Lessons Learned from Developing a Drug Evidence Base to Support Pharmacovigilance

J.C. Smith[1]; J.C. Denny[1,2]; Q. Chen[1,3]; H. Nian[3]; A. Spickard III[1,2]; S. T. Rosenbloom[1,2,4,5]; R. A. Miller[1,2,4]

[1]Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA;
[2]Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA;
[3]Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA;
[4]School of Nursing, Vanderbilt University, Nashville, Tennessee, USA;
[5]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA;

Summary

Objectives: This work identified challenges associated with extraction and representation of medication-related information from publicly available electronic sources.
Methods: We gained direct observational experience through creating and evaluating the Drug Evidence Base (DEB), a repository of drug indications and adverse effects (ADEs), and supplemented this through literature review. We extracted DEB content from the National Drug File Reference Terminology, from aggregated MEDLINE co-occurrence data, and from the National Library of Medicine's DailyMed. To understand better the similarities, differences and problems with the content of DEB and the SIDER Side Effect Resource, and Vanderbilt's MEDI Indication Resource, we carried out statistical evaluations and human expert reviews.
Results: While DEB, SIDER, and MEDI often agreed on medication indications and side effects, cross-system shortcomings limit their current utility. The drug information resources we evaluated frequently employed multiple, disparate vaguely related UMLS concepts to represent a single specific clinical drug indication or adverse effect. Thus, evaluations comparing drug-indication and drug-ADE coverage for such resources will encounter substantial numbers of false negative and false positive matches. Furthermore, our review found that many indication and ADE relationships are too complex – logically and temporally – to represent within existing systems.
Conclusion: To enhance applicability and utility, future drug information systems deriving indications and ADEs from public resources must represent clinical concepts uniformly and as precisely as possible. Future systems must also better represent the inherent complexity of indications and ADEs.

Correspondence to:
Joshua C. Smith
Eskind Biomedical Library – B003D
2209 Garland Avenue
Nashville, TN 37232–8340
joshua.smith@vanderbilt.edu
Phone: (615) 936–3690
Fax: (615) 936–5900

1. Introduction

Pharmacovigilance projects and clinical decision support systems require comprehensive, authoritative databases of medication indications and known adverse drug effects (ADEs). Deriving a high quality, widely-available medication database from publically available resources would benefit clinicians and researchers [1–4]. This paper examines problems with current drug information resources that future system developers must overcome to create widely disseminated useful non-commercial drug indication and side effect databases. Existing resources for such information are incomplete, variably out of date, in disagreement with one another [2, 5], or fail to store drug information in structured computable formats [5].

The authors have begun development of a pharmacovigilance system using electronic medical record (EMR) data, analogous to research efforts at the FDA [6, 7], Columbia University [8, 9], Stanford University [4, 10], and elsewhere. Through developing the first component of our system, the Drug Evidence Base (DEB), the authors have gained insight into the difficulties associated with extracting, combining, and representing imperfect drug knowledge from multiple public sources. By examining the evaluations and knowledgebase construction processes of our own and others' similar systems, we have compiled a list of challenges that current systems must overcome to gain widespread applicability.

2. Background

2.1 Motivation

Previous projects have compiled knowledge from multiple drug sources to address specific clinical informatics goals. In 2010, Wang, et al. [2], compiled drug indication information for use in automated pharmacovigilance and decision support systems from a combination of sources: the FDA Adverse Effect Reporting System (AERS), NDF-RT, and SemMed – a database generated from natural language processing (NLP) on MEDLINE abstracts. For a set of 20 drugs, they extracted indication knowledge comparable to a manually curated gold standard. In 2011, Li, et al. [3], expanded on Wang's work to determine the ability of combined indication resources in identifying medication indication in EMRs. They applied information from Micromedex, NDF-RT, and the AERS to infer the reasons for prescriptions for drugs mentioned in EHR discharge summaries, revealing promising results but only focusing on a limited sample of six drugs. In 2013, Wei, et al. [11], developed MEDI, a medication indication resource linking data from RxNorm, SIDER (defined below), MedlinePlus, and Wikipedia. They further refined their work to include a high-precision subset of indications retrieved from RxNorm or at least two out of three sources.

Previous projects have demonstrated potential utility of using EMRs to discover known and novel correlations among drugs, indications, and ADEs [4, 8–10]. While accurate reference standards are critical to development and evaluation of pharmacovigilance systems, most "gold standards" for pharmacologic studies have been created solely for specific individual investigations. Wang, et al. [8], illustrated the feasibility of using NLP on EMRs to identify potential drug-ADE associations. However, the reference standard was created using only a single expert, and contained only seven drugs. Tatonetti, et al., developed their own reference standard of drug effects and drug-drug interactions from significant associations found in the FDA AERS database [10]. LePendu, et al., adjusted for known confounders to identify associations for several known ADEs, also recognizing that these associations could have been discovered before FDA action. However, the manually created reference standard contained only 12 ADEs of interest and 78 drugs [4]. It is possible to confirm significant associations that have already been discovered with a small reference standard, but a comprehensive database of drug indications and adverse effects would benefit the hypothesis-free examination of a large EMR corpus. An accurate reference standard is necessary for both prioritizing ADE associations under consideration and enabling prospective evaluation of signal detection methods [12, 13].

2.2. The DEB Project as a Basis for Authors' Observations

The authors' envisioned future pharmacovigilance system will identify correlations between drugs and clinical findings in EMR-based clinical notes. This goal required creation and evaluation of the Drug Evidence Base (DEB), a resource to catalog known drug relationships so that potentially novel associations could be recognized. The DEB algorithmically combines data from multiple publicly available drug information sources to derive known drug-indication and drug-ADE pairs. The DEB information sources include the US Department of Veterans Affairs National Drug File Reference Terminology (NDF-RT) [14], information contained in the biomedical literature, represented using MEDLINE [15] data in the Unified Medical Language System (UMLS) [16] co-occurrence of concepts table (MRCOC), and structured product labels (SPLs) from the US Food and Drug Administration (FDA) [17]. Through the process of developing the DEB, and via statistical evaluations and human expert reviews, we gained valuable insights into the similarities, differences, and problems with the content of DEB, SIDER, and Vanderbilt's MEDI.

2.3 Drug Knowledge Resources

Commercial resources, including Micromedex®, First Databank®, and ePocrates®, among others [18–21], contain drug indications and ADEs, and additional information, including drug-drug interaction and dosing instructions. However, commercial medication databases are expensive, not widely available, and often lack published validations. They vary in scope, content, and reliability; are often not available in formats suitable for computational processing, and have unknown frequencies of updating [1–3, 22]. While no single, comprehensive drug knowledge source yet exists [1, 3], useful subsets of drug information are available from frequently updated, public resources [5].

Within UMLS, drug knowledge resources include RxNorm, NDF-RT, and the MRCOC Co-Occurrence of Concepts table. RxNorm provides a medication nomenclature, and mappings among drug concepts, dose forms, brand names, and generic ingredients [23, 24]. The NDF-RT provides formal structured representations for medications, including ingredients, dose forms and classification, physiologic effects, mechanisms of action, and relationships such as indications and ADEs [14]. The NLM's MEDLINE topically indexes millions of journal articles using MeSH (Medical Subject Headings) [15], and also provides abstracts when available. The UMLS MRCOC table captures aggregations of co-occurrences of MeSH concepts in MEDLINE-indexed articles from the prior ten years [25, 26]. Researchers have extracted drug knowledge from MEDLINE using co-occurrences of MeSH terms, text mining, and other automated methods [27–32]. The NLM's DailyMed website provides access to the FDA's XML-formatted Structured Product Labels (SPLs) for prescription drugs sold in the United States [17]. The SPLs include approved indications, known ADEs, and potential drug-drug interactions. The content of many SPL sections (e.g., *Adverse Effects, Indications and Usage*, etc.) occur as blocks of unstructured text. The SIDER Side-Effect Resource is a public research database primarily containing ADEs text-mined from the SPLs [33, 34].

