an instance of one concept may or may not be an instance of the other). Such ambiguities were resolved by separating out the different concepts (i.e., unflattening the structure) and then declaring an additional new concept or relationship as the intersection (with appropriate constraints specified). For example, NCCMERP includes health care settings (locations) and health care services within the same category of “Setting.” Although health care service and clinical location may seem similar, they are different. For example, an error in oncology (the health care service area) could be committed either in the ICU, the operating room, or the rehabilitation center (the location). This ambiguity was resolved by creating two separate concepts, “Health Care Service” and “Error Location.” The intersection of these two concepts represents the “Setting” in which an error can occur.
- Occasionally, concepts that existed as separate categories in the source ontologies were merged into a single concept. For example, “Contributing Factors” and/or “Cause” (leading to a medical error) are included as top-level categories in many of our source ontologies (NCCMERP, PSET, JSER, TNE).
- In some cases, concepts that were present in lower layers of classification in the source ontologies were elevated to higher levels, and vice versa, based on decisions made by the project team about the degree of significance and applicability of these concepts to the purpose of the ontology.

During the process of developing the ontology, once the concept was clarified, its location on an axis had to be assigned. To the extent possible, we took direction from the existing taxonomies as to where concepts were placed in their hierarchal arrangement. Using the Contributing Factors Axis, concepts were aligned into Action Domains and Action Types using the following structure:

- Contributing factors
  - Action domain
    - Communication
      - Communication with nonphysician colleagues
      - Communication with physician colleagues
      - Communication with patients
    - Investigation
      - Diagnosis imaging
      - Laboratory
      - Other
      - Office administration
      - Appointments
      - Chart completeness
      - Filing system
      - Message handling
      - Patient flow
      - Action type
      - Documenting
      - Followup
      - Implementing
      - Ordering
      - Receiving
      - Reporting
      - Responding to a correct action
      - Responding to an incorrect action

The Contributing Factor Axis covers actions that were taken, not taken, or should have been taken. It does not reflect the person or people who were involved with the action. Personnel are addressed in the Practitioners Involved Axis. At first, it seems counterintuitive to separate actions from the personnel. However, in reviewing the published literature, we found many instances where the personnel were listed but not the actions that were contributing factors to a medical error or near miss report. Likewise, we found reports where the actions were listed, but the personnel were either not explicitly stated or had to be inferred. Since the goal of this ontology was to index the published literature, we felt it was critical that the ontology not force the user to index information that was not explicitly stated in the article. Yet, because the information exists on the two axes, it is quite easy to link the information when both are reported in a given article. The published literature was the consistent gold standard for decisions about how to separate or relate concepts and when to join concepts in order to represent the medical errors and near-miss events.

Reference Ontology Implementation

Implementation of the reference ontology was done using Protégé-OWL. Concepts and subconcepts within each multidimensional axis were defined as classes and subclasses by using the "necessary conditions." At the top of the ontology was the class MedicalErrorEvent, which represented the point of linkage between all the axes. Attributes for each class were defined as properties (functional, transitive, or symmetric). In OWL-DL, properties are binary relations,[26] linking individuals (instances) of two classes (e.g., the expression, "MedicalErrorEvent *hasContributingFactor* ContributingFactors" indicates that an individual belonging to the class MedicalErrorEvent has the property *hasContributingFactor* relating it to another individual from the class ContributingFactors). Properties were also used to create restrictions on a class (e.g., quantifier restrictions, cardinality restrictions), specified within the necessary conditions for that class (e.g., to specify that a medical error event must have at least one contributing factor associated with it, we represented it as, "*hasContributingFactor* **owl:minCardinality** 1"). Figure 2 shows a graphic representation of some of the high-level concepts and relationships in the reference ontology.

Apart from directly creating named classes, new classes were also created as logical combinations of source classes using OWL expressions, such as *unionOf* , *intersectionOf* , *complementOf*, and others. For example, the concept of "Setting" mentioned earlier was defined as the logical combination of HealthcareService and ErrorLocation through *owl:intersectionOf*. The resulting anonymous class was then related to class MedicalErrorEvent by specifying the relationship within MedicalErrorEvent's necessary conditions list as: "*occurredinSetting* (Health careService **owl:unionOf** ErrorLocation)."

