

Research and Applications

A value set for documenting adverse reactions in electronic health records

Foster R Goss,¹ Kenneth H Lai,² Maxim Topaz,^{3,4} Warren W Acker,³ Leigh Kowalski,³ Joseph M Plasek,^{2,3} Kimberly G Blumenthal,^{5,6} Diane L Seger,⁷ Sarah P Slight,^{3,4,8} Kin Wah Fung,⁹ Frank Y Chang,⁷ David W Bates,^{2,4,6} and Li Zhou^{3,6,10*}

¹Department of Emergency Medicine, University of Colorado, Aurora, CO, USA, ²Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT, USA, ³Division of General Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, USA, ⁴Division of Pharmacy, School of Medicine, Pharmacy and Health, Durham University, Durham, UK, ⁵Division of Rheumatology, Allergy and Immunology, and Medical Practice Evaluation Center, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA, ⁶Harvard Medical School, Boston, MA, USA, ⁷Clinical and Quality Analysis, Partners HealthCare System, Boston, MA, USA, ⁸Newcastle upon Tyne Hospitals NHS Foundation Trust, UK, ⁹National Library of Medicine, Bethesda, MD, USA and ¹⁰Clinical Informatics, Partners eCare, Partners HealthCare System, Boston, MA, USA

Corresponding Author: Foster R Goss, B215, Leprino Building, 12401 East 17th Ave, Aurora, CO 80045, USA.

E-mail: foster.goss@ucdenver.edu

Received 14 April 2017; Revised 18 October 2017; Editorial Decision 19 October 2017; Accepted 4 November 2017

ABSTRACT

Objective: To develop a comprehensive value set for documenting and encoding adverse reactions in the allergy module of an electronic health record.

Materials and Methods: We analyzed 2 471 004 adverse reactions stored in Partners Healthcare's Enterprise-wide Allergy Repository (PEAR) of 2.7 million patients. Using the Medical Text Extraction, Reasoning, and Mapping System, we processed both structured and free-text reaction entries and mapped them to Systematized Nomenclature of Medicine – Clinical Terms. We calculated the frequencies of reaction concepts, including rare, severe, and hypersensitivity reactions. We compared PEAR concepts to a Federal Health Information Modeling and Standards value set and University of Nebraska Medical Center data, and then created an integrated value set.

Results: We identified 787 reaction concepts in PEAR. Frequently reported reactions included: rash (14.0%), hives (8.2%), gastrointestinal irritation (5.5%), itching (3.2%), and anaphylaxis (2.5%). We identified an additional 320 concepts from Federal Health Information Modeling and Standards and the University of Nebraska Medical Center to resolve gaps due to missing and partial matches when comparing these external resources to PEAR. This yielded 1106 concepts in our final integrated value set. The presence of rare, severe, and hypersensitivity reactions was limited in both external datasets. Hypersensitivity reactions represented roughly 20% of the reactions within our data.

Discussion: We developed a value set for encoding adverse reactions using a large dataset from one health system, enriched by reactions from 2 large external resources. This integrated value set includes clinically important severe and hypersensitivity reactions.

Conclusion: This work contributes a value set, harmonized with existing data, to improve the consistency and accuracy of reaction documentation in electronic health records, providing the necessary building blocks for more intelligent clinical decision support for allergies and adverse reactions.

Key words: drug-related side effects and adverse reactions, allergy and immunology, hypersensitivity, natural language processing, electronic health records, vocabulary, controlled.

OBJECTIVE, BACKGROUND, AND SIGNIFICANCE

Adverse reactions to foods, pharmaceuticals, and diagnostic products create significant costs, morbidity, and mortality in our health care system.^{1–3} Adverse drug reactions (ADRs) have been reported to affect 10%–20% of hospitalized patients and up to 25% of outpatients.^{4–6} Some reactions, while rare, can be life-threatening; for example, toxic epidermal necrolysis (TEN), Stevens-Johnson syndrome (SJS), drug reaction with eosinophilia and systemic symptoms syndrome (DRESS), and immune hepatitis. In the United States, it has been estimated that nearly 1 in 300 hospitalized patients dies from an ADR every year.¹ Accurate documentation of a patient's adverse reactions to medications, products, or foods in the electronic health record (EHR) represents an important part of patient safety.

Allergy modules within EHRs provide a location for clinicians to document patients' adverse reactions. Although such modules use the term “allergy,” many reactions entered in the allergy module are not immunologically mediated.⁷ Non-immunologically mediated reactions such as intolerances, toxicities, and idiosyncratic and pseudoallergic reactions can also be clinically important and are often documented here. In fact, it is estimated that only about 5%–10% of adverse reactions to drugs are allergic (immune-mediated).⁷ Some EHR systems provide coded options in the allergy module for clinicians to indicate a reaction as an allergy, intolerance, or contraindication, but studies have found that reaction type and classification are poorly understood by clinicians.⁸ In addition, in most EHR systems, code sets for reactions are typically provided by third-party content vendors and can vary considerably based on the vendor and/or local management of terminologies. Well-developed documentation standards for EHR systems exist in some areas, such as medications, but not in others, such as allergies.⁹ Vendors have found this to be a problem, where one health care institution may use 87 codes to encode allergic reactions and another may have 12.^{9,10}

Standard value sets for encoding adverse reactions offer a solution to this problem by providing a list of codes to improve consistent data capture, clinical documentation, and quality reporting.¹¹ The National Quality Forum defines a *value set* as a common group of codes used to define a clinical concept drawn from standard terminologies (eg, Systematized Nomenclature of Medicine – Clinical Terms [SNOMED-CT] and International Classification of Diseases, Tenth Revision).¹¹ Clinical decision support (CDS) and quality measures hinge on the accuracy and completeness of these value sets to drive adverse event guidance/avoidance; however, there are many inherent challenges in creating a value set, including choosing a terminology with comprehensive coverage and appropriate granularity, integrating concepts across disparate systems, and ensuring that concepts are semantically consistent.¹¹