Lacking a comprehensive "gold" standard for drug-indication and drug-ADE pairs [1, 2], we evaluated the DEB by comparing its content to two existing reference standards of convenience – a version of the European SIDER Side-Effect Resource [33], an ADE database mined from FDA SPLs, and the Vanderbilt-developed MEDI Indication Resource [11], a collection of drug indications extracted from multiple sources. We enlisted expert clinician reviewers to evaluate both the accuracy and shortcomings of DEB, SIDER, and MEDI. We believe a detailed accounting of the problems facing current drug knowledgebases is essential to guide future work extracting, combining, and representing drug information for biomedical informatics applications.

3. Methods

3.1. Definition of DEB Terms

The DEB comprises a knowledge base of drug indications and ADEs. Within the DEB, a "clinical manifestation" (CM) can represent a disease or finding that is an indication for drug therapy or an

adverse effect of drug therapy. The project defined an indication (IND) relationship as "drug treats or prevents *CM*" and an ADE relationship as "drug predisposes to, causes, or exacerbates *CM*."

The project defined a *drug* as any UMLS concept that had at least one of the following UMLS semantic types: *Antibiotic, Pharmacologic Substance*, or *Clinical Drug*. A CM includes any IND or ADE concept related to drug administration that appears in clinical text. The project constrained *CM terms* to UMLS concepts having at least one of the following semantic types: *Anatomical Abnormality, Injury or Poisoning, Congenital Abnormality, Finding, Sign or Symptom, Acquired Abnormality, Clinical Attribute, Disease or Syndrome, Mental or Behavioral Dysfunction, Neoplastic Process*, and *Pathologic Function*. A *drug-CM* pair involves a *drug* CUI and a *CM* CUI connected by an IND or ADE relationship.

3.2. Extracting Drug-CM Pairs from NDF-RT

The DEB algorithms extracted all NDF-RT entries in the UMLS Relationships (MRREL) table where the row contained a *drug* and a *CM* and had at least one of the following asserted relationships: "has physiologic effect" or "induces" (indicating a potential ADE), and "may prevent" or "may treat" (indicating a likely IND). The system stored data extracted from each source in a MySQL database and combined that data into the full DEB (described below). ▶Figure 1 illustrates DEB construction process and results.

3.3. Extracting Drug-CM Pairs from MRCOC

The DEB algorithms extracted all MRCOC table entries representing the co-occurrence of a *drug* and a *CM* in indexed journal articles. The DEB retained pairs where the entry contained at least one of following relevant MeSH subheadings: Adverse Effect (AE) or Therapeutic Use (TU) qualifying a *drug*, and Drug Therapy (DT) or Etiology (ET) qualifying a *CM*. The combination (**drug**/*TU* + **CM**) and (**drug** + **CM**/*DT*) together implies an IND relationship; the combination (**drug**/*AE* + **CM**) and (**drug** + **CM**/*ET*) together implies an ADE. To exclude unfounded relationships, the DEB only retained drug-CM pairs that had co-occurred in at least four MEDLINE-indexed articles.

3.4. Extracting Drug-CM Pairs from SPLs

Available SPLs for human prescription drugs from DailyMed comprised ~56,000 entries associated with approximately 2400 unique medications [35]. From the SPL Data Elements index file, DEB algorithms selected single-ingredient human prescription drugs and automatically mapped drug ingredient or brand names to CUIs by regular expression string matching. For strings that did not match exactly, the DEB matched partial names (for example, "Fluoxetine Hydrochloride" mapped to "Fluoxetine"). We manually reviewed these matches to confirm accuracy, and corrected DEB algorithms to eliminate mismatches or unmatched terms whenever possible.

For every DailyMed drug that mapped to a CUI, the DEB algorithms parsed the corresponding SPL structure, extracting the "Adverse Reactions" and "Indications and Usage" sections (when present). The DEB algorithms used the KnowledgeMap Concept Identifier (KMCI), a Vanderbilt-developed general NLP tool [36–38], to identify all unique CM concepts in those sections, discarding any negated CMs (e.g., "no fever"). In order to capture all relevant concepts, we did not restrict KMCI to any particular UMLS source vocabularies. Drugs with multiple SPLs had their extracted information combined. Concepts extracted from "Adverse Reactions" were tagged as ADEs and those from "Indications and Usage" were tagged as INDs.

3.5. Integrating Source Information into the DEB

Multiple unique drug concept CUIs in the UMLS used by DEB refer to the same generic drug (for example, *C0000970 – Acetaminophen, C0699142 – Tylenol*, and *C1640784 – Tylenol 160 mg*). The DEB algorithms normalized such drug concepts by mapping each to the CUI corresponding to its main or active generic ingredient. From RxNorm and NDF-RT, the algorithms used the "ingredient of" and "has ingredient" relationships from the UMLS MRREL table to map dose forms of a drug to

the drug name, and used the "tradename of" and "has tradename" relationships to map brand names to corresponding generic drug CUIs.

Through an iterative process, the authors developed a differential scoring system for IND and ADE pairs extracted from the three sources (NDF-RT, MRCOC, SPL), calculating both an "IND score" and "ADE score" for each pair. We weighted the NDF-RT highest, as a trusted knowledge source that directly indicates whether a drug-CM pair comprises an IND or an ADE. A drug-CM pair validated by NDF-RT contributed 10 points to DEB's ADE or IND score. We weighted MRCOC second because it captures expert NLM indexers' specific assessments of peer-reviewed literature contents. The exact MRCOC weight in DEB involved the ratio of co-occurrences suggesting either an ADE or IND. For example, if MRCOC indicated that a drug-CM pair co-occurred in 12 articles, 6 of which implied an ADE and 3 of which implied an IND, $(6/12*10) = 5$ points were added to the AE score and $(3/12*10) = 2.5$ points were added to the IND score. We weighted the SPL component lowest because, independent of the authoritativeness of each SPL entry per se, the extraction of SPL information via NLP made it potentially less reliable. Furthermore, SPLs mention all findings reported during preliminary drug evaluation clinical trials, many of which lack subsequent validation Thus, DEB assigned five points for drug-CM pair categorizations supported by SPLs. The sum of the scores from all three sources determined the final DEB drug-CM classification. Thus, if the sum of the IND scores was higher than the sum of the ADE scores for a given pair, the drug-CM pair was classified as an indication (and vice versa).

The authors recognize that a CM concept can represent both an ADE and an IND for a specific drug (e.g., warfarin is used for stroke prophylaxis for atrial fibrillation, but excessive dosages can cause intracerebral hemorrhage). Nevertheless, the current version of DEB only classifies a drug-CM pair as ADE or IND, and not "both." When IND and ADE scores were equal, the DEB algorithm assigned an IND classification.

3.6. DEB Implementation

►Figure 2 illustrates a portion of the reconciled, extracted DEB data for Warfarin. The DEB algorithms used Perl scripts to extract and reconcile the data from each DEB source, storing intermediate and final results into a MySQL database. To maintain compatibility with the UMLS version used by KMCI at the time of this study, the project used UMLS version 2009AA and corresponding CUIs throughout.

3.7. Evaluation

3.7.1 DEB Evaluation

Expert reviewers evaluated the accuracy of DEB drug-CM pairs. Four Vanderbilt faculty physician reviewers (JCD, AS, STR, RAM), each board-certified in internal medicine with at least 10 years of clinical experience, rated 125 pairs each. The pairs were randomly selected from those in DEB but not in SIDER; 25 pairs were identical across all reviewers to enable calculation of inter-rater agreement. Reviewers were blinded to the DEB categorizations of the pairs. The experts received instructions to mark each pair's relationship as ADE, IND, both (indicating the CM concept could reasonably represent either an ADE or IND for the given drug), or neither. Experts optionally could comment freehand about any relationship.

3.7.2. Comparison of DEB with SIDER

Due to the lack of a widely agreed upon gold standard resource for drug indications and adverse effects, the DEB evaluation selected SIDER (downloaded January 2012) as an external reference standard of convenience. After completion of our study, an updated version of SIDER appeared; the comparisons reported herein do not involve SIDER2. The study converted SIDER entries into the same format as DEB (drug-CM pairs). This mapping went from SIDER drug names to UMLS drug CUIs (as with the SPLs); CM concepts in SIDER were already stored as UMLS CUIs. If SIDER classified a drug-CM pair as both ADE and IND, the evaluation considered it an IND (as was done for DEB).

