

MeTMapS - Medical Terminology Mapping System

Shao Fen Liang^{1,*}, Jean-Francois Ethier², Talya Porat¹, Archana Tapuria¹, Brendan C. Delaney³ and Vasa Curcin¹

¹ Division of Health and Social Care Research, King's College London, London, UK

² French Institute of Health and Medical Research, Centre de Recherche des Cordeliers, Paris, France

³ Department of Surgery and Cancer, Imperial College London, London, UK

ABSTRACT

Motivation: Medical terminology mapping is a long-standing challenge for projects requiring retrieval, querying and integration of heterogeneous patient data. Current tools fail to fully utilise the richness of the underlying coding systems, and can be difficult to install and maintain. For example, National Library of Medicine's UMLS provides a rich collection of terminology mapping, however, its search results are displayed in a simplistic general purpose interface that cannot easily be navigated and results filtered according to user's preferences. Specifically, returned results cannot be visualised in a tree to show positions and relationships. BioPortal offers a large number of terminologies and ontologies, each of which can be viewed in a tree structure, however it does not allow for multiple ontologies to be viewed and compared on a single page. Our work aims to address these issues and provide a simple and easy to use terminology mapping software.

Results: MeTMapS was evaluated with academic and clinical research users. The users have tested the mapping between ICD10, Read CTV2, V3 in Hypertension. It was also tested on a list of clinical terms from the inclusion and exclusion criteria of the INFORM clinical trial protocol. Our initial evaluation produced positive results.

Availability: We are currently in the process of updating the design based on some improvements suggested by the participants. MeTMapS is developed under Apache V2 license and is currently hosted at KCL for internal use and will shortly be opened to the public once the internal security concerns are resolved. In the meantime, the tool is available from the author upon request.

* **Contact:** fennie.liang@kcl.ac.uk

1 INTRODUCTION

Medical terminology mapping is a long-standing challenge when designing observational studies from Electronic Health Records, particularly when working with multiple databases employing different coding systems. Mapping from one terminology to another can rarely be done automatically due to many-to-many mappings that frequently occur between terminologies, and it typically requires a user with medical knowledge and a good understanding of each terminology to manage the cardinality issues. This is one of the reasons that the US National Library of Medicine has built the UMLS Metathesaurus with over a hundred national and international terminologies in different languages and their mappings. The UTS Metathesaurus browser¹ is provided as an interface for navigating the mappings. Users can either simply search for a term or

select a term from the full UMLS collection tree to get results if they have the sufficient knowledge of UMLS. However, the results are displayed in a long list so as to cover all mappings of the UMLS collection. The system provides good completion and detailed information but it is also likely to overload users with information. LexEVS² is another system for terminology mapping and has been used in TRANSFoRm (Ethier, 2013) and BioPortal (Salvadorea, 2013). However, setting up the LexEVS platform and training the users are non-trivial tasks because LexEVS requires users to set up a server and load data into its database so the data can be transformed into LexEVS data objects. We developed the Medical Terminology Mapping System (MeTMapS) with the aim of addressing these concerns and producing a usable system for clinical researchers.

2 METHOD

In the construction of MeTMapS (see Figure 1) we have utilised UTS APIs and BioPortal Widgets, some of which have been modified to achieve better performance, e.g. the tree widget used to display hierarchical terminologies. Any terminology not covered by UMLS can be added into MeTMapS via BioPortal, if the paired mapping file and ontology are ready.

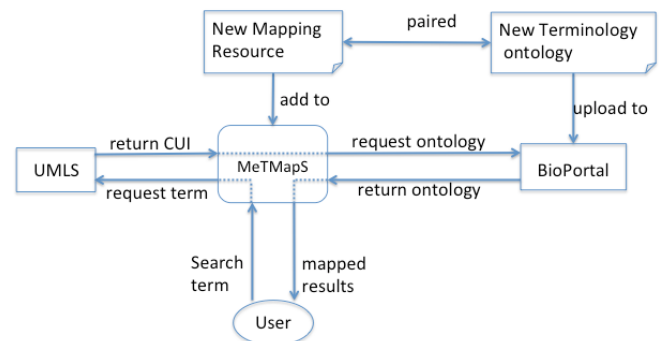


Fig. 1. MeTMapS architecture shows a user inputs a search term to the system. The system requests the term from UMLS and gets a CUI returned. The relevant ontology, which contains the search term is then requested from BioPortal and returned to the system. The returned ontology and the mapped results can then be viewed by the user.

For example, we have generated a Read CTV2 (used by most of the primary care systems in the UK) ontology and uploaded it into BioPortal. We have also produced a mapping file, which contained Concept Unique Identifiers (CUI) from UMLS for MeTMapS.