Multiple terminologies have been used to code adverse reactions. Although the World Health Organization Adverse Reaction Terminology¹² and the Medical Dictionary for Regulatory Activities¹³ are used by pharmaceutical manufacturers to report adverse drug events to regulatory agencies^{14,15} and the International Classification of Diseases, Clinical Modification is used for pharmacovigilance purposes,^{16–19} these terminologies are not optimized for clinical practice and use within EHRs. Recently, SNOMED-CT was recommended by the Federal Health Information Modeling and Standards (FHIMS) program for coding adverse reactions. An Adverse Clinical Reaction value set created by FHIMS using SNOMED-CT codes was recently released by the National Library

of Medicine Value Set Authority Center (VSAC).²⁰ VSAC provides official versions of value sets used in clinical quality measures and other use cases.²⁰ The current published value set comprises reactions collected from Intermountain Healthcare, Kaiser Permanente, and the Veterans Administration's systems based on reported frequencies, but the underlying processes and methods for creating this value set are unclear. In addition, there are no published studies that perform external validation of this value set.

In this study, we developed a comprehensive value set for documenting and encoding adverse reactions in an EHR by examining frequencies of ADRs, including hypersensitivity reactions (HSRs),^{21,22} using data stored in a large allergy repository. The goal was to create a value set that is consistent with those found in the patient record, integrated with the most current terminology updates and aligned with the current data models needed to support CDS. We then compare this with existing value sets to generate an integrated value set for encoding adverse reactions. This approach aligns with initiatives at many health care professional and standards organizations (eg, the American Academy of Allergy, Asthma, and Immunology,²³ Health Level Seven,²⁴ the National Quality Forum,¹¹ and the International Health Terminology Standards Development Organization²⁵) and institutions (eg, the US Pharmacopeia Convention²⁶) to develop allergy/intolerance standards and knowledge in EHRs to improve care. We used a natural language processing (NLP) tool called the Medical Text Extraction, Reasoning, and Mapping System (MTERMS)^{27,28} to process reaction entries in Partners Healthcare's Enterprise-wide Allergy Repository (PEAR). Because most reaction entries were short-length free-text descriptions, we used MTERMS' tokenization, misspelling correction,²⁹ lexical lookup, and terminology mapping functions to process and map reaction terms to SNOMED-CT. The feasibility of extracting and mapping allergies has been demonstrated in prior work on encoding food allergens^{30,31} and extracting allergy information from clinical notes.²⁸

MATERIALS, METHODS, AND RESULTS

Our approach included 2 phases (Figure 1). In the first phase, we processed patient reaction entries in PEAR using MTERMS, created a PEAR reaction lexicon, and mapped these reactions to SNOMED-CT concepts. As described in prior work,^{27,28,30} MTERMS uses regular expression rules and a lexicon to: (1) process and tag allergens and reactions, (2) correct misspellings,²⁹ (3) handle lexical variations (eg, abbreviations), and (4) map terms to concepts within standard terminologies. In the second phase, we calculated the frequencies of reaction concepts, including rare or severe reactions and HSRs, in PEAR; compared PEAR to FHIMS and a set of reactions provided by the University of Nebraska Medical Center (UNMC); and created an integrated value set. As our methods involved multiple steps and each step generated corresponding results, we present them in one section accordingly.

Definitions

Reactions are described at the entry, term, and concept levels. We defined an “entry” as word(s), phrase(s), or sentence(s) entered by a clinician into the reaction field of a patient's allergy record. A “term” was the portion of a reaction entry corresponding to the actual reaction, which is not normalized and may contain misspellings, acronyms, or syntactic variations. A “concept” was a collection of synonymous terms that represent a specific reaction. We defined a

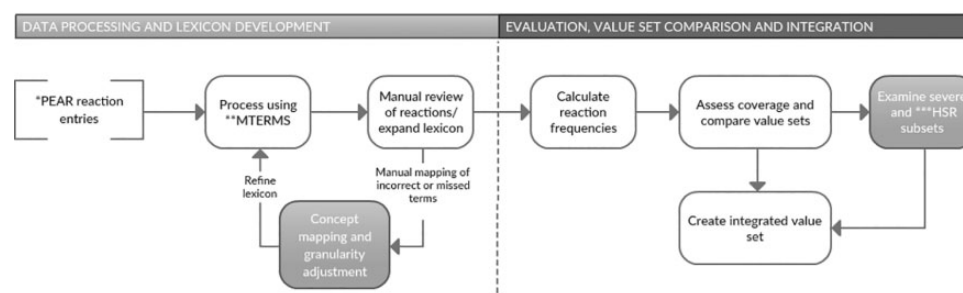


Figure 1. Overview of methods.

PEAR: Partners Enterprise-wide Allergy Repository; MTERMS: Medical Text Extraction, Reasoning and Mapping System; HSRs: hypersensitivity reactions.

“value set” as a common group of numerical codes used to define a specific clinical concept, each derived from a standard terminology (ie, SNOMED-CT) for use within a clinical domain (ie, adverse reactions).^{11,20} A “lexicon” was the vocabulary of our NLP system that was used to identify reaction terms in free-text entries. Our reaction lexicon contained additional terms collected from free-text that are not currently included in SNOMED-CT. For example, in the reaction entry “he was switched from APAP [acetaminophen] to ASA [aspirin] due to elev [elevated] LFTs [liver function tests] and cough,” we extracted the terms “elev LFTs” and “cough.” Since SNOMED-CT currently does not have a concept for “elevated liver function tests,” we therefore mapped it to the semantically closest SNOMED-CT concept ID, 707724006, “elevated liver enzymes level” (using its preferred term) along with other synonyms such as “LFT elevation.”