For drug-CM pairs present in both SIDER and DEB, we determined whether both resources categorized the relationship in the same way, or if they differed. Among drug-CM pairs that differed, the four reviewers each categorized the relationships for 75 different pairs (per above).

3.7.3. Quantitative Evaluation

To evaluate reviewers' pairwise inter-rater agreement, we calculated Cohen's Kappa using the 25 common drug-CM pairs. We used Fleiss' Kappa to measure agreement across all four reviewers. We also calculated Kappas separately for those drug-CM pairs on which no reviewer had indicated a "both" relationship (to determine agreement on "non-ambiguous" pairs). We calculated 95% confidence intervals for the estimated Kappas based on z statistics [39, 40]. For 75 randomly chosen pairs where DEB-SIDER disagreed (IND versus ADE) expert ratings used "IND", "ADE", "both" or "neither." We evaluated DEB versus SIDER performance by calculating how often experts "agreed-with-DEB" versus the proportion they "agreed-with-SIDER". For each reviewer, the 95% confidence intervals of the differences between the two proportions (multinomial distribution) were constructed using z statistics. The 95% confidence interval for the averaged difference over the reviewers was built using a bootstrap approach with 5000 resamples. To evaluate the performance of DEB, reviewers rated 100 random pairs and their ratings were categorized as "agreed with DEB", "disagreed with DEB", "both" and "neither". We reported the proportions of each of the four categories as well as the ratio of "agreed" over "disagreed". The corresponding confidence intervals were constructed using the bootstrap approach with 5000 resamples as well. All the analyses were performed using statistical software R 2.15.2 [41].

3.7.4. Qualitative Evaluation

To identify common themes regarding problems with DEB, we compiled and compared expert reviewers' comments about DEB and SIDER. Additionally, one reviewer (RAM) empirically analyzed all DEB and SIDER drug-CM pairs for the drug *abacavir*. After our original study was completed, Wei, et al., published an article describing the MEDI Indication Resource [11]. To further elucidate difficulties in compiling drug knowledge from multiple sources, we performed a similar qualitative comparison of all DEB and MEDI indications for a single drug, *metoprolol*.

3.8 Ability to Automatically Update the DEB

Since more than a year has elapsed between initial construction of the DEB and the present time, we have had the opportunity to evaluate the ease of updating DEB using new releases of the DEB primary data sources. Specifically, we replaced the original 2009 version of the UMLS sources with the 2013AA release and previous version of SPLs (downloaded August 2011) with SPLs downloaded October 2013 from DailyMed. We performed all work (except for processing the SPLs) on a MacBook Pro with a 2.6 GHz Intel Core i7 processor with 16 GB of RAM. To process the SPLs, we ran KMCI on a Linux server with forty-eight 2.2 GHz AMD Opteron cores and 256 GB RAM. Note that the UMLS MRCOC tables use a moving 10-year window for reported co-occurrences of concepts, so that "previously known" drug-CM pairs derived from older versions of MRCOC may drop out when one uses newer versions of MRCOC, and conversely, new drug-CM pairs will occur in newer versions of MRCOC. Thus, newly "discovered" (from the October 2013 DEB version) DEB drug-CM pairs were combined with those already present in the earlier version of the DEB.

4. Results

4.1. Quantitative Evaluation Results

After representing drugs as their generic ingredients, the DEB algorithms extracted 149,197 unique drug-CM pairs from the three DEB sources, consisting of 3291 drugs and 8579 CMs. ▶Figure 3 illustrates the intersection among drugs, CMs, and drug-CM pairs for the three DEB sources. Fully 144,532 pairs (97%) had data from just one source (distributed across multiple possible sources – i.e., not all came from the same one source), 4180 pairs (~3%) had data based on two knowledge

sources, and 485 pairs (<1%) came from all three sources. Overall, the DEB classified 85,610 pairs as ADEs and 63,587 pairs as INDs. A total of 84,174 ADEs and 60,358 INDs came from one source, 1430 ADEs and 2750 INDs came from two sources, and 6 ADEs and 479 INDs came from three sources.

Mapping SIDER information to the DEB format resulted in 66,612 SIDER drug-CM pairs consisting of 871 drugs and 1684 CMs. Both SIDER and DEB contained 37,335 overlapping drug-CM pairs, comprising 56% of SIDER and 25% of DEB. For those pairs, SIDER and DEB ratings agreed on 97% (as IND or ADE). ▶Figure 4 illustrates the intersection of Drug-CM pairs in DEB and SIDER.

The pairwise and group Kappas for experts' rating of 25 drug-CM pairs appear in ▶Table 1A. Since reviewer 2 tended to disagree with the others most often, Kappa among all other reviewers is shown to illustrate their consensus. ▶Table 1B shows reviewers' ratings for 300 drug-CM pairs where DEB and SIDER disagreed (comprising ~30% of all DEB-SIDER disagreements). Reviewers agreed with the DEB classification (IND or ADE) significantly more often than SIDER (P≤0.01 for the reviewer group as a whole). On average, reviewers agreed with DEB 30% more of the time for DEB-SIDER disagreements (95% CI, 20% - 40%). ▶Table 1C shows reviewer categorizations for 400 randomly selected DEB drug-CM pairs not in SIDER. Reviewers agreed with DEB categorizations 61% of the time on average; each reviewer agreed with DEB on more than 50% of the pairs and disagreed on less than 20% of the pairs. On average, reviewers were 9-fold more likely to agree with a given DEB categorization than to disagree (average ratio of agreed to disagreed was [4.88 + 4.73 + 7.25 + 19.00]/4 = 8.96; 95% CI, 5.6 - 20.9 fold). Note that DEB did not include "both" or "neither" for its IND/ADE categorizations, even though reviewers used them.

4.2. Qualitative Evaluation Results

▶Table 2 shows the reviewer's categorizations of the DEB and SIDER drug-CM pairs for one sample drug, abacavir. The reviewer indicated where discrepancies were due to use of differing terms/concepts for the same or similar CMs. For example, SIDER listed only *Acidosis (C0001122)* as an ADE while DEB more specifically listed *Lactic Acidosis (C0001125)*; the SPL only listed *Lactic Acidosis*. Similarly, only DEB listed *Myalgia (C0231528)* as an ADE, only SIDER listed *Arthralgia (C0003862)*, and both listed *Musculoskeletal Pain (C0026858).* Of apparent DEB-SIDER discrepancies in ▶Table 2, many involved low-information, non-specific CMs.

▶Table 3 excerpts reviewer comments for several drug-CM pairs with inter-reviewer disagreements. The table illustrates both the subjective nature of some drug-CM pairs and problems with CMs, including those that are not useful (e.g., the CM concept *Adverse Event*), too general (*Anti-Bacterial Agents* and *Heart Diseases*), and irrelevant (*Sheep Diseases*).

▶Table 4 shows indications for *metoprolol* that both DEB and MEDI identified. A physician reviewer determined if these listed indications could stand alone as valid clinical indications for the drug (i.e., if all that was known about the patient was the CM, would it be appropriate to prescribe *metoprolol*). Both DEB and MEDI captured all of the FDA approved indications from the package inserts (SPLs), as well as many accepted off-label uses. Both sources also included numerous redundant CMs, incorrect indications, and incomplete partial indications in the form of CMs that would be prescribed but only in certain situations. For example, *metoprolol* is indicated as prophylaxis for *ventricular fibrillation*, but only post *myocardial infarction*, and not as primary therapy for *ventricular fibrillation*.

4.3 DEB Update

After downloading and installing the UMLS 2013AA release and downloading the SPLs from DailyMed, updating DEB required less than 2 hours. This includes extracting MRCOC and NDF-RT relationships from the UMLS, parsing the SPLs to extract the "Indications" and "Adverse Reactions" sections, processing those sections using KMCI, and loading the data into a MySQL database. The updated version of DEB contained a total of 190,789 drug-CM pairs, with 3842 distinct drugs and 9693 distinct CMs. The updated DEB includes 112,651, 1897, and 9 ADE pairs and 72,223 plus 3366 plus 643 IND pairs from one, two, and three sources, respectively. This represents a 2% increase in

ADE pairs and a 15% increase in IND pairs within DEB, as well as a 27% increase in pairs from 2 or more sources. The new version contained 42,168 drug-CM pairs that were not previously in DEB. A total of 44 drug CUIs and 581 CM CUIs in the original DEB were deleted or merged in the newer release of the UMLS.