Following the above manner, the MedicalErrorEvent class was related, through its properties and restrictions, to every top-level class corresponding to each of the multidimensional axes (e.g., *involvedPatient* PatientInvolved; *wasInitiatedBy* PractitionersInvolved; *wasDiscoveredBy* PractitionersInvolved; *involvedMedicalProduct* [MedicalDevice **owl:unionOf** MedicationProdu]; *ledToPatientOutcome* PatientOutcome; etc.), thus bringing all the axes together at the time of instantiation of this class (an example is provided in the next section).

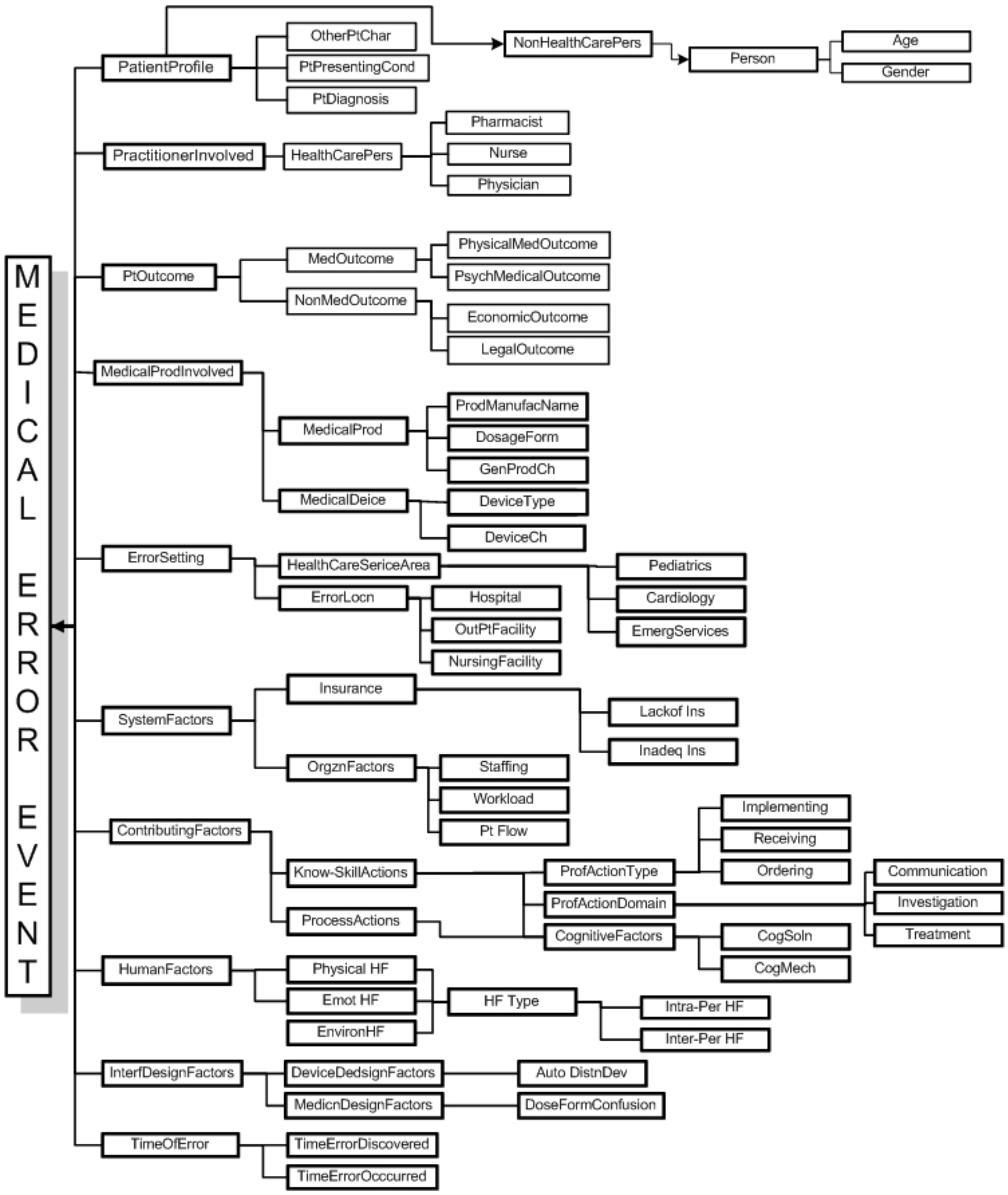


Figure 2. Graphic representation of some high level concepts and relationships in the reference ontology. This reference ontology integrates the various concepts from the source ontology and serves as an integrated ontology.

