

Formalizing Evidence Type Definitions for Drug-drug Interaction Studies to Improve Evidence Base Curation

Joseph Utecht^a, Mathias Brochhausen^a, John Judkins^a, Jodi Schneider^b, Richard D. Boyce^c

^a Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, USA,

^b School of Information Sciences, University of Illinois at Urbana-Champaign, Champaign, IL, USA,

^c Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

In this research we aim to demonstrate that an ontology-based system can categorize potential drug-drug interaction (PDDI) evidence items into complex types based on a small set of simple questions. Such a method could increase the transparency and reliability of PDDI evidence evaluation, while also reducing the variations in content and seriousness ratings present in PDDI knowledge bases. We extended the DIDEO ontology with 44 formal evidence type definitions. We then manually annotated the evidence types of 30 evidence items. We tested an RDF/OWL representation of answers to a small number of simple questions about each of these 30 evidence items and showed that automatic inference can determine the detailed evidence types based on this small number of simpler questions. These results show proof-of-concept for a decision support infrastructure that frees the evidence evaluator from mastering relatively complex written evidence type definitions.

Keywords:

Drug Interactions, Ontologies, Artificial Intelligence

Introduction

While medication therapies are generally beneficial to a patient's health, they can also result in harm. A recent review of epidemiologic studies published over a 14 year period found that 3.5% of hospital admissions are the result of an adverse drug reaction [1]. The United States Department of Health and Human Services recently stated that reducing the rate of adverse drug events should be a national priority [2]. Research on the safety of any given drug starts during its development and continues after marketing. During these phases, studies examine the potential for one drug to alter the pharmacokinetic properties (absorption, distribution, metabolism, or distribution) or clinical effect (pharmacodynamics) of another drug. The results of these studies can suggest PDDIs that might lead to preventable harm to patients without proper management [3].

The spectrum of study types used to research PDDIs is broad and complex. These include in vitro experiments, population pharmacokinetic analyses, randomized controlled clinical trials, and observational epidemiologic studies [4]. Data mining adverse event reports and case report evaluation can also be included because these are research activities that generate PDDI hypotheses [5]. Not only are there many different types of PDDI studies, there are also numerous considerations that can influence the validity of a given study [2,4]. The range and complexity of study designs and the resulting evidence types can make it very difficult to evaluate a body of evidence to determine if a PDDI exists. This difficulty might be an

important factor influencing the large differences in content found in PDDI knowledge bases that are designed to help clinicians make management decisions [6,7]. Indeed, a more systematic approach to evaluating evidence for the existence of PDDIs was one of the recommendations put forth by a recent expert consensus conference series [8].

We believe that an ontology-based system can categorize PDDI evidence items into complex types based on a small set of simple questions, thereby reducing the cognitive load experienced by evidence evaluators. Our approach is to have a computer program infer the specific evidence type of a PDDI study based on simple data provided by the evidence evaluator. Here, we report our progress creating an infrastructure for this kind of decision support. We demonstrate a proof-of-concept for using an off the shelf OWL reasoner and formal evidence type definitions so that an evidence base curator only has to identify the study type at a high level (e.g., clinical study) and provide some simple design features (e.g., 'randomization'). With this information, the OWL reasoner can efficiently and consistently infer the specific evidence type with which to tag the study (see Fig. 1).

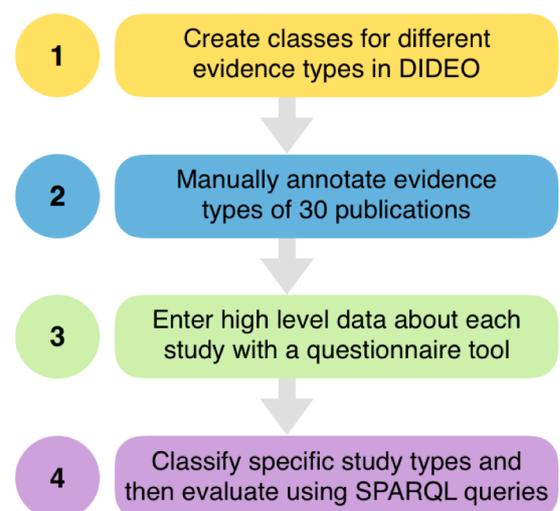


Figure 1 – The steps taken to build the proof-of-concept for categorizing PDDI evidence.