5. Discussion

The study results demonstrate the potential for a drug-CM repository compiled from multiple public sources to provide information that might support drug safety and pharmacovigilance projects. Nevertheless, the results document shortcomings that must be overcome before attempting any real world implementation. Problems the study uncovered include the necessity of collapsing similar and related CMs into a single "aggregate" CM, the need to eliminate inappropriate drug-CM relationships (e.g., pairs which are overly broad), and the need to understand and represent complex indications (e.g., "useful in preventing XYZ in the setting of condition ABC").

The most important result of our investigation involves an increased understanding of the desiderata for, and pitfalls related to, construction of drug information resources. At one level, our study results indicate that drug-CM pairs present in both DEB and SIDER had 97% agreement and that reviewers rated DEB correct 30% more often when the two disagreed. Previous studies in this area presented their results in similar terms [2, 3, 11]. But what do such results really mean? When comparing "indications" represented in DEB, SIDER, or MEDI, the most relevant question clinically is, "Should a physician prescribe the 'indicated' medication if all that is known about the patient is a randomly selected finding/disease term from the system's list of indications?" In our study, expert clinician reviewers rated *chest pain* as a correct indication for nitroglycerin with respect to DEB and MEDI. But *chest pain* is a much broader superset of the FDA-approved indication of *angina pectoris*. The undifferentiated term *chest pain* includes pain due to fractured ribs, which one should not treat with nitroglycerin. Researchers in this field must therefore develop more precise NLP algorithms to capture exactly the narrow indications listed in SPLs. Nevertheless, doing so would miss the large number of "off label" indications for commonly used drugs. Furthermore, a pharmacovigilance application that only "knew" about the very specific FDA-approved indications for nitroglycerin might "discover" chest pain as an unexplained side effect (or possible new off-label indication) for nitroglycerin – even for patients with *angina pectoris* documented elsewhere in their EMRs. Representing exact indications is critical, because *esophageal spasm*, another cause of chest pain, is an off-label indication for nitroglycerin. Future systems must represent specific indications and also the set of related (and usually broader) concepts that clinicians might use in EMRs to refer to the more specific indications.

Additionally, many indications are logically and temporally complex. For example, beta blockers such as metoprolol are indicated following *myocardial infarction* to prevent *ventricular fibrillation*. It would be improper during a cardiac arrest to state "the patient has ventricular fibrillation, give metoprolol now."

While DEB, SIDER, and MEDI agreed on the core indications for most medications, the intersections of indications and ADEs among drug knowledge sources were small. To state that one database or one source was superior based on the number of indication or adverse effect terms each contained could be highly misleading. If two drug information resources both listed *angina pectoris* as an indication for *nitroglycerin*, but one listed ten other variant and less specific terms for chest pain as well, it is not immediately clear which resource actually contains "better" information in the absence of a more in-depth analysis. In the DEB evaluation, we observed that similar (or the same) CMs were often represented by different UMLS concepts in different sources. Frequently, one source lists a specific concept and the other a more general concept. Future systems must represent narrow, specific indications for drugs explicitly, and then relate to each indication the multiple CMs that represent broader terms in common use for that single concept. While the knowledge that *angina* is a type of *chest pain* would allow automated methods to infer that *chest pain* mentioned in an EMR note might be a reference to *angina* if the latter was a previously established diagnosis, not all UMLS vocabularies define robust conceptual relationships. Automatically determining concept relatedness is a difficult problem.

Conversely, reliable methods to normalize similar CMs are also important. Within UMLS, closely related concepts have different CUIs. For example, the CUIs for *neutropenia* and *leukopenia*, and for *myalgia* and *musculoskeletal pain* are different; the closeness of their relationships is not captured. Systems such as DEB that depend on external terminologies such as UMLS will identify such pairs as two independent drug-CM relationships instead of overlapping ones. Additionally, CMs that include qualifiers have different CUIs as well, such as *moderate fever* and *fever, or acute myocardial infarction* and *acute myocardial infarction, site unspecified*. Normalizing these CMs to canonical "consensus" terms would potentially resolve many of the discrepancies between sources, and increase evidence for the particular relationship.

There are a variety of methods that attempt to accomplish these tasks. Methods range from traversing the UMLS MRREL and MRHIER to discover relationships, to walking up individual ontologies parent-child hierarchies, to sophisticated methods such as k-Neighborhood decentralization [42–46]. Most of these solutions, however, are partial and do not necessarily perform the required tasks in a uniformly reliable way. We did not attempt to normalize CM concepts in this study because doing so involves complexity greater than that of normalizing medication concepts. No perfect hierarchy of clinical terms (diseases and findings) exists. Under certain conditions, a concept should be normalized to a more or less specific concept, but in other situations, no changes are required. For example, CMs such as "obese" and "morbidly obese" might be merged to "obesity" when categorizing patients with acute sinusitis, but should remain separate when studying predictors for myocardial infarction [46]. We believe that the DEB should retain all CM concept categories prima facie, but that in specific applications, those concepts should be aggregated post-hoc, using NLP methods tailored to the particular task at hand. More work is needed to perfect current methods.

Additionally, a deeper representation system for indications and ADEs might help to resolve current ambiguities in determining whether a drug-CM pair comprises an indication, an ADE, both, or neither. The DEB clinician-reviewers often disagreed with one another, and their stated reasons for categorizing pairs a certain way provided insight. Some reviewers would rate the *clonidine-essential hypertension* pair as "indication" since that is an FDA-approved use mentioned in the package insert. Other reviewers, aware that abrupt clonidine discontinuation in hypertensive patients can potentially exacerbate hypertension, might rate the pair as "both IND and ADE". In another context, a clinician reviewing an EMR record where a patient was treated with clonidine for opiate withdrawal, and the patient developed hypertension for the first time when clonidine was discontinued, would rate the relationship as purely ADE for that case.

Reviewers classified almost one quarter of drug-CM pairs as neither indication nor ADE. We believe this represents noise in DEB, not necessarily incorrect associations. Pairs classified as neither were almost exclusively too broad (drug classes or CMs such as "cardiovascular diseases") or from animal studies indexed in MRCOC (the CM "sheep diseases"). This represents noise that ideally should be removed from DEB, but we do not believe it will prevent use of the current system; when using DEB to classify drug-CM correlations from an EMR, one is unlikely to encounter the concept "sheep diseases." However, future iterations of DEB will focus on eliminating this noise, potentially through the use of complete MEDLINE data instead of MRCOC.

Finally, some of the DEB, SIDER, and MEDI indications were incomplete, in that some drug indications require concurrence of multiple CMs (e.g., ACE inhibitors are preferred for treating patients with diabetes who have hypertension and/or albuminuria). For completeness and correctness, resources such as the DEB should represent drug-CM associations that involve multiple drugs and multiple CMs with specified logical and temporal relationships. To address ambiguity, resources such as the DEB should also incorporate a certainty metric, based upon number of independent sources from which the relationship was derived, and the strength of the evidence.

6. Conclusion

This preliminary study illustrated the potential utility of using public domain sources to create automatically a drug indication and adverse effect knowledge base. Many such public sources are frequently updated, enabling DEB-like databases to algorithmically generate new, improved versions

without manual intervention. The authors believe limitations encountered in DEB construction and the review of similar systems are important, but that they can eventually be overcome. To do so requires methods to relate broad concepts to more specific concepts, a normalization method to combine similar CM concepts, and improved indication and ADE definitions that represent the inherent complexity of those entities. The results of this preliminary evaluation will hopefully enable researchers in this field to improve drug information databases for use in decision support and pharmacovigilance.

Clinical Relevance Statement

Our study identified important problems regarding drug knowledge extraction from multiple sources. Using this information, researchers can develop more accurate drug indication and ADE databases necessary for decision support systems and pharmacovigilance studies.