Validating the Reference Ontology

Since the task of aligning and merging the ontologies was done on an ad hoc basis, such an approach bears the risk of inconsistency and incompleteness. Therefore, once the merger process was complete, an iterative process of validating followed by refining the ontology ensued. As mentioned earlier, DL reasoners can help detect inconsistencies and discrepancies in ontologies.[26] Racer,[32] the DL reasoner to which Protégé-OWL provides access, was used for verifying the logical consistency of the reference ontology. Based on the description (conditions) of a class, Racer can check whether it is possible for the class to have any instances. A class is deemed inconsistent if it cannot possibly have any instances (e.g., a class with two superclasses, where the superclasses are disjoint from each other). Besides consistency checking, Racer also enables automatic computation of the classification hierarchy by determining whether one class is a subclass of another class (multiple inheritance). This helped to keep the ontology structure simple and maintainable, while also minimizing human errors that are inherent in maintaining a multiple inheritance hierarchy.[26] In addition, to ensure best design practices, we performed tests on the ontology on a periodic basis using Protégé-OWL's test framework that provides a standard set of ontology tests (e.g., checking that a property's characteristics correspond correctly with its inverse property's characteristics).

Finally, the ontology was tested for its utility and completeness by instantiating a number of medical error cases selected from the published literature. The cases helped the researchers verify that all the characteristics of a given error case were adequately represented in the reference ontology and provided feedback on further modifications to the ontology. For example, below is a brief description of a medical error case from the Agency for Healthcare Research and Quality Morbidity and Mortality rounds (AHRQ Web M&M) case archive:[33]

> A patient was admitted to the intensive care unit (ICU) with septic shock, and required vasopressors. He appeared to have suffered from a myocardial infarction (MI) in the course of his treatment. It was determined that the cause of the MI was related to a wrong dosage of the drug, Vasopressin (the patient had been receiving 0.4 units/min of vasopressin, rather than the intended dose of 0.04 units/min). The error was discovered the next day, and the patient recovered. Other factors that appeared to have contributed to the case included, (i) a verbal direction was given by the ICU fellow to the resident to order vasopressin; (ii) the resident directly entered the wrong dosage for vasopressin into the computerized physician order entry (CPOE) system, which had a menu of several possible doses of vasopressin; (iii) the error persisted through the next day's multidisciplinary team rounds; and (iv) the error was discovered when one of the ICU nurses was discussing the medication dosing with nursing students, and the incorrect dose was overheard by the ICU fellow and the error recognized.

The various details of this case were first coded in the reference ontology as individuals (instances) of classes and subclasses within the different axes. For example, details about the patient were entered as individuals of classes, PatientAge, PatientGender, PatientPresentingConditions, PatientCo-existingConditions, PatientDiagnosis, etc, which were then related, through properties, to "Pt-1," an individual created for class PatientProfile (corresponding to the "Patient Profile" axis).

Figure 3(a) shows a visualization of Pt-1, using Protégé-OWL's Ontoviz[34] tool (properties listed on the left and individuals of the related classes listed on the right). Similarly, information on the contributing factors leading to the error; personnel involved in initiating, perpetuating, and discovering the error; patient outcome; interface design issues involved in the error; and system factors, etc., were entered as individuals of the corresponding classes.

In the final step, an individual ("Case-1") was created for the class, MedicalErrorEvent, which took in the above individuals as its properties, thus bringing all the multidimensional axes together in a single instance. Figure 3(b) shows the Ontoviz visualization of "Case-1."

Since the original goal for the ontology was to codify the published literature of medical errors and near-miss events, we continue to use published error reports to test the completeness and organization of the ontology. As we enter the individual published articles, we are slowly building a searchable knowledge base of published medical errors and near-miss events.