Methods

DIDEO (<https://github.com/DIDEO>) is an OWL ontology developed to facilitate managing information about PDDIs from multiple sources (clinical studies, case studies, in vitro experiments, etc.). It is freely accessible from <http://purl.obolibrary.org/obo/dideo.owl>. The fundamental ontological commitments of DIDEO have been described in detail [9]. One key contribution of DIDEO is to provide OWL classes for evidence types (see Fig. 2).

The process creating the textual definitions for each evidence type has been previously reported [10]. For Step 1 of the current project, we wrote necessary and sufficient conditions (i.e.,

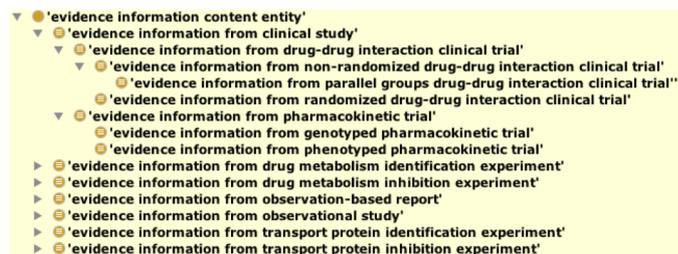


Figure 2 - The evidence information types from DIDEO. All subclasses of “evidence information from clinical study” currently existing in DIDEO are displayed.

Table 1 – (Step 1) Clinical trial evidence types represented in DIDEO, their textual definitions, and their necessary and sufficient conditions.

Evidence type (IRI and rdfs:label)	Definition (iao:definition)	Axiom
DIDEO_0000056 evidence information from clinical study	An evidence information content entity that is about a clinical drug trial.	'evidence information content entity' and ('is about' some (assay and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some ('chemical entity' or 'drug product'))))
DIDEO_0000071 evidence information from drug-drug interaction clinical trial	An evidence information content entity that is about a clinical drug trial that has at least two drugs as its specified input.	'evidence information content entity' and ('is about' some ('scientific observation' and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some (('chemical entity' or 'drug product') and ('is bearer of' some 'object drug role')))) and (has_specified_input some (('chemical entity' or 'drug product') and ('is bearer of' some 'precipitant drug role'))))
DIDEO_0000072 evidence information from non-randomized drug-drug interaction clinical trial	An evidence information content entity that is about a clinical drug trial that has at least two drugs as its specified input.	'evidence information content entity' and ('is about' some (assay and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some (('chemical entity' or 'drug product') and ('bearer of' some 'object drug role')))) and (has_specified_input some (('chemical entity' or 'drug product') and ('bearer of' some 'precipitant drug role'))))
DIDEO_0000073 evidence information from parallel groups drug-drug interaction clinical trial	An evidence information content entity that is about a clinical drug trial that has at least two drugs as its specified input, and that does not have group randomization as a part, and that realizes a clinical study design that has parallel group design as a part.	'evidence information content entity' and ('is about' some ('scientific observation' and (not (has_part value 'group randomization'))) and (realizes some (concretizes some ('clinical study design' and (has_part some 'independent measure design')))) and (has_specified_input some (('chemical entity' or 'drug product') and ('is bearer of' some 'object drug role')))) and (has_specified_input some (('chemical entity' or 'drug product') and ('is bearer of' some 'precipitant drug role'))))
DIDEO_0000074 evidence information from randomized drug-drug interaction clinical trial	An evidence information content entity that is about a clinical drug trial that has at least two drugs as its specified input and does have group randomization as a part.	'evidence information content entity' and ('is about' some ('scientific observation' and (has_part some 'group randomization') and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some (('chemical entity' or 'drug product') and ('is bearer of' some 'object drug role')))) and (has_specified_input some (('chemical entity' or 'drug product') and ('is bearer of' some 'precipitant drug role'))))
DIDEO_0000075 evidence information from pharmacokinetic trial	An evidence information content entity that is about a clinical drug trial that focusses on pharmacokinetics.	'evidence information content entity' and ('is about' some ('scientific observation' and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some (organism and (participates_in some 'pharmacokinetic process')))) and (has_specified_input some ('chemical entity' or 'drug product'))))
DIDEO_0000076 evidence information from genotyped pharmacokinetic trial	An evidence information content entity that is about a clinical drug trial that focusses on pharmacokinetics and that has organisms as participants that participated in genotyping.	'evidence information content entity' and ('is about' some ('scientific observation' and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some (organism and (participates_in some 'pharmacokinetic process') and (participates_in some genotyping))) and (has_specified_input some ('chemical entity' or 'drug product'))))
DIDEO_0000103 evidence information from phenotyped pharmacokinetic trial	An evidence information content entity that is about a clinical drug trial that focusses on pharmacokinetics and that has organisms as participants that participated in phenotyping.	'evidence information content entity' and ('is about' some ('scientific observation' and (realizes some (concretizes some 'clinical study design'))) and (has_specified_input some (organism and (participates_in some 'pharmacokinetic process') and (participates_in some 'phenotype characterization')))) and (has_specified_input some ('chemical entity' or 'drug product'))))