Conflict of Interest
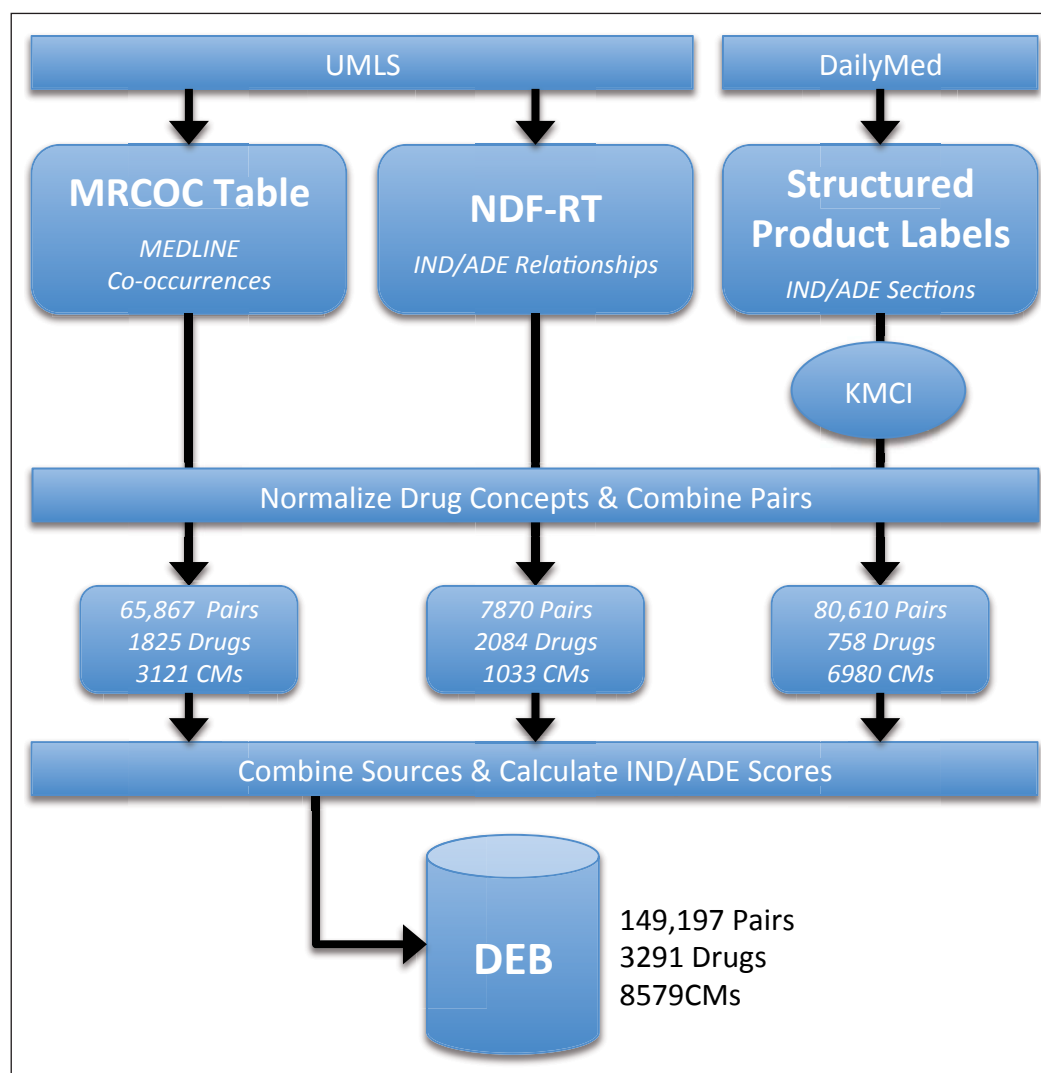
The authors declare that they have no conflicts of interest in the research.

Human Subjects Protections

No human subjects were involved in this research.

**Fig. 1** Flowchart for DEB (Drug Evidence Base) creation.

Drug	Clinical Manifestation	Sources	
C0043031 Warfarin	C0000737 Abdominal Pain	spl	AE
C0043031 Warfarin	C0877248 Adverse event	spl	AE
C0043031 Warfarin	C0002792 anaphylaxis	spl	AE
C0043031 Warfarin	C0002871 Anemia	spl	AE
C0043031 Warfarin	C0004238 Atrial Fibrillation	mrcoc,ndfrt,spl	IND
C0043031 Warfarin	C0038454 Stroke	mrcoc,ndfrt,spl	IND
C0043031 Warfarin	C0010072 Coronary Thrombosis	mrcoc,ndfrt	IND
C0043031 Warfarin	C0011991 Diarrhea	mrcoc,spl	AE
C0043031 Warfarin	C0019080 Hemorrhage	mrcoc,spl	AE
C0043031 Warfarin	C0019158 Hepatitis	spl	AE
C0043031 Warfarin	C0027051 Myocardial Infarction	mrcoc,ndfrt,spl	IND
C0043031 Warfarin	C0151744 Myocardial Ischemia	mrcoc	IND
C0043031 Warfarin	C0027497 Nausea	spl	AE
C0043031 Warfarin	C0151791 Nausea, vomiting, diarrhea	spl	AE
C0043031 Warfarin	C0027540 Necrosis	spl	AE
C0043031 Warfarin	C0032787 Postoperative Complications	mrcoc,ndfrt	IND
C0043031 Warfarin	C0033117 Priapism	mrcoc,spl	AE
C0043031 Warfarin	C0151872 Prothrombin time increased	spl	AE
C0043031 Warfarin	C0034065 Pulmonary Embolism	mrcoc,ndfrt,spl	IND
C0043031 Warfarin	C0039070 Syncope	spl	AE
C0043031 Warfarin	C0040034 Thrombocytopenia	mrcoc	IND
C0043031 Warfarin	C0857496 Thromboembolic event	spl	IND
C0043031 Warfarin	C0040038 Thromboembolism	mrcoc,ndfrt	IND
C0043031 Warfarin	C0398623 Thrombophilia	mrcoc	IND
C0043031 Warfarin	C0040046 Thrombophlebitis	mrcoc,ndfrt	IND
C0043031 Warfarin	C0007787 Transient Ischemic Attack	mrcoc,ndfrt	IND
C0043031 Warfarin	C0042373 Vascular Diseases	mrcoc	IND
C0043031 Warfarin	C0042384 Vasculitis	spl	AE
C0043031 Warfarin	C0042487 Venous Thrombosis	mrcoc,ndfrt,spl	IND
...			