¹ <https://uts.nlm.nih.gov//metathesaurus.html>

² <https://wiki.nci.nih.gov/display/LexEVS/LexEVS>

The idea behind the MeTMapS user interface (Figure 2) is to focus the users only on the terminologies they need. The search results are organised into tree structures, offering a better view of relationships among the selected terms, their parents, siblings and descendants. To minimise the setup needed and the learning curve for the users, the system is designed as a web application with a step-by-step workflow guide. The interface comprises three main sections: 1) Search of terms and selection of coding systems; 2) browsing of results and selection of mappings from structured trees; and 3) removal of unwanted mappings and storage of results for further use. Each section has a guide for users to follow, requiring no previous training. With all relevant information displayed on one page, it facilitates navigation and orientation. Auto-completion is also provided to help users quickly find their desired term.

3 RESULTS

The MeTMapS system was evaluated by seven academic and clinical research users including GPs, clinical informaticians and IT specialists at KCL in the last two months. A total of 54 sessions were completed, where the system was used by the researchers. Participants were asked to map the required clinical terms in the 'Hypertension' clinical domain either from ICD10 to Read CTV2 or from ICD10 to Read CTV2 and V3. Participants also tested a list of clinical terms extracted from the inclusion and exclusion criteria of the INFORM clinical trial protocol (Wilkinson, 2016). The clinical terms listed were mapped to the clinical terminologies Read CTV2 and V3 that are used by most of the GP systems recruiting patients for the trial. A screen shot of a search for Hypertensive disorder and the mapped results from Read CTV2, Read V3 and ICD10 is shown in Figure 2 (the listed results

contain the same CUI as Hypertensive disorder from UMLS). The correct term always appears first with the most relevant results being at the top, and suggestions provided while typing. The search terms are not order- or case-sensitive and handles terms separated by hyphen. (For example, search term 'sugarfree' includes 'sugar-free' and search term 'breast-cancer' includes 'breast cancer') It handles exact term matches for different concepts such as 'Fundus coloboma' and 'Gastric fundus structure' and multiple terminologies can be simultaneously selected for mapping. We have found that MeTMapS could be improved by providing suggestions on mis-spelled terms, and also enhanced on handling partial words such as 'Myo inf' as opposed to 'myocardial infarction' and on handling known synonyms such as 'kidney' and 'renal'.

ACKNOWLEDGEMENT

This research is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London.

REFERENCES

- Ethier, J-Fo and Dameron, O. et al. (2013) A unified structural/terminological interoperability framework based on LexEVS: application to TRANSFoRM. *Journal of the American Medical Informatics Association*, **2013**(0) 1-9.
- Salvadoresa M. and Alexander PR. Et al. (2013) BioPortal as a Dataset of Linked Biomedical Ontologies and Terminologies in RDF. *Semantic Web*, **4**(3) 277-84.
- Wilkinson Ian (2016) Randomised cross-over trial in a multi-ethnic cohort: AIM HY – INFORM. Work Strand 3, <http://www.aimhy.org.uk/our-research/ws3/>.

MeTMapS - Medical Terminology Mapping System

1 Type in or select a search term and its desired terminologies:

Search Term:

Terminology: Read CTV2 Read V3 (RCD) ICD10 SNOMED CT ICPC2 Plus

2 Select mapping from each ontology

- Hypertensive diseases (ICD10: I10-I15.9)
- High blood pressure (RCD: XM02V)
- Hypertension (RCD: XE0Ub)
- Hypertensive disease (RCD: G2...)
- BP - High blood pressure (RCD: XE0Ub)
- High blood pressure disorder (RCD: XE0Ub)
- Hypertensive disease NOS (RCD: G2z..)
- [X]Hypertensive diseases (RCD: Gyu2.)
- HBP - High blood pressure (RCD: XM02V)
- BP+ - Hypertension (RCD: XE0Ub)
- HT - Hypertension (RCD: XE0Ub)
- Hypertensive disease (RCTV2: G2...00)

Search for class...

- Hypertensive disease(G2...00)
 - Essential hypertension(G20..00)
 - Hypertension resistant to drug therapy(G27..00)
 - Hypertensive disease NOS(G2z..00)
 - Hypertensive heart and renal disease(G23..00)
 - Hypertensive heart disease(G21..00)

3 Final mapping

Preferred Name	Terminology Code	Remove
Hypertensive diseases	ICD10 I10-I15.9	<input checked="" type="checkbox"/>
Hypertension	RCD XE0Ub	<input checked="" type="checkbox"/>
Hypertensive disease	RCTV2 G2...00	<input checked="" type="checkbox"/>

Fig. 2. MeTMapS Interface