Setting and corpus

Our study used PEAR, which contains allergy information from all patients within the federated hospital/provider network entered by clinicians in the EHR’s allergy module.³² As of October 26, 2014, PEAR contained 3 949 996 active allergy entries for 2 730 250 unique patients, including drug, food, and environmental allergens, with 2 315 944 (58.6%) allergens having one or more reaction entries, accounting for 2 471 004 active allergy reaction entries. Among these reactions, 1 751 817 (70.9%) were coded entries (using 35 unique locally defined codes, including “Unknown”) and 719 187 (29.1%) were free-text entries. For external validation, we compared the concepts generated with 2 external datasets: (1) the FHIMS adverse reaction value set, which spans 599 concepts stemming from multiple hierarchies within SNOMED-CT, and (2) 157 806 (604 unique) adverse reaction entries from the UNMC. For this study, we mapped reaction terms in PEAR to the March 2016 release of the US edition of SNOMED-CT.³³ This study was approved by the Partners and University of Colorado Multiple Institutional Review Board.

Phase 1: data processing and lexicon development

The average length of PEAR free-text entries was 12.1 (range 1–255) characters. Entries of length >255 characters were likely truncated at some point in the data storage/extraction process. Many free-text entries were long narratives containing other contextual information, such as “after a shot of PEN [penicillin], walked across the room and passed out, in childhood.” To develop our lexicon, we utilized a subset of our corpus consisting of all entries with a frequency >10 entries, resulting in 539 610 (75.0% of total) free-text entries, corresponding to 3160 unique entries, with an average

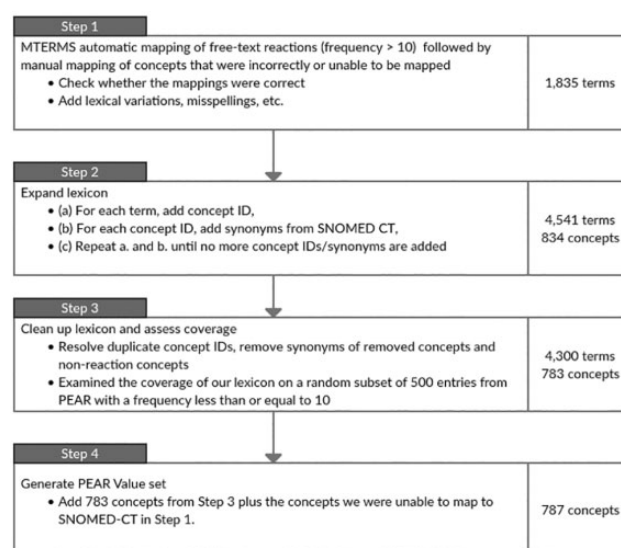


Figure 2. Adverse reaction lexicon development.

length of 8.7 (range 1–68) characters. To ensure that the lexicon did not miss important but rare concepts, we randomly selected a subset of 500 reaction entries with a frequency of 10 or less (average length 22.1 [range 1–255] characters) for an internal evaluation. Data processing and lexicon development involved 4 steps, as summarized in Figure 2 and described in detail below.

Step 1: Map reaction entries to SNOMED-CT descriptions

The first step focused on mapping PEAR reaction entries of frequency >10 to SNOMED-CT at the description (term) level via MTERMS. The 3160 unique terms and their mappings to SNOMED-CT were then reviewed by a research assistant or pharmacy student in their fifth year of training and then reviewed and approved by a physician informatician (FG). Because our task was to check whether the mapping was correct, we performed this manually in Excel. To decide whether a term was a reaction or not, we used the clinicians’ judgment. For cases that were unclear, we discussed with a panel that consisted of pharmacists, allergy specialists, and physicians to achieve consensus. We also discussed with a terminology expert from SNOMED-CT (KF) to check whether the reaction mapped to an appropriate concept in SNOMED-CT. The terms found in this step formed the basis of our lexicon.

MTERMS mapped 3160 unique free-text PEAR reaction entries to 757 SNOMED-CT descriptions. Thirty-one automatically

mapped terms were not reactions (eg, “male,” “near”) and therefore were removed from the lexicon, yielding 726 correctly mapped terms. We then manually mapped 1109 terms that were unable to be automatically mapped to SNOMED-CT (eg, they were too poorly misspelled for the spell checker or used abbreviations, eg, “ITP” for “idiopathic thrombocytopenic purpura”). Finally, we identified 4 terms (itchy tongue, sores in throat, throat tingling, and tingling in throat) that, although they represented reactions, we were unable to map to SNOMED-CT, resulting in 1835 terms in our lexicon.

Step 2: Map reaction terms to SNOMED-CT concepts

This step aimed to further expand the lexicon using the following 3 sub-steps. First, we used MTERMS to map these 1835 terms to their SNOMED-CT concept IDs. For example, “rash” was mapped to 2 SNOMED-CT concept IDs: 271807003, eruption of skin (disorder), and 112625008, cutaneous eruption (morphologic abnormality), because “rash” is listed as a synonym of these 2 concepts. Second, we included SNOMED-CT synonyms for each identified concept. For example, for concept ID 271807003, we included the synonyms “eruption of skin,” “eruption,” “exanthema,” etc. Third, we repeated the above 2 sub-steps until no more concept IDs or synonyms could be added.

This resulted in 4541 terms representing 834 concepts being included in our lexicon.

Step 3: Refine mapping hierarchies

It is possible that a term was mapped to multiple concept IDs within different top-level hierarchies (or axes) of SNOMED-CT. In the above example, “rash” was mapped to 2 concepts within 2 different hierarchies: disorder and morphologic abnormality, where the latter hierarchy was not related to adverse reactions and was therefore excluded from the lexicon. This enabled us to clean up the lexicon by excluding nonreaction concepts and unrelated hierarchies such as morphologic abnormality, qualifier value, and observable entity.

At this stage, our lexicon comprised 4300 terms representing 782 concepts, split between the disorder (47.2%, $n = 369$) and findings (52.8%, $n = 413$) hierarchies.