3(a)

Pt-1	
isDiagnosedWith =	RespiratoryDistress
cameWithPresentingConditions =	ShortnessOfBreath
hasCo-existingConditions =	AtrialFibrillation
	ChronicRenalInsufficiency
	LungCancer
isOfGender =	Male
has Age =	65

3(b)

Case-1	
wasInitiatedBy =	Resident1
involvedPatient =	Pt-1
occurredInSetting =	EmergencyServices-ICU
hasContributingFactors =	Incorrect-CPOE-OrderEntrry
	No-ReviewOfMedicationPrescription
	No-Follow-up-on-VerbalOrder
hasEventDescription =	Pt admitted to ICU with septic shock
hasAssociatedHumanFactors =	NoneIdentified
resultedInPatientOutcome	MI-Temporary
InvolvedInterfaceDesignIssue =	CPOE
wasPerpetuatedBy =	Nurse1
	ICU-Fellow1
involvedMedicalProduct =	ErrorDrug-Vasopressin
wasDiscoveredBy =	ICU-Fellow1
involvedSystemIssues =	ReviewOfDrugDoses
	ReviewOfKnowledgeAssessment

Figure 3. (a) Ontoviz visualization of "Pt-1," an instance of PatientProfile class; (b) Ontoviz visualization of "Case-1," an instance of MedicalErrorEvent class. (Ontoviz is a visualization plug-in in Protégé). The left boxes in each figure are the relations, and the right boxes are the values of the relations.

Conclusions and Future Work

A critical step in the direction of reducing errors and adverse events in health care is the collection, aggregation, and assessment of medical error data, so that potential harm to patients can be identified and steps taken to prevent or mitigate harm. Despite the efforts now devoted to reporting and collecting medical error data, no standardized framework for classifying and evaluating such data is currently available. This paper describes our experience with developing a comprehensive medical error ontology to serve as a standard representation for medical error concepts gleaned from various existing published taxonomies, with the ultimate goal of preventing errors and improving patient safety.

Following a literature review of existing and published taxonomies related to medical errors, we selected eight candidate taxonomies to supply the sum of domains for our ontology. The taxonomies were modeled as ontologies in Protégé-OWL and then aligned with one another through identification of semantic relations between concepts contained in the source ontologies. Once aligned, the source ontologies were merged to produce a single reference ontology with 12 multidimensional axes, the intersection of which characterizes a medical error event. The classes and relationships in the reference ontology were implemented using Protégé-OWL.

The next step in our project will be to create mappings among the source ontologies via the reference ontology. Once mapped, the source ontologies will form the integrated medical error ontology that combines the domains of the source ontologies, while preserving their integrity. The mechanism for mapping will involve defining equivalencies between classes and properties—i.e., specifying that a particular class or property in one ontology is equivalent to a class or property in a second ontology. Protégé-OWL provides built-in mapping support[35] for defining such equivalence classes using "necessary & sufficient" conditions.

The medical error ontology will continue to evolve as we map additional relevant ontologies to the reference ontology as they become available. Candidate ontologies for consideration will include those that have potential to be made open source and are designed to encompass one or more clinical domains and/or settings of care. For example, Stetson, et al., are developing an ontology that models medical errors as an intersection of three domains: human/system errors, information needs, and communication space.[36] Once developed and published, the concepts defined in this ontology will help enrich and extend the representation of these three domains in our integrated medical error ontology.

We have seen how combined use of ontologies can be hindered by language-level and ontology-level (semantic) heterogeneities. While language-level heterogeneities did not a pose a problem for us, since all the source ontologies were coded using OWL-DL, they might pose a problem when integrating new ontologies represented in other languages. However, one of the major benefits of using Protégé-OWL for our medical error ontology is that its underlying knowledge model is designed to be compatible with Open Knowledge Base Connectivity (OKBC),[19] which provides reasonable support for language-level interoperability. Semantic heterogeneities during the integration process will be identified in the manner described in the ontology alignment section. The merger of new ontologies into the reference ontology will entail using a set of change operations that may affect the present ontology structure. Our choice for a change model will be based on experiences and ideas of other researchers[37, 38, 39] and our own ideas.

The final step in the project involves using the shared medical error ontology to build a knowledge base of medical error events from various published data sources, using the *Forms* and *Individuals* tab provided by Protégé-OWL. Once the knowledge base is built, medical error events can be analyzed, interpreted, and understood. The "Queries tab" provided by Protégé-OWL enables running simple queries on the ontology knowledge base at an instance level, while more complex queries can be built using a more powerful query language such as RDQL (www.w3.org/Submission/2004/SUBM-RDQL-20040109/). For example, one could obtain all the instances of medical error events that were initiated in a particular health care setting (e.g., ICU) and involved a particular interface design issue (e.g., in the computerized physician order entry system). Such querying techniques would help in identifying potential targets for which patient safety interventions could be developed to prevent future adverse events.