OWL equivalent class axioms) for each textual definition using entities defined or imported into DIDEO (see Table 1). We then ran an OWL reasoner (e.g. Hermit (<http://www.hermit-reasoner.com>)) to categorize the evidence types into a multi-level hierarchy.

Next for Step 2, we created a dataset that would enable automated decision support for people who need to evaluate PDDI evidence items. As Table 1 shows, evidence definitions for more specific kinds of evidence can become fairly complicated. Decision support should simplify the cognitive load of a person evaluating an evidence item.

We developed a proof-of-concept system that was used to test whether an OWL reasoner and formal evidence type definitions can efficiently and consistently infer the specific evidence type from Table 1 based on basic evidence types (e.g., clinical study) and design features (e.g., 'randomization'). Combining the evidence instance data with the OWL equivalent class axioms from DIDEO should enable an OWL reasoner to infer the specific evidence type of each instance. We tested this in Step 4 shown in Figure 1.

For Step 1 we used classes for the detailed evidence types shown in Table 1 in DIDEO (created in prior work) [11]. For Step 2, we drew on prior work [10] that had manually assigned a study type from Table 1: 30 publications were queried from the Drug Interaction Knowledge Base (<http://dikb.org>), representing five publications for each of the six clinical study evidence type defined in DIDEO (see Table 2). Listing 1 shows the SPARQL query used to query a single evidence type from Table 2. The interested reader can modify the `dikbEvidence:Evidence_type` and PubMed identifier to retrieve any of the other items listed.

As part of Step 3 a questionnaire tool, previously developed for the CAFE project (<https://cafe-trauma.com>), was modified and used to manually enter high level type and simple design features for each study in Table 2 (see Fig. 3). We then exported RDF instances created by the tool for each evidence item. We put these into a single file for import into an RDF store. Our RDF store was the community version of the Stardog (<http://stardog.com/>) RDF store, which comes with a built-in OWL description logic reasoner. We loaded RDF file of evidence instance data along with following OWL files:

- DIDEO: <http://purl.obolibrary.org/obo/dideo/release/2016-07-29/dideo.owl>
- RO core: <http://purl.obolibrary.org/obo/ro/releases/2016-09-02/core.owl>
- RO BFO classes minimal: <http://purl.obolibrary.org/obo/ro/releases/2016-09-02/bfo-classes-minimal.owl>
- IAO: <http://purl.obolibrary.org/obo/iao/2015-02-23/iao.owl>
- BFO (classes only) <http://purl.obolibrary.org/obo/bfo/2014-05-03/classes-only.owl>
- RO annotations: <http://purl.obolibrary.org/obo/ro/releases/2016-09-02/annotations.owl>

Listing 1. (Step 1) A query used to retrieve one of the 30 evidence types from the Drug Interaction Knowledge Base (<https://dbmi-icode-01.dbmi.pitt.edu/dikb/snorql>)

```
SELECT DISTINCT ?ev ?quote
WHERE {
  ?ev dikbEvidence:Evidence_type
      dikbEvidence:EV_PK_DDI_NR.
  ?ev rdfs:seeAlso
    <http://www.ncbi.nlm.nih.gov/pubmed/10445377>.
  ?ev siocns:content ?quote.
}
```

Finally for Step 4, we developed a set of competency questions and associated SPARQL queries (see Table 3) that could be implemented in an application similar to that shown in Figure 3 in order to provide automated support to evidence evaluators for determining an evidence item's type. We tested that each SPARQL query ran successfully over the inferred RDF store.