Step 3a: Assess lexicon coverage by processing less frequent reaction entries

To ensure that our lexicon was not missing rare reactions with lower frequencies, we examined its coverage on a random subset of 500 free-text entries from PEAR with a frequency ≤ 10 . Using MTERMS, we processed this subset and calculated its coverage. The commonly used statistical measures of precision, recall, and F1-measure were also calculated.³⁴

After manual review, 928 reaction terms were identified in those 500 free-text entries. MTERMS achieved 98.0% precision (887/905) and 95.6% recall (887/928), yielding an F1-measure of 96.8% at the term level on this subset. MTERMS incorrectly identified 8 terms that were not reactions (eg, “glaucoma” in “can’t take for glaucoma” and “faint” in “faint rash”). There were 41 reaction terms that MTERMS was not able to recognize, mainly due to lexical variations (eg, “racing heartbeat” for “palpitations”), abbreviations (eg, “renal dys” for “renal dysfunction”), or misspellings (eg, “pancritis” for “pancreatitis”). However, each of these 41 terms already had a corresponding concept in our lexicon. Only one reaction concept, “livedo reticularis,” was not in our lexicon and was subsequently added. Most of the 41 terms identified by MTERMS were low in frequency, having only 1 or 2 documented entries

(average = 1.7). Overall, MTERMS performed well at identifying terms that existed in these longer narratives and mapping them to existing concepts in our lexicon. Our lexicon covered the majority of reactions in this subset, supporting our belief that few new (and rare) reactions were likely to be found in these longer narratives and increasing our confidence in the lexicon and methods.

Step 4: Create PEAR value set (PEAR-VS)

Our PEAR value set (PEAR-VS) included a total of 787 concepts (ie, the 782 concepts from Step 3 plus 1 concept from Step 3a and the 4 concepts that we were unable to map to SNOMED-CT in Step 1).

Phase 2: Evaluation, value set comparison, and integration

Frequencies of reaction concepts in PEAR

To assess the frequency of reaction concepts, we processed all reaction entries in PEAR (ie, all free-text and structured entries) using MTERMS and our final lexicon. We identified 2 584 112 terms (1 770 418 in coded entries and 813 694 in free-text entries), mapped these terms to their corresponding SNOMED-CT concepts in PEAR-VS, and calculated the frequency of each concept in PEAR (Table 1). We found the most frequently observed reactions were rash (eruption of skin) (13.96%, $n = 360\,859$), hives (weal) (8.25%, $n = 213\,228$), and gastrointestinal upset (irritation) (5.47%, $n = 141\,389$). Of interest, some of the most frequent concepts were only entered as free-text, including dizziness and palpitations.

Comparison with existing reaction value sets

Comparison among value sets was performed using methods similar to Zhou et al.,¹⁵ where we mapped concepts between our lexicon and the target terminology (FHIMS, UNMC), classifying each match as exact, partial (broad or narrow), or missing. Exact matches were defined as instances where the *concept IDs* were identical or the *terms* were similar in their meaning and granularity (eg, “bleeding from nose” → “nose bleed”). Broad matches were defined as instances where the term in our lexicon was less specific than that of our target terminology (eg, “liver damage” → “cirrhosis of liver”). Conversely, narrow matches were instances when the term in our lexicon was more specific than the target terminology (eg, “GI upset” → “gastrointestinal symptom”). Broad and narrow matches involved only one reaction, and the reaction could have included a modifier (eg, “acute”) or qualifier (eg, “severe”). When 2 or more reactions existed within a single concept, they were classified as pre-coordinated (eg, “nausea, vomiting, and diarrhea”). A percent match was calculated for each category to the target terminologies. For example, in FHIMS, there were 92 partial matches that were classified as broad out of a total of 599 FHIMS concepts, representing 15.4% of concepts. As we knew the frequency of concepts in each target terminology, we calculated a frequency-weighted percent match by concept. For example, the frequency of gastrointestinal symptom within UNMC was 0.113% (179 reaction entries in 157 806 total reactions). These concepts were summed across each value set and stratified by match type.

When comparing FHIMS to PEAR-VS at the *concept* level, we found 321 exact matches at the *concept ID* level and 59 exact matches at the *term* level, resulting in a total of 380 exact matches, representing 63.4% of the concepts in FHIMS (Table 2). By frequency of documentation, the number of concepts covered by PEAR-VS was 97.48% at the exact level and 99.35% including partial matches.

Table 1. Top 20 reaction concepts in PEAR

Reaction	SNOMED-CT ConceptID	Total Frequency (%)	Coded Frequency (%)	Free-text Frequency (%)
Eruption of skin (ie, rash)	271807003	360 859 (13.96)	251 840 (14.22)	109 019 (13.40)
Weal (ie, hives)	247472004	213 228 (8.25)	165 108 (9.33)	48 120 (5.91)
Gastrointestinal irritation	95516005	141 389 (5.47)	113 323 (6.40)	28 066 (3.45)
Itching	418290006	81 462 (3.15)	67 707 (3.82)	13 755 (1.69)
Anaphylaxis	39579001	63 632 (2.46)	59 915 (3.38)	3717 (0.46)
Nausea	422587007	45 300 (1.75)	20 035 (1.13)	25 265 (3.10)
Swelling	65124004	39 522 (1.53)	20 067 (1.13)	19 455 (2.39)
Cough	49727002	37 510 (1.45)	14 860 (0.84)	22 650 (2.78)
Vomiting	422400008	34 894 (1.35)	18 854 (1.06)	16 040 (1.97)
Angioedema	41291007	32 619 (1.26)	27 196 (1.54)	5423 (0.67)
Altered mental status	419284004	29 709 (1.15)	27 788 (1.57)	1921 (0.24)
Dyspnea	267036007	25 013 (0.97)	18 235 (1.03)	6778 (0.83)
Muscle pain	68962001	23 568 (0.91)	8539 (0.48)	15 029 (1.85)
Headache	25064002	22 534 (0.87)	8084 (0.46)	14 450 (1.78)
Bronchospasm	4386001	18 521 (0.72)	17 934 (1.01)	587 (0.07)
Diarrhea	62315008	18 365 (0.71)	9106 (0.51)	9259 (1.14)
Wheezing	56018004	17 871 (0.69)	15 869 (0.90)	2002 (0.25)
Sneezing	76067001	16 856 (0.65)	11 792 (0.67)	5064 (0.62)
Dizziness	404640003	12 258 (0.47)	0 (0.00)	12 258 (1.51)
Palpitations	80313002	8724 (0.34)	0 (0.00)	8724 (1.07)
Others	–	403 556 (15.62)	36 664 (2.07)	366 892 (45.09)
Unknown	–	936 722 (36.25)	857 502 (48.44)	79 220 (9.74)
Total	–	2 584 112 (100)	1 770 418 (100)	813 694 (100)