Finally, although the shared medical error ontology itself is not a reporting system, it is designed to support integration of data across varied reporting systems and can provide systematic, principled methods for the design of improved medical error reporting systems in the future.

Acknowledgments

This project was supported by Grant No. R01 LM07894 and Grant No. 5 T15 LM07093 from the National Library of Medicine. We thank Soni H. Gupta, Yang Gong, and Arunkumar Srinivasan for their contributions in the project.

Author Affiliations

School of Health Information Sciences, University of Texas Health Science Center at Houston (Mr. Mokkarala, Dr. Johnson, Dr. Zhang, Dr. Turley); School of Nursing, University of Kansas (Dr. Brixey); Department of Biomedical Informatics, Arizona State University (Dr. Patel).

Address correspondence to: James P. Turley or Jiajie Zhang, School of Health Information Sciences, University of Texas Health Science Center at Houston, 7000 Fannin, Suite 600, Houston, TX 77030; e-mail: James.P.Turley@uth.tmc.edu; Jiajie.Zhang@uth.tmc.edu.

References

1. Kohn LT, Corrigan JM, Donaldson MS, eds. To err is human. Washington, DC: National Academies Press; 1999.
2. FDA. MEDWATCH - The FDA Safety Information and Adverse Event Reporting Program. Available at: www.accessdata.fda.gov/scripts/medwatch/. Accessed April 18, 2008.
3. Dovey SM, Meyers DS, Phillips RL Jr, et al. A preliminary taxonomy of medical errors in family practice. Qual Saf Health Care 2002; 11: 233-238.
4. Battles JB, Shea CE. A system of analyzing medical errors to improve GME curricula and programs. Acad Med 2001; 76: 125-133.
5. Runciman WB, Helps SC, Sexton EJ, et al. A classification for incidents and accidents in the health care system. J Qual Clin Pract 1998; 18: 199-211.
6. Benner P, Sheets V, Uris P, et al. Individual, practice, and system causes of errors in nursing: A taxonomy. J Nurs Adm 2002; 32: 509-523.
7. Chang A, Schyve PM, Croteau RJ, et al. The JCAHO patient safety event taxonomy: A standardized terminology and classification schema for near misses and adverse events. Int J Qual Health Care 2005; 17: 95-105.
8. Elder NC, Dovey SM. Classification of medical errors and preventable adverse events in primary care: A synthesis of the literature. J Fam Pract 2002; 51: 927-932.
9. Joint Commission. Sentinel event statistics. 2005 Mar 31. Available at: http://www.jcaho.org/accredited+organizations/sentinel+event/sentinel+event+statistics.htm. Accessed April 28, 2005.

10. National Coordinating Council on Medication Error Reporting and Prevention. Taxonomy of medication errors. Available at: www.nccmerp.org/pdf/taxo2001-07-31.pdf. Accessed April 18, 2008.

11. Boxwala AA, Dierks M, Keenan M, et al. Organization and representation of patient safety data: Current status and issues around generalizability and scalability. J Am Med Inform Assoc 2004; 11: 468-478.

12. Quality Interagency Coordination Task Force. Doing what counts for patient safety: Federal actions to reduce medical errors and their impact. Report of the Quality Interagency Coordination Task Force (QuIC) to the President. Available at: http://www.quic.gov/report/toc.htm. Accessed April 18, 2008.

13. Gruber TR. A translational approach to portable ontologies. Knowl Acquis 1993; 5: 199-220.

14. Noy NF, McGuinness DL. Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880. Stanford, CA: Knowledge Systems Laboratory, Stanford University; 2001.

15. Stearns MQ, Price C, Spackman KA, et al. SNOMED clinical terms: Overview of the development process and project status. AMIA Annu Symp Proc 2001: 662-666.

16. Humphreys BL, Lindberg DA. The UMLS project: Making the conceptual connection between users and the information they need. Bull Med Libr Assoc 1993; 81: 170-177.

17. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998; 37: 394-403.

18. Noy NF, Musen MA. An algorithm for merging and aligning ontologies: Automation and tool support. Sixteenth National Conference on Artificial Intelligence (AAAI99), Workshop on Ontology Management; July 18-22, 1999; Orlando, FL.