Figure 3 – (Step 3) The questionnaire tool used to enter high level type and simple design features for each study in Table 2.

Table 3 – (Step 4) A set of competency questions and associated SPARQL queries that could be used by an application similar to that shown in Figure 3 to implement automated support to evidence evaluators for determining an evidence item's type.

Competency Question	SPARQL Query
Which evidence items stem from non-randomized clinical studies on drug-drug interactions without parallel group design?	<pre>select * { ?evidence a obo:DIDEO_00000072 . filter not exists { ?evidence a obo:DIDEO_00000073 . } }</pre>
Which evidence items stem from non-randomized clinical studies on drug-drug interactions with a parallel group design?	<pre>select * { ?evidence a obo:DIDEO_00000073 . }</pre>
Which evidence items stem from randomized clinical studies on drug-drug interactions?	<pre>select * { ?evidence a obo:DIDEO_00000074 . }</pre>
Which evidence items stem from clinical studies targeting pharmacokinetics without genotyping or phenotyping?	<pre>select * { ?evidence a obo:DIDEO_00000075 . filter not exists { ?evidence a obo:DIDEO_00000076 . } . filter not exists { ?evidence a obo:DIDEO_00000103 . } . }</pre>
Which evidence items stem from clinical studies targeting pharmacokinetics and using genotyping?	<pre>select * { ?evidence a obo:DIDEO_00000076 . }</pre>
Which evidence items stem from clinical studies targeting pharmacokinetics and using phenotyping?	<pre>select * { ?evidence a obo:DIDEO_00000103 . }</pre>

Table 4 – (Step 4) Retrieval results from running the SPARQL queries (Table 3, Column B) on the Stardog triplestore.

DIDEO annotation	PMID of Publications
Evidence information from non-randomized DDI clinical trial (DIDEO_00000072)	10445377, 10907965, 15876900, 11147928, 8801057
Evidence information from parallel groups DDI (DIDEO_00000073)	12911366, 11910262, 17571477, 9855322, 15518608
Randomized drug-drug interaction clinical trial (DIDEO_00000074)	9542477, 16778714, 11563412, 9757151, 19242403
Pharmacokinetic trial (DIDEO_00000075)	8911886, 15834460, 14747427, 1487561, 1438031
Genotyped pharmacokinetic trial (DIDEO_00000076)	11452243, 19142106, 16765147, 17429316, 8689810
Phenotyped pharmacokinetic trial (DIDEO_00000103)	8513845, 7690693, 2007317, 1412613, 8823236

Results

RDF descriptions, generated by the tool in Step 3, for the 30 clinical studies is freely available from <https://goo.gl/Wk84gg>. Table 4 shows the results of running the competency question SPARQL queries (see Table 3) over the inferred RDF store. Each evidence type classification corresponds exactly to the

manual classification from Table 2, thereby validating the data model and OWL ontology while showing proof of concept that complex evidence type classification can be obtained from the answers to simple questions.

Discussion

DIDEO now provides OWL classes for many of the evidence types within the PDDI domain of scientific discourse. The current work shows the feasibility of using these formal definitions to build decision support that helps evidence evaluators to determine the specific types of studies or experiments they review. Instead of having to master complex written evidence type definitions, such as those shown in Table 1, evaluators will only need to answer a small number of simple questions about the study, such as those shown Figure 3. With that information, simple SPARQL queries can be run over an inferred RDF dataset similar to the one used in this study to return the specific evidence type as a URI from DIDEO. By simplifying the task of evidence type assignment, a team of curators should be able to produce more correct and consistent work over the many hundreds of evidence items they need to manage.

Conclusions

In this work we showed proof-of-concept for the technical infrastructure, showing that PDDI evidence item data we added into a triple store together with DIDEO evidence types enable an OWL reasoner to infer an evidence item's specific type. In future work we plan to conduct a user study to test the hypothesis that this approach improves the correctness and consistency of evidence type assignment by evidence base curators. We also plan to extend the decision support to include questions that can be used to infer the quality of a study. We then plan to conduct a user study that compares interrater

agreement between two groups of PDDI experts – one provided with evidence evaluation decision support, the other using their usual procedures. Based on our hypothesis, the group of participants given decision support should show much greater interrater agreement than the comparison group. We will also compare speed, ease, and user preference. If our hypothesis is shown to be true, it would be a promising step forward toward reducing the large differences in content found in PDDI knowledge bases that are designed to help clinicians make management decisions [6,7].