Table 2. Concept coverage between PEAR-VS and FHIMS/UNMC

Match Type	FHIMS		UNMC		Examples
	No. of concepts (% concepts)	% frequency of concepts (source database)	No. of concepts (% concepts)	% frequency of concepts (source database)	
Exact Match	380 (63.4)	97.48	454 (75.2)	98.73	Heart irregular irregular heartbeat
Partial Broad ^a	92 (15.4)	1.55	75 (12.4)	0.72	Swelling of lower jaw region face swelling
Partial Narrow ^b	25 (4.17)	0.32	31 (5.1)	0.33	Disease of liver liver damage
Missing	96 (16.03)	0.54	39 (6.5)	0.16	Mass, elevated INR
Pre-coordinated ^c	6 (1.0)	0.09	5 (0.83)	0.07	“Nausea, vomiting, and diarrhea?”
Total (partial+exact)	497 (82.97)	99.35	560 (92.7)	99.78	

^aPartial Broad: Broad matches in our study were defined when the terms in our lexicon were less specific than our target terminology (FHIMS, UNMC) (eg, “liver damage” → “cirrhosis of liver”).

^bPartial Narrow: Narrow matches were defined when the terms in our lexicon were more specific than the target terminology (eg, “GI upset” → “gastrointestinal symptom”).

^cPre-coordinated: where 2 or more reactions existed within a single concept.

When comparing the UNMC reactions to PEAR-VS at the *concept ID* level, 408 concepts were an exact match by *concept ID* and 46 concepts were an exact match at the *term* level (Table 2). By frequency, PEAR-VS covered 98.73% of concepts at the exact level and 99.78% including partial matches.

From the 135 (96 FHIMS, 39 UNMC) unique missing concepts, there were 12 (10 from FHIMS and 2 from UNMC) that, upon review, we believed should be excluded from the integrated reaction value set. Reasons for possible exclusion are shown in Table 3.

Rare or severe reactions

Severe reactions, while low in frequency, can be the most critical for ensuring patient safety. Using a list of 15 rare or severe reactions compiled by expert review (KB, FG) in PEAR, we examined the coverage of the FHIMS value set and UNMC corpus on these reactions.

We found that FHIMS included 5 out of 15 of the rare or severe reaction concepts, while UNMC included 6 out of 15 concepts (Table 4). Common among all were SJS and serum sickness, methemoglobinemia, and neuroleptic malignant syndrome. However, other important reactions, such as TEN and DRESS, while included in PEAR, were not present in either FHIMS or UNMC. Other important rare or severe reactions absent from FHIMS and UNMC included erythema nodosum, drug-induced hepatitis, fixed drug eruption, leukocytoclastic vasculitis, lichen planus, and aseptic meningitis.

Hypersensitivity reactions (HSRs)

We classified HSRs as *immediate* or *nonimmediate*. Immediate HSRs have a time to onset <1 h and are typically IgE-mediated, manifesting as urticaria, angioedema, rhinitis, conjunctivitis,

Table 3. Reasons for possible exclusion

Concept	Source ^a	Frequency in source (%)	Reason for possible exclusion
Acute relapsing multiple sclerosis	FHIMS	0	Chronicity not necessary per se, use multiple sclerosis concept
Traffic accident on public road	FHIMS	0	Event, not a clinical finding or disorder
Malignant tumor of breast	FHIMS	0	Diagnosis, not a reaction
Fracture of femur	FHIMS	0	Cause is unknown, eg, pathological fracture?, result of fall due to adverse reaction to medication?
Infectious disease	FHIMS	0	Very broad, nonspecific
Drug intolerance	FHIMS	0.0037	Classification of reaction, not a symptom of a reaction to an allergen
General health deterioration	FHIMS	0	Too broad to be useful, rarely documented
Patient's condition worsened	FHIMS	0	Unclear what reaction worsened
Traumatic or nontraumatic injury	FHIMS	0	Reaction type unclear – ? fall
Course of illness	FHIMS	0	Concept is an attribute concept in SNOMED-CT and should not be used to encode clinical information
Infection	UNMC	0.04	Too broad, rarely documented
Treatment not tolerated	UNMC	0.001	Reaction not specified, classified as intolerance

^aSource database: referring to FHIMS or UNMC databases.

Table 4. Presence of rare or severe reactions among value sets

Reactions in PEAR	Freq. in PEAR (%) ^a	FHIMS ^b	UNMC ^c
Acute interstitial nephritis	0.0212	Yes	No
Drug reaction with eosinophilia and systemic symptoms	0.0069	No	No
Drug-induced hepatitis	0.0027	No	No
Erythema multiforme	0.0195	No	Yes
Erythema nodosum	0.0072	No	No
Fixed drug eruption	0.0061	No	No
Leukocytoclastic vasculitis	0.0027	No	No
Lichen planus	0.0023	No	No
Meningitis (aseptic)	0.0040	No	No
Methemoglobinemia	0.0058	Yes	Yes
Neuroleptic malignant syndrome	0.0070	Yes	Yes
Pneumonitis	0.0073	No	Yes
Serum sickness	0.0439	Yes	Yes
Stevens-Johnson syndrome	0.0685	Yes	Yes
Toxic epidermal necrolysis	0.0104	No	No

^aPEAR frequency: number of reaction entries for a concept divided by the total number of reactions in PEAR ($n = 1\,647\,390$), excluding unknown reactions.