19. Klein M. Combining and relating ontologies: Problems and solutions. IJCAI workshop on ontologies; 2001Aug 4; Seattle, WA.

20. Pinto HS, Martins JP. A methodology for ontology integration. Proceedings of the First International Conference on Knowledge Capture (K-CAP); 2001 Oct 21-23; Victoria, BC.

21. Su X. A text categorization perspective for ontology mapping. Norway: Department of Computer and Information Science, Norwegian University of Science and Technology; 2002.

22. Hameed A, Preece AD, Sleeman DH. Ontology reconciliation. In: Staab S, Studer R, eds. Handbook on ontologies. International handbook on information systems. New York: Springer; 2004. p. 231-250.

23. Zhang J, Patel VL, Johnson TR, et al. Toward an action-based taxonomy of human errors in medicine. In: Gray W, Schunn C, eds. Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum Associates; 2002. p. 970-975.

24. Suresh G, Horbar JD, Plsek P, et al. Voluntary anonymous reporting of medical errors for neonatal intensive care. Pediatrics 2004; 113: 1609-1618.

25. Meadows M. MedWatch: Managing risks at the FDA. FDA Consum 2003; 37: 10-11.

26. Knublauch H, Fergerson RW, Noy NF, et al. The Protégé OWL plugin: An open development environment for semantic web applications. Third International Semantic Web Conference - An architectural overview for developers and decision-makers. Hiroshima, Japan: ISWC; 2004.

27. Noy NF, Musen MA. PROMPT: Algorithm and tool for automated ontology merging and alignment. Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000). Austin, TX: AAAI/MIT Press; 2000.

28. McGuinness DL, Fikes R, Rice J, et al. An environment for merging and testing large ontologies. In: Cohn AG, Giunchiglia F, Selman B, eds. KR2000: Principles of knowledge representation and reasoning. San Francisco: Morgan Kaufmann; 2000. p. 483-493.

29. Kavouras M. A unified ontological framework for semantic integration. Proceedings of the International Workshop on Next Generation Geospatial Information; Boston, MA; 2003 Oct 19-21.

30. Hakimpour F, Geppert A. Resolving semantic heterogeneity in schema integration. Proceedings of the international conference on formal ontology in information systems; 2001 Oct 17-19. Ogunquit, ME. p. 297-308.

31. Hammer J, McLeod D. An approach to resolving semantic heterogeneity in a federation of autonomous, heterogeneous database systems. Int J Intel Coop Inform Syst 1993; 2: 51-83.

32. Liebig T, Pfeifer H, Henke FW. Reasoning services for an OWL authoring tool: An experience report. In: Haarslev V, Möller R, eds. Proceedings of the 2004 International Workshop on Description Logics (DL2004). CEUR Workshop; 2004 June 6-8; Whistler, BC: CEUR-WS.org; 2004.

33. Flanders SA, Saint S. Getting to the root of the matter. AHRQ Web M&M. Available at: www.webmm.ahrq.gov/case.aspx?caseID=98. Accessed April 18, 2008.

34. Storey M, Lintern R, Ernst N, Perrin D. Visualization and Protégé. Victoria, BC, Canada: University of Victoria;. Available at: http://www.smi.stanford.edu/projects/protege/conference/2004/abstracts/Storey.pdf. Accessed June 16, 2008.

35. Knublauch H, Musen MA, Noy NF. Tutorial: Creating semantic Web (OWL) ontologies with Protégé. 2nd International Semantic Web Conference (ISWC2003); 2003 Oct 20-23; Sanibel Island, FL.

36. Stetson PD, McKnight LK, Bakken S, et al. Development of an ontology to model medical errors, information needs, and the clinical communication space. AMIA Annu Symp Proc 2001: 672-676.

37. Cimino JJ. Formal descriptions and adaptive mechanisms for changes in controlled medical vocabularies. Methods Inf Med 1996; 35: 202-210.

38. Oliver DE, Shahar Y. Development of a change model for a controlled medical vocabulary. Proc AMIA Annu Fall Symp 1997: 605-609.

39. Bernauer J. Subsumption principles underlying medical concept systems and their formal reconstruction. Proc Annu Symp Comput Appl Med Care 1994: 140-144.