Evidence annotation is already a significant part of the biomedical enterprise, with computer-supported manual annotation used by an entire community and additional efforts to synthesize clinical research (e.g. Cochrane Reviews). We believe that approaches to simplify annotation processes, using incremental formalization and granular information, will be essential for increasing the availability of searchable (and in some cases algorithmically-synthesized) information. Generalizations of the work presented in this paper have the potential to greatly increase the impact of the curation enterprise.

Acknowledgements

This project was supported by a grant from the National Library of Medicine: “Addressing gaps in clinically useful evidence on drug-drug interactions” (R01LM011838). JS was partially supported by training grant T15LM007059 from the National Library of Medicine/National Institute of Dental and Craniofacial Research.

References

- [1] J.C. Bouvy, M.L.D. Bruin, M.A. Koopmanschap, Epidemiology of adverse drug reactions in Europe: A review of recent observational studies. *Drug Saf* **38** (2015), 437–453.
- [2] US Department of Health and Human Services, *National Action Plan for Adverse Drug Event Prevention* [Internet]. 2014. Available from: <https://health.gov/hcq/pdfs/ade-action-plan-508c.pdf>
- [3] L. Magro, U. Moretti, R. Leone, Epidemiology and characteristics of adverse drug reactions caused by drug-drug interactions, *Expert Opin Drug Saf* **11** (2012), 83–94.
- [4] S. Hennessy, C.E. Leonard, J.J. Gagne, J.H. Flory, X. Han, C.M. Brensinger, W.B. Bilker, Pharmacoepidemiologic methods for studying the health effects of drug-drug interactions, *Clin Pharmacol Ther* **99** (2016), 92–100.
- [5] B. Percha, R.B. Altman, Informatics confronts drug-drug interactions, *Trends Pharmacol Sci* **34** (2013), 178–184.
- [6] L.M. Wang, M. Wong, J.M. Lightwood, C.M. Cheng, Black box warning contraindicated comedications: concordance among three major drug interaction screening programs, *Ann Pharmacother* **44** (2010), 28–34.
- [7] K.R. Saverno, L.E. Hines, T.L. Warholak, A.J. Grizzle, L. Babits, C. Clark, A.M. Taylor, D.C. Malone, Ability of pharmacy clinical decision-support software to alert users about clinically important drug–drug interactions, *J Am Med Inform Assoc* **18** (2011), 32–37.
- [8] R.T. Scheife, L.E. Hines, R.D. Boyce, S.P. Chung, J.D. Momper, C.D. Sommer, D.R. Abernethy, J.R. Horn, S.J. Sklar, S.K. Wong, G. Jones, M.L. Brown, A.J. Grizzle, S. Comes, T.L. Wilkins, C. Borst, M.A. Wittie, D.C. Malone. Consensus recommendations for systematic evaluation of drug-drug interaction evidence for clinical decision support, *Drug Saf* **38** (2015), 197–206.
- [9] M. Brochhausen, J. Schneider, D. Malone, P. Empey, W.R. Hogan, R.D. Boyce, Towards a foundational representation of potential drug-drug interaction knowledge, *Proc. of DIKR 2014/IWOOD 2014/OBIB 2014 – Workshops of ICBO 2014*, Houston TX, USA, October 6-7, 2014, <http://ceur-ws.org/Vol-1309/paper2.pdf>.
- [10] R. Boyce, C. Collins, J. Horn, I. Kalet, Computing with evidence: Part I: A drug-mechanism evidence taxonomy oriented toward confidence assignment, *J Biomed Inform* **42**, 979–989.
- [11] M. Brochhausen, P.E. Empey, J. Schneider, W.R. Hogan, R.D. Boyce, Adding Evidence Type Representation to DIDEO, *Proc. Of the Joint International Conference on Biological Ontology and BioCreative (2016)*, Corvallis OR, USA, August 1-4, 2016, http://ceur-ws.org/Vol-1747/IP02_ICBO2016.pdf.

Address for correspondence

Mathias Brochhausen: MBrochhausen@uams.edu
Department of Biomedical Informatics
University of Arkansas for Medical Sciences
4301 W. Markham St., #782
Little Rock, AR 72205-7199
USA