^bFHIMS: showing the presence or absence of severe reactions in FHIMS value set.

^cUNMC: showing the presence or absence of severe reactions in UNMC reaction list.

bronchospasm, or anaphylaxis.³⁵ Nonimmediate HSRs have a time to onset >1 h and are commonly T-cell mediated, manifesting with cutaneous symptoms including late-onset urticaria, maculopapular eruptions, fixed drug eruptions, vasculitis, TEN, SJS, or DRESS.^{21,22}

Within PEAR, HSRs represented 19.2% of reactions ($n = 150$ concepts). Common HSRs are shown in Figure 3, classified as immediate and nonimmediate.

Creation of an integrated value set

We compiled an integrated value set that combines PEAR-VS with nonexact matches found in either FHIMS or UNMC. We (FG, KW, LK) manually reviewed each concept, removed duplicates, and removed excluded reactions listed in Table 3.

PEAR-VS includes 787 concepts, to which we added 219 and 150 concepts from FHIMS and UNMC, respectively, resulting in 1156 candidate concepts. With duplicates ($n = 37$) and excluded reactions ($n = 12$) removed, the resulting value set included 1107 concepts within the clinical findings and disorder hierarchies. These 1107 concepts constituted our integrated adverse reaction value set,

of which 1103 concepts were mapped to SNOMED-CT and harmonized with those in FHIMS and UNMC.

DISCUSSION

We created a reaction value set by processing and analyzing reaction entries contained within a large enterprise-wide allergy repository collected from over 2 decades of documented reaction records using a semiautomated approach, combining computer-assisted coding with NLP and manual review. PEAR-VS covered 63% and 75% of reactions by exact match, representing 97.5 and 98.7% of reactions by frequency in FHIMS and UNMC, respectively. Partial matches represented 17%–20% of reactions, and the percentage of missing reactions ranged from 6% to 16%, with a higher number of missing terms in FHIMS compared to UNMC. An integrated value set of 1107 reaction concepts comprising PEAR-VS, FHIMS, and UNMC was curated to include missing concepts and partial matches, and to remove ambiguous or duplicative concepts. This work demonstrates how we processed a large amount of raw EHR data (in both

HYPERSENSITIVITY REACTIONS	
IMMEDIATE ^a	NON-IMMEDIATE ^b
Anaphylaxis	Serum sickness
Hypotension	Drug rash eosinophilia and systemic symptoms syndrome
Angioedema	Erythema multiforme
Swelling	Acute interstitial nephritis
Rhinitis	Fixed drug eruption
Bronchospasm	Erythema nodosum
Asthma	Stevens-Johnson syndrome
Wheezing	Toxic epidermal necrolysis
Shortness of breath	Pneumonitis
Hives	Meningitis
Urticaria	Lichenoid drug eruptions
Itching	Leukocytoclastic vasculitis
Rash	

Figure 3: Hypersensitivity reactions (HSRs).

^aImmediate HSRs: Immediate HSRs (time to onset <1h) are typically IgE-mediated.

^bNonimmediate HSRs: Nonimmediate HSRs (time to onset >1h) are commonly T-cell mediated, but may be antibody- or immune complex-mediated.

structured and free-text form), mapped these raw data to a standard terminology, and analyzed the aggregated data to generate a value set, based on patterns, frequencies, and importance of the information as well as enriching it using existing open sources. Our data align with National Quality Forum recommendations for value sets, and are consistent with data found in the patient record, integrated with the most current terminology updates, and aligned with current data models needed to support clinical decision support. Our final integrated reaction value set, we believe, captures the most frequently documented concepts, including severe reactions and HSRs, and will be valuable for supporting adverse reaction documentation and allergy-related clinical decision support.

SNOMED-CT for encoding reactions

We encountered several challenges during our evaluation, including defining the appropriate hierarchy and reconciling ambiguous or duplicative concepts. With hierarchies, for example, a reaction to an allergen might include the *clinical finding* of irregular heartbeat or the specific *disorder* of cardiac arrhythmia. While either might be correct, it would be desirable to default to one (eg, clinical finding over disorder) to avoid unintentional divergence in coding. Duplicative concepts were another challenge where there could be 2 concepts with similar strings but with different concept IDs. For example, the concept “red eye” was represented by 2 different concept IDs, one that referred to red eye as a “Finding of general observation of [an] eye” and one that referred to red eye as an “Ill-defined disorder of [the] eye.” While debatable, the former was thought to be a better concept, both in its ontological representation and by its frequency of documentation (825 entries vs 0 entries). While nearly every reaction term in our lexicon had a corresponding concept ID, some gaps did exist. A small number of reactions that were frequently documented were not present in SNOMED-CT, including “itchy tongue,” “sores in throat,” “throat tingling,” and “tingling in throat.” Based on their frequency of occurrence, we would suggest adding these concepts to SNOMED CT using the

concept “throat irritation” with the synonym “throat tingling” or “tingling in throat” and creating a new concept, “tongue irritation,” with the potential synonyms “tongue itching,” “tongue tingling,” and “tongue burning.”