Fig. 2 Sample data from the combined DEB entry for the drug warfarin, including the sources containing the pair and the determined IND/ADE relationship.
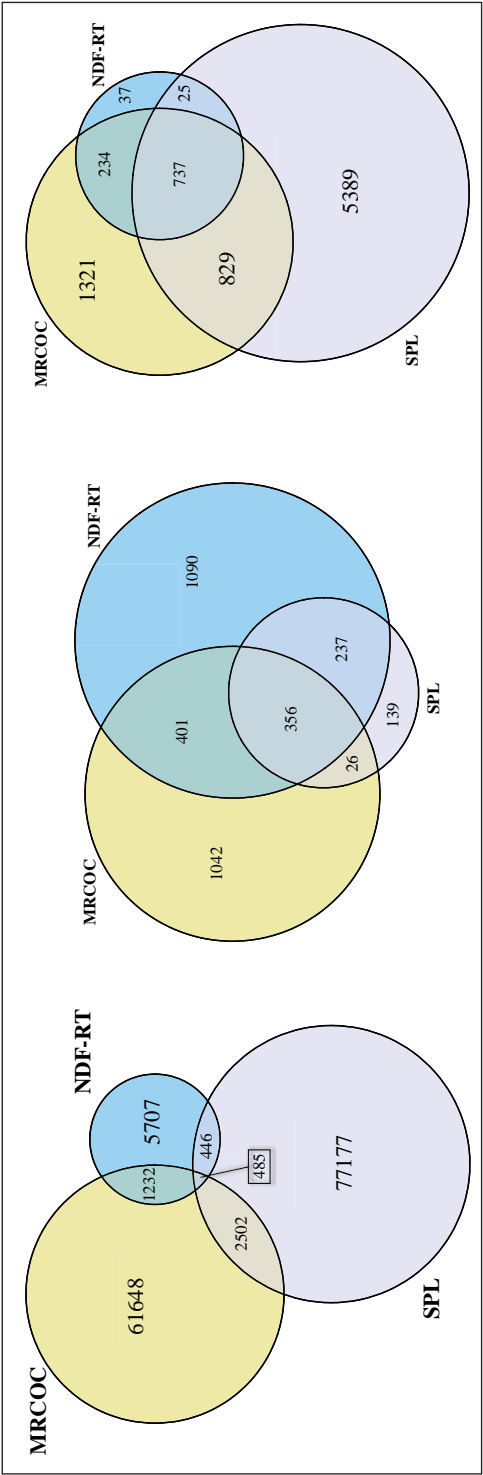
Fig. 3   Venn diagrams illustrating the intersection of drug-CM pairs (a), drugs (b), and CMs (c) from the three DEB sources.
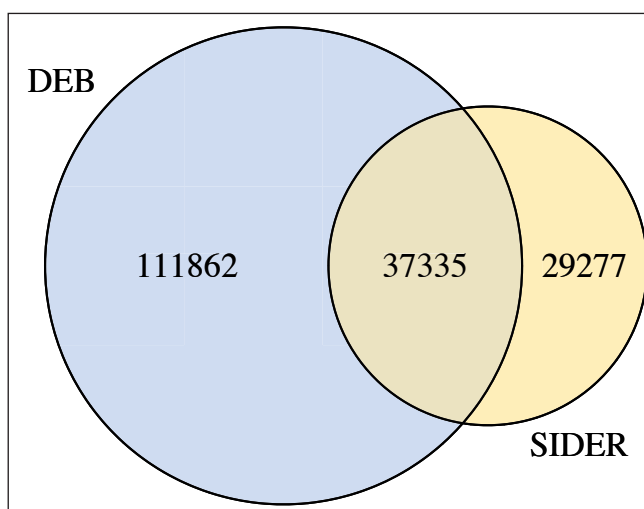
Fig. 4 Venn diagram illustrating the intersection of drug-CM pairs between DEB and the SIDER Side Effect Resource (January 2012 version).

DEB

111862   37335   29277

SIDER

Table 1A   DEB and DEB/SIDER Evaluation: Agreement between reviewers.

Reviewers	N = 25 (all)		N = 21 ("Both" removed)	
	Kappa	**95% CI**	**Kappa**	**95% CI**
1 and 2	0.11	(-0.13, 0.36)	0.17	(-0.16, 0.49)
1 and 3	0.58	(0.28, 0.87)	0.59	(0.27, 0.90)
1 and 4	0.43	(0.18, 0.68)	0.59	(0.27, 0.90)
2 and 3	0.13	(-0.11, 0.38)	0.18	(-0.14, 0.49)
2 and 4	0.47	(0.23, 0.71)	0.34	(0.03, 0.66)
3 and 4	0.50	(0.25, 0.75)	0.67	(0.36, 0.99)
1, 2, 3, & 4	0.36	(0.26, 0.47)	0.42	(0.29, 0.55)
1, 3, & 4	0.50	(0.34, 0.65)	0.62	(0.43, 0.80)

Table 1B   DEB and DEB/SIDER Evaluation: Reviewers' assessment of disagreements between DEB and SIDER (n = 75 each).

Reviewer	Agreed w/ DEB (Pr_1)	Agreed w/ SIDER (Pr_2)	Both (Pr_3)	Neither (Pr_4)	$Pr_1 - Pr_2$	
					Est.	**95% CI**
1	0.64	0.28	0.03	0.05	0.36	(0.16, 0.56)
2	0.55	0.24	0.09	0.12	0.31	(0.12, 0.50)
3	0.53	0.28	0.01	0.17	0.25	(0.06, 0.45)
4	0.52	0.25	0.07	0.16	0.27	(0.08, 0.46)
Average	0.56	0.26	0.05	0.13	0.30	(0.20, 0.40)

Table 1C   DEB and DEB/SIDER Evaluation: Reviewers' categorizations of random selection from DEB only (n = 100 each).

Reviewer	Agreed (Pr_1)	Disagreed (Pr_2)	Both (Pr_3)	Neither (Pr_4)	Pr_1 / Pr_2	
					Est.	**95% CI**
1	0.78	0.16	0.03	0.03	4.88	(3.00, 9.33)
2	0.52	0.11	0.08	0.29	4.73	(2.71, 10.60)
3	0.58	0.08	0.00	0.34	7.25	(3.79, 20.36)
4	0.57	0.03	0.02	0.38	19.00	(7.86, 63.00)
Average	0.61	0.10	.0.3	0.26	8.96	(5.60, 20.86)

Table 2 DEB/SIDER categorizations for the drug Abacavir and reviewer comments (abridged). Blank entries indicate the pair was not present.

CM	SIDER	DEB	Comments
Abdominal Pain	**ADE**	**ADE**	
Acidosis	ADE		See below
Acidosis, Lactic		ADE	See above
AIDS	**IND**	**IND**	
Adverse event		ADE	Non-specific, low-information concept
Alanine aminotransferase increased		ADE	See "Liver function…" & "Raised liver…" below
Allergy Severity – Severe		ADE	See below
Anaphylaxis	ADE		See above
Anemia	**ADE**	**ADE**	
Anorexia	ADE		
Anxiety	ADE		
Arthralgia	ADE		See "Musculoskeletal Pain" above
Blind Vision		ADE	
Bronchitis	**ADE**	**ADE**	
Chills	**ADE**	**ADE**	
Conjunctivitis	ADE		
Coughing	ADE		See "Bronchitis" above
Creatine phosphokinase increased		ADE	
Depressive disorder		ADE	
Diarrhea	**ADE**	**ADE**	See "Severe Diarrhea" below
Dizziness	**ADE**	**ADE**	
Dream disorder		ADE	See "Sleep Disorders" below
Dyspnea	ADE		See "Shortness of breath" below
Edema	ADE		
Enlargement of lymph nodes	ADE		
Erythema Multiforme	**ADE**	**ADE**	
Exanthema	**ADE**	**ADE**	See multiple skin disorders listed elsewhere
Fatigue	**ADE**	**ADE**	Non-specific, low-information concept
Fatty Liver	**ADE**	**ADE**	
Fever	**ADE**	**ADE**	
Gastritis	**ADE**	**ADE**	See below
Gastroenteritis	ADE		See above
Gastrointestinal sign		ADE	See above
Gastrointestinal symptoms NOS		ADE	See above
Headache	**ADE**	**ADE**	
HIV Infections	ADE	IND	
Hyperamylasemia	**ADE**	**ADE**	See "Pancreatitis" below

Table 2   Continued

CM	SIDER	DEB	Comments
Hyperglycemia	**ADE**	**ADE**	
Hypersensitivity	**IND**	**IND**	
Hypertriglyceridemia	**ADE**	**ADE**	
Hypotension	ADE		Part of "Anaphylaxis" above or independent?
Infection	IND		Non-specific, low-information concept
Infective pharyngitis		ADE	See "Pharyngitis" below
Influenza	ADE		See "viral respiratory infection" below
Kidney Failure	ADE		
Leukopenia	**ADE**	**ADE**	See "WBC…" below
Lipid Metabolism Disorders	ADE		See "Hypertriglyceridemia" above
Liver Failure	ADE		See below
Liver function tests abnormal find.		ADE	See above AND "Raised liver …"
Lymphopenia	ADE		See "Leukopenia" above
Malaise	**ADE**	**ADE**	Non-specific, low-information concept
Migraine Disorders	**ADE**	**ADE**	
Morular Metaplasia of the Endometrium		ADE	Non-specific, low-information concept
Musculoskeletal pain	**ADE**	**ADE**	See below
Myalgia	ADE		See above
Myocardial Infarction		ADE	
Nasal infection		ADE	
Nausea	**ADE**	**ADE**	
Neutropenia	**ADE**	**ADE**	See "Leukopenia" above
Oral Ulcer	ADE		
Pain	ADE	ADE	
Pancreatitis	ADE	ADE	
Paresthesia	ADE		
Pharyngitis	ADE		See "Infective pharyngitis" above
Pneumonia	ADE	ADE	
Raised liver function tests	ADE		See "Liver function…" and "Alanine…" above
Respiratory Distress Syndrome, Adult	ADE		See below
Respiratory Failure	ADE		See above
Severe diarrhea		ADE	See "Diarrhea" above
Shortness of Breath	ADE		See "Dyspnea" above
Sleep Disorders	**ADE**	**ADE**	See below
Sleeplessness	ADE		See above
Sore Throat	ADE		See "Pharyngitis" and "Infective Pharyngitis" above
Spondylolisthesis, grade 2		ADE	

Table 2   Continued

CM	SIDER	DEB	Comments
Stevens-Johnson Syndrome	**ADE**	**ADE**	
Therapy naive		ADE	Non-specific, low-information concept
Thrombocytopenia	**ADE**	**ADE**	
Toxic Epidermal Necrolysis	**ADE**	**ADE**	
Urticaria	ADE		See "Allergy" above
Viral respiratory infection		ADE	See "Influenza" above
Vomiting	**ADE**	**ADE**	
White blood cell count increased		ADE	

Table 3   Illustrative sample of reviewer comments, including reviewer and DEB categorizations.

