



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Bacterial clinical infectious diseases ontology (BCIDO) dataset

Claire L. Gordon^{a,b,c}, Chunhua Weng^{b,*}^a Department of Medicine, Columbia University Medical Center, 630 West 168th Street, New York, USA^b Department of Biomedical Informatics, Columbia University Medical Center, 622 West 168th Street, New York, NY 10032, USA^c Department of Medicine, University of Melbourne, Melbourne, VIC 3010, Australia

ARTICLE INFO

Article history:

Received 24 December 2015

Received in revised form

28 June 2016

Accepted 11 July 2016

Available online 16 July 2016

Keywords:

Biomedical informatics

Ontology

Infectious diseases

ABSTRACT

This article describes the Bacterial Infectious Diseases Ontology (BCIDO) dataset related to research published in <http://dx.doi.org/10.1016/j.jbi.2015.07.014> [1], and contains the Protégé OWL files required to run BCIDO in the Protégé environment. BCIDO contains 1719 classes and 39 object properties.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications Table

Subject area	Medicine, Biomedical informatics
More specific subject area	Bacterial Clinical Infectious Diseases Ontology
Type of data	Figure, Protégé source files
How data was acquired	Ontology was developed by Claire L. Gordon and includes imports from the OBO Foundry, Infectious Disease Ontology, Foundational Model of Anatomy and NCBI Taxon. Classes were mapped to Unified Medical Language System (UMLS) concept unique identifiers (CUIs) where possible.
Data format	Formatted

DOI of original article: <http://dx.doi.org/10.1016/j.dib.2016.07.018>

* Corresponding author.

E-mail address: cw2384@cumc.columbia.edu (C. Weng).<http://dx.doi.org/10.1016/j.dib.2016.07.018>2352-3409/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Experimental factors	<i>Preparation of BCIDO was as follows: (1) determination of the domain and scope of the ontology; (2) review of the literature and related ontologies to evaluate them for reuse; (3) knowledge representation; and (4) evaluation.</i>
Experimental features	<i>BCIDO data is represented in the Web Ontology Language (OWL) as a single hierarchical structure using the Protégé-OWL editor Version 4.1 (http://protege.stanford.edu). Clinical ID concepts and antimicrobials in BCIDO were mapped to the reference resource Unified Medical Language System concept unique identifiers where possible. Bacterial terms were imported from the National Center for Biotechnology Information Organismal Classification (NCBITaxon). Anatomical terms were imported from The Foundational Model of Anatomy (FMA) (http://sig.biostr.washington.edu/projects/fm/index.html).</i>
Data source location	<i>n/a</i>
Data accessibility	<i>Data is submitted with this article</i>

Value of the data

- BCIDO may be useful for improving interoperability of antibiotic decision support systems.
 - BCIDO may be used as a knowledge representation framework for clinical infectious disease data.
 - BCIDO can be compared with other infectious disease ontologies to obtain further insight.
 - BCIDO may be reused for designing an antibiotic decision support system.
-

1. Data

BCIDO is represented in the Web Ontology Language (OWL) as a single hierarchical structure using the Protégé-OWL editor Version 4.1 (<http://protege.stanford.edu>). Fig. 1 shows the infectious disease domain class hierarchy. BCIDO contains 1719 classes, 39 object properties, 18 individuals, 2247 subsumption relations (SubClassOf axioms), 2770 logical axioms, 86 EquivalentClasses axioms and 350 DisjointClasses axioms.

2. Experimental design, materials and methods

The design of BCIDO has been described previously [2]. The data contained in BCIDO broadly covers the domain of clinical infectious diseases, and integrates the three major determinants of clinical infectious disease management (e.g. the infectious disease, the causative bacteria and the treating antibiotic). The accuracy and coverage of the data in BCIDO was assessed using a semi-automated method, as described [1]. To open BCIDO in Protégé Version 4.1, an open source collaborative ontology editing environment that is downloadable from Stanford University (<http://protege.stanford.edu>), open the file “BCIDO FINAL DIB.owl”. The required imported files are also contained within the “BCIDO” folder and will import automatically. Click “No” if asked to resolve missing imports.

The screenshot displays the Protégé ontology editor interface. The top menu bar includes 'Active Ontology', 'Entities', 'Classes', 'Object Properties', 'Data Properties', 'Individuals', and 'DL Query'. The main window is divided into two panes. The left pane, titled 'Class hierarchy: acute_cholecystitis', shows a tree view of classes. The root is 'disease', which branches into 'infectious_disease' and several other categories. Under 'infectious_disease', there are sub-classes for various body systems (ear, skin, eye, musculoskeletal, nervous, respiratory, cardiovascular, gastrointestinal). Under 'gastroenteritis', there are sub-classes for colitis, inflammatory enteritides, and cholecystitis. 'acute_cholecystitis' is highlighted in blue. The right pane, titled 'Annotations: acute_cholecystitis', shows the 'Annotations' section with the identifier 'UMLS CUI C0149520'. Below it, the 'Description: acute_cholecystitis' section shows the class's relationships: 'Equivalent classes' (none), 'Superclasses' (cholecystitis), and 'Inherited anonymous classes' (disease and has_material_basis_in some 'infectious disorder' and (inheres_in some (organism and (part_of some ('extended organism' and (has_part some 'infectious disorder')))) and (has_material_basis_in only 'infectious disorder')) and (inheres_in only (organism and (part_of some ('extended organism' and (has_part some 'infectious disorder')))) and (realized_by only 'infectious disease course')) and has_material_basis_in some (intra-abdominal_and_gastrointestinal_infecti on and (located_in some Gallbladder))).

Fig. 1. Bacterial infectious disease ontology (BCIDO) domain class of infectious disease showing "acute cholecystitis" as an example.

Acknowledgements

We are very grateful to Lindsay Cowell and Albert Goldfain for assistance with designing BCIDO. This work was conducted using the Protégé resource, which is supported by Grant GM10331601 from the National Institute of General Medical Sciences of the United States National Institutes of Health. This research was partially supported by National Library of Medicine Grant R01LM009886 to Dr. Weng. Dr. Gordon is supported by a Hutchins Family Fellowship, Fulbright Postgraduate Scholarship and Australian American Association Fellowship. The article's contents are solely the responsibility of the authors and do not necessarily represent the official view of NIH.

Transparency document. Supplementary material

Transparency document associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.07.018>.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.07.018>.

References

- [1] C.L. Gordon, C. Weng, Combining expert knowledge and knowledge automatically acquired from electronic data sources for continued ontology evaluation and improvement, *J. Biomed. Inform.* 2015 (Jul) (2015) 23. <http://dx.doi.org/10.1016/j.jbi.2015.07.014> [Epub ahead of print].
- [2] C.L. Gordon, S. Pouch, L.G. Cowell, M.R. Boland, H.L. Platt, A. Goldfain, et al. Design and evaluation of a bacterial clinical infectious diseases ontology, in: *Proceedings of the AMIA Annual Symposium proceedings/AMIA Symposium AMIA Symposium, 2013; 2013*. pp. 502–11.