Value set comparison

Overlap between PEAR-VS with FHIMS and UNMC reactions ranged between 82.9% and 92.7%. Manual review proved to be critical, as concepts could be similar at the *term* level with different concept IDs, resulting in loss of content and coverage. One gap identified was in the coverage of rare or severe reactions, which was notably lower in FHIMS and UNMC compared to PEAR-VS. These findings underscore the need for value sets to include rare or severe reactions, even though they may be low in frequency, highlighting the limitation of using frequency alone as a criterion for value set creation. With full representation of rare or severe concepts in SNOMED-CT, we would suggest including them in a reaction value set, given their importance to patient safety and adverse event avoidance. Comparison of value sets allowed identification of gaps present in each, their missing or similar concepts, and the frequency of each concept’s use. After adding partial and missing concepts, the integrated value set added 507 concepts to FHIMS, suggesting that more than half of the content between reaction entries across multiple organizations is at least similar at the *term* level. At the ontology level, future efforts will be needed to ensure consistency among concept IDs and perhaps default to one SNOMED-CT hierarchy to avoid divergence in coding. Computer-assisted coding leveraging NLP could assist in this process in addition to automating value set creation and syntactic classification (ie, hypersensitivity reactions, severe reactions, or intolerances), or exploring causal relationships between allergens and their associated reactions.

Applications to clinical decision support

With current override rates of 90% for inpatient and 77% for outpatient allergy alerts,^{36–39} redesigning CDS for allergy alerting could not be more important.⁴⁰ Value sets for encoding reactions provide the necessary discrete data to realign the type of CDS alert for a potential allergy or adverse reaction. This is particularly true for immune-mediated responses, which are associated with high risk. We found that nearly 20% of reactions within our value set were HSRs (immediate and nonimmediate). This estimate may be conservative, as reactions by themselves may not represent HSRs but in combination with other reactions can be HSRs. For example, the gastrointestinal symptoms nausea, vomiting, and diarrhea in combination with shortness of breath may represent anaphylaxis (an HSR). Serum sickness can be described by the combination of fever, myalgias, arthralgias, and rash. Surveillance for reaction documentation patterns can help clinicians identify rare but important reactions. Better understanding of the immunologic nature of reactions (eg, HSRs) can help inform the type of alert to be presented to clinicians.

Documentation of the frequency of allergies can also be valuable for CDS. With known frequencies of reactions in PEAR and other value sets (FHIMS, UNMC), EHRs could conceivably populate a dynamic reaction quick-pick list based on the most commonly associated reactions for a particular allergen. Using a data-driven approach to populate the reaction list could help limit the amount of time clinicians spend searching for a specific reaction, improve the accuracy of documentation, and reduce inappropriate downstream alerting. The ideal EHR CDS module would know the most

probable reactions for a given allergen; tier the allergy alerts based on the reaction entered, severity, and prior results from allergy testing (skin tests, challenge results); and allow exceptions to account for medications previously tolerated or not cross-reactive. Contraindications would be handled using simple rule-based advisories that could be triggered from any discrete data in the patient's record, be it religious preference, surgical history, or a problem on the patient's problem list. Robust value sets for encoding patient reactions to culprit agents (or "allergens"), we believe, are key to developing the infrastructure necessary to achieve a more intelligent advisory and alerting system for allergy and adverse reactions, limiting alert fatigue and reducing override rates.

Limitations

This work may be limited by the particular demographic of the population, which was primarily localized to the New England area. Allergens, particularly food and environmental allergens, in this area may be different than those in other localities. The frequency of rare or severe reactions may be higher in our data due to the presence of 2 large tertiary referral centers, where transfers are often initiated from outside facilities for specialty care (burn consults, dermatology consults, etc.). While our reaction value set contains the majority of reactions within PEAR, our lexicon was limited to those types of adverse reactions documented as free-text in the allergy module of the EHR. Other types of adverse reactions contained outside the EHR or in clinical notes may differ, and could enrich the reactions documented within the allergy module.

CONCLUSION

We processed and encoded reactions contained within a large allergy repository to harmonize, validate, and inform maintenance of the Adverse Clinical Reaction value set by the National Library of Medicine VSAC for interoperable use by EHRs. Our value set in addition to FHIMS will provide new insights into reaction and culprit agent associations, innovative CDS solutions for designing more intelligent alerting algorithms, and improved allergy documentation for adverse reactions.

FUNDING

This study was funded by Agency for HealthCare Research and Quality grant R01HS022728.

COMPETING INTERESTS

None.

CONTRIBUTIONS

All authors contributed substantially to the conception and design of this work and its data analysis and interpretation, and helped draft and revise the manuscript. All authors are accountable for the integrity of this work.

ACKNOWLEDGMENTS

Thanks to James Campbell, MD, for sharing the UNMC reaction data. We also thank George Robinson, Shelly Spiro, and Robert McClure for their valuable advice on this study.