Drug	CM	Review	DEB	Comments
Acetylcysteine	Heart Diseases	IND	IND	"Heart Diseases" vague; N-Acetlycysteine used to prevent damage due to myocardial ischemia, mostly in research
almotriptan	Nausea	Both	IND	Either both or ADE only; Can cause nausea, and indicated for migraine which has nausea as a symptom often
Anti-Bacterial Agents	Theileriasis	Neither	IND	"Anti-bacterial agents" too broad; disease only affects cattle
Anticonvulsants	Ketogenic Diet	Neither	IND	Both are treatments for seizures
Antioxidants	Pathologic Neo-vascularization	IND	IND	"Antioxidants" too general; pathologic neovascularization is not much better
Bupropion	Weight Gain	IND	IND	Indirect association; helps with smoking but prevents weight gain experienced during smoking cessation
Cardiovascular Agents	Atrial Fibrillation	Both	IND	"CV agents" category too general; mostly treat; digoxin does both
Corticotropin	Contracture	Neither	ADE	Corticotropin is a natural substance in humans; its deficiency can lead to flexion contractures
Dantrolene	Tachypnea	Neither	IND	Dantrolene treats malignant hyperthermia, itself a very rare cause of tachypnea
Estrogens	Stroke	ADE	IND	Very weak association in the literature.
Imiquimod	Carcinoma	IND	IND	"Carcinoma" too general; this is a topical agent used in various forms of skin cancer
Lactulose	Diarrhea	ADE	IND	Both desired and adverse effect
Levalbuterol	Adverse event	Neither	ADE	"Adverse Event" is too nonspecific, ignore this term
Praziquantel	Sheep Diseases	Neither	IND	Non-human
Psychotropic Drugs	Substance-Related Disorders	IND	ADE	"Substance-related disorders" and psychotropic drugs; both too broad as categories

Table 4   Reviewer analysis for metoprolol indications from DEB and MEDI.

Source	Metoprolol Indication	Comments
MEDI	Acute myocardial infarction	Indication, on-label
MEDI	Acute myocardial infarction, unspecified site	indication, on-label; redundant
MEDI	Acute myocardial infarction, unspecified site, episode of care unspecified	indication, on-label; redundant
DEB	Myocardial Infarction	Indication, on-label
DEB	Myocardial Ischemia	Indication, related to MI; redundant
DEB	Myocardial Reperfusion Injury	Indication, related to MI; redundant
DEB	Myocarditis	Indication, off-label
DEB	Hemodynamically stable	Incorrect; Label states patient should be hemodynamically stable before treated with metoprolol post-MI
DEB/MEDI	Angina Pectoris	Indication, on-label
DEB	Anginal attack	Indication, on-label
MEDI	Other and unspecified angina pectoris	Indication, on-label
MEDI	Chest pain	Incomplete IND; only for certain types of chest pain
MEDI	Unspecified chest pain	Incomplete IND; only for certain types of chest pain
DEB	Cardiomyopathy, Dilated	Indication, off-label
DEB	Cardiomyopathies	Indication, off-label; redundant
DEB/MEDI	Congestive heart failure	Indication, off-label
DEB/MEDI	Heart failure	Indication, off-label; redundant
MEDI	Heart failure unspecified	Indication, off-label; redundant
DEB	Low Cardiac Output	Indication, off-label; redundant
DEB	Coronary Artery Disease	Indication, off-label
DEB	Coronary Heart Disease	Indication, off-label; redundant
DEB	Coronary Stenosis	Indication, off-label; redundant
MEDI	Long QT syndrome	Indication, off-label
DEB	Mitral Valve Insufficiency	Incorrect
DEB	Stroke Volume	Incorrect; CM too broad and vague
DEB	Cardiovascular Diseases	Incorrect; CM too broad and vague
DEB/MEDI	Heart Diseases	Incorrect; CM too broad and vague
DEB/MEDI	Atrial Fibrillation	Indication, off-label
DEB/MEDI	Atrial Flutter	Indication, off-label
DEB/MEDI	Cardiac Arrhythmia	Incomplete IND; only certain arrhythmias
MEDI	Other specified cardiac dysrhythmias	Incomplete IND; only certain arrhythmias
DEB/MEDI	Ventricular Fibrillation	Incomplete IND; only as prophylaxis post-MI
DEB	Tachycardia, Ventricular	Indication, off-label
MEDI	Paroxysmal ventricular tachycardia	Indication, off-label
DEB	Supraventricular tachycardia	Indication, off-label
DEB	Ventricular Dysfunction, Left	Indication, off-label
MEDI	Essential Hypertension	Indication, on-label
MEDI	Hypertension, NOS	Indication, on-label

Table 4   Continued

Source	Metoprolol Indication	Comments
MEDI	Hypertensive disease	Indication, on-label
DEB	Hypertensive (finding)	Indication, on-label
DEB	Hypertensive disease	Indication, on-label
MEDI	Hypotension	Incorrect
MEDI	Hypotension NOS	Incorrect
MEDI	Migraine	Indication, off-label
DEB	Migraine	Indication, off-label
MEDI	Migraine, unspecified	Indication, off-label; redundant
MEDI	Migraine, unspecified, without mention of intractable migraine without mention of status migrainosus	Indication, off-label; redundant
MEDI	Anxiety state unspecified	Incorrect; CM too broad and vague
MEDI	Social phobia	Incorrect; CM too broad and vague
MEDI	Personality disorder NOS	Incorrect
MEDI	Unspecified nonpsychotic mental disorder	Incorrect
MEDI	Asthma	Incorrect
MEDI	Asthma, unspecified	Incorrect
MEDI	Asthma, unspecified type, without mention of status asthmaticus	Incorrect
MEDI	Diabetes insipidus	Incorrect
DEB	Diabetes Mellitus, Insulin-Dependent	Incorrect
DEB	Diabetes Mellitus, Non-Insulin-Dependent	Incorrect
DEB	Diabetic Angiopathies	Incorrect
DEB	Complications of Diabetes Mellitus	Incorrect
DEB	Insulin Resistance	Incorrect
MEDI	Thyrotoxicosis without mention of goiter or other cause, and without mention of thyrotoxic crisis or storm	Indication, off-label
DEB	Parkinson Disease	Indication, off-label
DEB	Vasovagal syncope	Indication, off-label
DEB	Albuminuria	Incomplete IND; Indication for Beta-blockers in patients with diabetes, but other drugs are preferred
MEDI	Calculus of kidney	Incorrect
MEDI	Syncope and collapse	Incorrect
MEDI	Electrolyte and fluid disorders not elsewhere classified	Incorrect
MEDI	Other fluid overload	Incorrect
DEB	Polycystic Kidney, Autosomal Dominant	Incorrect
MEDI	Unspecified extrapyramidal disease and abnormal movement disorder	Incorrect
DEB	Postoperative Complications	Incorrect; CM too broad and vague
MEDI	Other unknown and unspecified cause of morbidity or mortality	Incorrect; CM too broad and vague

References

1. Miller RA, Gardner RM, Johnson KB, Hripcsak G. Clinical decision support and electronic prescribing systems: a time for responsible thought and action. J Am Med Inform Assoc 2005; 12(4): 403–409.

2. Wang X, Chase HS, Li J, Hripcsak G, Friedman C. Integrating heterogeneous knowledge sources to acquire executable drug-related knowledge. AMIA Annu Symp Proc 2010; 2010: 852–856.

3. Li Y, Salmasian H, Harpaz R, Chase H, Friedman C. Determining the Reasons for Medication Prescriptions in the EHR using Knowledge and Natural Language Processing. AMIA Annu Symp Proc 2011; 2011: 768–776.

4. LePendu P, Iyer SV, Bauer-Mehren A, Harpaz R, Mortensen JM, Podchiyska T, Ferris TA, Shah NH. Pharmacovigilance Using Clinical Notes. Clin Pharmacol Ther [Internet]. 2013 Mar 4 [cited 2013 Apr 11]; Available from: http://www.nature.com/doifinder/10.1038/clpt.2013.47

5. Sharp M, Bodenreider O, Wacholder N. A framework for characterizing drug information sources. AMIA Annu Symp Proc 2008; 2008: 662–666.

6. FDA's Sentinel Initiative [Internet]. [cited 2012 Apr 3]. Available from: http://www.fda.gov/Safety/FDAs SentinelInitiative/default.htm

7. Mini-Sentinel [Internet]. [cited 2013 Aug 23]. Available from: http://www.mini-sentinel.org/default.aspx

8. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: A feasibility study. J Am Med Inform Assoc 2009; 16(3): 328–337.

9. Wang X, Chase H, Markatou M, Hripcsak G, Friedman C. Selecting information in electronic health records for knowledge acquisition. J Biomed Inform 2010; 43(4): 595–601.

10. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med 2012; 4(125): 125ra31.

11. Wei W-Q, Cronin RM, Xu H, Lasko TA, Bastarache L, Denny JC. Development and evaluation of an ensemble resource linking medications to their indications. J Am Med Inform Assoc [Internet]. 2013 Apr 10 [cited 2013 Apr 11]; Available from: http://jamia.bmj.com/cgi/doi/10.1136/amiajnl-2012–001431

12. Coloma PM, Avillach P, Salvo F, Schuemie MJ, Ferrajolo C, Pariente A, Fourrier-Réglat A, Molokhia M, Patadia V, Van der Lei J, Sturkenboom M, Trifirò G. A reference standard for evaluation of methods for drug safety signal detection using electronic healthcare record databases. Drug Saf 2013; 36(1): 13–23.

13. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. Clin Pharmacol Ther 2012; 91(6): 1010–1021.

14. 2011AA National Drug File – Reference Terminology Source Information [Internet]. [cited 2012 Feb 8]. Available from: http://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/NDFRT/

15. MEDLINE Fact Sheet [Internet]. [cited 2013 Apr 3]. Available from: http://www.nlm.nih.gov/pubs/fact sheets/medline.html

16. Unified Medical Language System (UMLS) – Home [Internet]. [cited 2013 Apr 3]. Available from: http://www.nlm.nih.gov/research/umls/

17. DailyMed [Internet]. [cited 2012 Feb 8]. Available from: http://dailymed.nlm.nih.gov/dailymed/about.cfm

18. Micromedex [Internet]. [cited 2012 Apr 23]. Available from: http://www.micromedex.com/

19. Drug Data | FDB (First Databank) [Internet]. [cited 2012 May 4]. Available from: http://www.fdbhealth.com/

20. Point of Care Medical Applications | Epocrates [Internet]. [cited 2013 Mar 19]. Available from: http://www.epocrates.com/

21. UpToDate Inc. [Internet]. [cited 2012 May 4]. Available from: http://www.uptodate.com/index

22. Schadow G. Assessing the impact of HL7/FDA Structured Product Label (SPL) content for medication knowledge management. AMIA Annu Symp Proc. 2007;2007:646–50.

23. Pathak J, Chute CG. Analyzing categorical information in two publicly available drug terminologies: RxNorm and NDF-RT. J Am Med Inform Assoc 2010; 17(4): 432–439.

24. Nelson SJ, Zeng K, Kilbourne J, Powell T, Moore R. Normalized names for clinical drugs: RxNorm at 6 years. J Am Med Inform Assoc 2011; 18(4): 441–448.

25. UMLS® Reference Manual – NCBI Bookshelf [Internet]. National Library of Medicine; 2009 [cited 2012 Feb 8]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK9676/

26. MEDLINE Backfiles Source Information [Internet]. [cited 2012 Apr 10]. Available from: http://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/MBD/

27. Wang W, Haerian K, Salmasian H, Harpaz R, Chase H, Friedman C. A drug-adverse event extraction algorithm to support pharmacovigilance knowledge mining from PubMed citations. AMIA Annu Symp Proc 2011; 2011: 1464–1470.

28. Zeng Q, Cimino JJ. Automated knowledge extraction from the UMLS. Proc AMIA Symp 1998; 1998: 568–572.
29. Duda S, Aliferis C, Miller R, Statnikov A, Johnson K. Extracting drug-drug interaction articles from MEDLINE to improve the content of drug databases. AMIA Annu Symp Proc 2005; 2005: 216–220.
30. Avillach P, Dufour J-C, Diallo G, Salvo F, Joubert M, Thiessard F, Mougin F, Trifirò G, Fourrier-Réglat A, Pariente A, Fieschi M. Design and validation of an automated method to detect known adverse drug reactions in MEDLINE: a contribution from the EU-ADR project. J Am Med Inform Assoc 2013; 20(3): 446–452.
31. Shetty KD, Dalal SR. Using information mining of the medical literature to improve drug safety. J Am Med Inform Assoc 2011; 18(5): 668–674.
32. Chen ES, Hripcsak G, Xu H, Markatou M, Friedman C. Automated acquisition of disease drug knowledge from biomedical and clinical documents: an initial study. J Am Med Inform Assoc 2008; 15(1): 87–98.
33. SIDER Side Effect Resource [Internet]. [cited 2012 Feb 22]. Available from: http://sideeffects.embl.de/
34. Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol 2010; 6: 343.
35. Structured Product Labeling > SPL – Downloadable Data [Internet]. [cited 2012 Feb 8]. Available from: http://www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling/ucm240580.htm
36. Denny JC, Smithers JD, Miller RA, Spickard A. "Understanding" medical school curriculum content using KnowledgeMap. J Am Med Inform Assoc 2003; 10(4): 351–362.
37. Denny JC, Miller RA, Waitman LR, Arrieta MA, Peterson JF. Identifying QT prolongation from ECG impressions using a general-purpose Natural Language Processor. Int J Med Inform 2009; 78 (Suppl 1): S34–S42.
38. Denny JC, Choma NN, Peterson JF, Miller RA, Bastarache L, Li M, Peterson NB. Natural language processing improves identification of colorectal cancer testing in the electronic medical record. Med Decis Making 2012; 32(1): 188–197.
39. Fleiss JL, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. Psychol Bull 1969; 72(5): 323–327.
40. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull 1971; 76(5): 378–382.
41. The R Project for Statistical Computing [Internet]. [cited 2013 Apr 14]. Available from: http://www.r-project.org/
42. Ghazvinian A, Noy NF, Musen MA. Creating mappings for ontologies in biomedicine: simple methods work. AMIA Annu Symp Proc 2009; 2009: 198–202.
43. Shah NH, Shah N, Muse MA, Musen M. UMLS-Query: a perl module for querying the UMLS. AMIA Annu Symp Proc 2008; 652–656.
44. McInnes BT, Pedersen T, Pakhomov SVS. UMLS-Interface and UMLS-Similarity : open source software for measuring paths and semantic similarity. AMIA Annu Symp Proc 2009; 2009: 431–435.
45. Xiang Y, Lu K, James SL, Borlawsky TB, Huang K, Payne PRO. k-Neighborhood decentralization: a comprehensive solution to index the UMLS for large scale knowledge discovery. J Biomed Inform 2012; 45(2): 323–336.
46. Pivovarov R, Elhadad N. A hybrid knowledge-based and data-driven approach to identifying semantically similar concepts. J Biomed Inform 2012; 45(3): 471–481.