REFERENCES

1. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*. 1998;279(15):1200–05.
2. McDonnell PJ, Jacobs MR. Hospital admissions resulting from preventable adverse drug reactions. *Ann Pharmacother*. 2002;36(9):1331–36.
3. Patel DA, Holdford DA, Edwards E, Carroll NV. Estimating the economic burden of food-induced allergic reactions and anaphylaxis in the United States. *J Allergy Clin Immunol*. 2011;128(1):110–15.
4. Gomes ER, Demoly P. Epidemiology of hypersensitivity drug reactions. *Curr Opin Allergy Clin Immunol*. 2005;5(4):309–16.
5. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. *Br J Clin Pharmacol*. 2001;52(1):77–83.
6. Gandhi TK, Weingart SN, Borus J, et al. Adverse drug events in ambulatory care. *N Engl J Med*. 2003;348(16):1556–64.
7. American Academy of Allergy, Asthma and Immunology. *Medications and drug allergic reactions*. www.aaaai.org/conditions-and-treatments/library/at-a-glance/medications-and-drug-allergic-reactions. Accessed September 25, 2016.
8. Blumenthal KG, Shenoy ES, Hurwitz S, Varughese CA, Hooper DC, Banerji A. Effect of a drug allergy educational program and antibiotic prescribing guideline on inpatient clinical providers' antibiotic prescribing knowledge. *J Allergy Clin Immunol Pract*. 2014;2(4):407–13.
9. Tanner A. *Our Bodies, Our Data: How Companies Make Billions Selling Our Medical Records*. Boston: Beacon Press; 2017.
10. Tanner A. Patient power through records. *Boston Globe*. July 1, 2017.
11. National Quality Forum. *Value set harmonization*. March 2016. www.qualityforum.org/ProjectDescription.aspx?projectID=78591. Accessed October 4, 2017.
12. WHO-ART – WHO Adverse Reaction Terminology. www.unc-products.com/DynPage.aspx?id=73589&mn1=1107&mn2=1664. Accessed April 12, 2017.
13. International Conference on Harmonization. www.ich.org/. Accessed April 12, 2017.
14. Linder JA, Haas JS, Iyer A, et al. Secondary use of electronic health record data: spontaneous triggered adverse drug event reporting. *Pharmacoepidemiol Drug Safety*. 2010;19(12):1211–15.
15. FDA Sentinel Initiative. www.fda.gov/Safety/FDASentinelInitiative. Accessed April 12, 2017.
16. Avillach P, Coloma PM, Gini R, et al. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. *J Am Med Inform Assoc*. 2013;20(1):184–92.
17. US Department of Health and Human Services. *International Classification of Diseases, 9th Revision, Clinical Modifications (ICD-9-CM)*. 6th ed. Washington, DC: DHHS Publication No. (PHS) 96–1260; 1997.
18. Hougland P, Xu W, Pickard S, Masheter C, Williams SD. Performance of International Classification of Diseases, 9th Revision, Clinical Modification codes as an adverse drug event surveillance system. *Med Care*. 2006;44(7):629–36.
19. Hougland P, Nebeker J, Pickard S, et al. Using ICD-9-CM codes in hospital claims data to detect adverse events in patient safety surveillance. In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment)*. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
20. National Library of Medicine. *Value Set Authority Center*. 2015. https://vsac.nlm.nih.gov/. Accessed April 12, 2017.
21. Pichler WJ. Delayed drug hypersensitivity reactions. *Ann Intern Med*. 2003;139(8):683–93.
22. Posadas SJ, Pichler WJ. Delayed drug hypersensitivity reactions – new concepts. *Clin Exp Allergy*. 2007;37(7):989–99.
23. American Academy of Allergy, Asthma, and Immunology. *Health Information Technology*. www.aaaai.org/practice-resources/running-your-practice/practice-management-resources/health-information-technology. Accessed September 25, 2016.

24. HL7 Working Group. *Allergy and intolerance*. March 20, 2012. http://wiki.hl7.org/index.php?title=Allergy_%26_Intolerance. Accessed March 22, 2017.
25. Donnelly K. SNOMED-CT: The advanced terminology and coding system for eHealth. *Stud Health Technol Inform*. 2006;121:279–90.
26. US Pharmacopeial Convention. *2015–2020 Healthcare Quality and Safety Committee Work Plan*. www.usp.org/healthcare-quality-committee-work-plan. Accessed September 28, 2016.
27. Zhou L, Plasek JM, Mahoney LM, *et al*. Using Medical Text Extraction, Reasoning and Mapping System (MTERMS) to process medication information in outpatient clinical notes. *AMIA Annu Symp Proc*. 2011;2011:1639–48.
28. Goss FR, Plasek JM, Lau JJ, Seger DL, Chang FY, Zhou L. An evaluation of a natural language processing tool for identifying and encoding allergy information in emergency department clinical notes. *AMIA Annu Symp Proc*. 2014;2014:580–88.
29. Lai KH, Topaz M, Goss FR, Zhou L. Automated misspelling detection and correction in clinical free-text records. *J Biomed Inform*. 2015;55:188–95.
30. Plasek JM, Goss FR, Lai KH, *et al*. Food entries in a large allergy data repository. *J Am Med Inform Assoc*. 2016;23(e1):e79–87.
31. Acker WW, Plasek JM, Blumenthal KG, *et al*. Prevalence of food allergies and intolerances documented in electronic health records. *J Allergy Clin Immunol*. 2017. [Epub ahead of print]
32. Kuperman GJ, Marston E, Paterno M, *et al*. Creating an enterprise-wide allergy repository at Partners HealthCare System. *AMIA Annu Symp Proc*. 2003;376–80.
33. SNOMED-CT (Systematized Nomenclature of Medicine – Clinical Terms). www.ihtsdo.org/snomed-ct. Accessed March 22, 2017.
34. Baeza-Yates R, Ribeiro-Neto B. *Modern Information Retrieval*. New York: ACM Press; 1999.
35. Sanchez E, Torres MJ, Mayorga C, *et al*. Adverse drug reactions with an immunological basis: from clinical practice to basic research. *Allergy*. 2002;57 (Suppl 72):41–44.
36. Genco EK, Forster JE, Flaten H, *et al*. Clinically inconsequential alerts: the characteristics of opioid drug alerts and their utility in preventing adverse drug events in the emergency department. *Ann Emerg Med*. 2016;67(2):240–48 e3.
37. Topaz M, Seger DL, Lai K, *et al*. High override rate for opioid drug-allergy interaction alerts: current trends and recommendations for future. *Stud Health Technol Inform*. 2015;216:242–46.
38. Topaz M, Seger DL, Slight SP, *et al*. Rising drug allergy alert overrides in electronic health records: an observational retrospective study of a decade of experience. *J Am Med Inform Assoc*. 2016;23(3):601–08.
39. Slight SP, Beeler P, Seger DL, *et al*. Override rates of drug allergy alerts in inpatient and outpatient settings. *BMJ Qual Saf*. 2016;97:1–9.
40. Topaz M, Goss F, Blumenthal K, *et al*. Towards improved drug allergy alerts: multidisciplinary expert recommendations. *Int J Med Inform*. 2017;97:353–55.