

# Modelling Medications for Public Health Research

D. van Gaans<sup>1</sup>, S. Ahmed<sup>1</sup>, K. D'Onise<sup>1</sup>, J. Moyon<sup>2</sup>, G. Caughey<sup>3</sup>, R. McDermott<sup>4</sup>

1. Centre for Research Excellence in the Prevention of Chronic Conditions in Rural and Remote Populations, University of South Australia, South Australia
2. Information Strategy and Technology, University of South Australia, South Australia
3. Pharmacy and Medical Sciences, University of South Australia, South Australia
4. College of Public Health, Medical and Veterinary Sciences, James Cook University, Queensland

## Abstract

Most patients with chronic disease are prescribed multiple medications, which are recorded in their personal health records. This is rich information for clinical public health researchers but also a challenge to analyse. This paper describes the method that was undertaken within the Public Health Research Data Management System (PHReDMS) to map medication data retrieved from individual patient health records for population health researcher's use. The PHReDMS manages clinical, health service, community and survey research data within a secure web environment that allows for data sharing amongst researchers. The PHReDMS is currently used by researchers to answer a broad range of questions, including monitoring of prescription patterns in different population groups and geographic areas with high incidence/prevalence of chronic renal, cardiovascular, metabolic and mental health issues. In this paper, we present the general notion of abstraction network, a higher level network that sits above a terminology and offers compact and more easily understandable view of its content. We demonstrate the utilisation of abstraction network methodology to examine medication data from electronic medical records to allow a compact and more easily understandable view of its content.

**Keywords:** Medication, Public Health, Modelling, Data Mapping

**DOI:** 10.5210/ojphi.v8i2.6809

**Copyright ©2016 the author(s)**

This is an Open Access article. Authors own copyright of their articles appearing in the Online Journal of Public Health Informatics. Readers may copy articles without permission of the copyright owner(s), as long as the author and OJPHI are acknowledged in the copy and the copy is used for educational, not-for-profit purposes.

## Introduction

Population health scientists aim to understand disease patterns and develop approaches for disease prevention, detection, and diagnosis at an early stage to reduce the burden of disease [1]. Within the last decade there has been a rapid increase in the availability of health care data and never before have population scientists had the capacity to collect, share, and analyse data as they have today. Population health researchers often collect data from patients directly or from their health service records. Personal health records (PHRs) of patients contain a wealth of information, but can be challenging to analyse [2]. There is a variety of paper and electronic medical records (EMRs) in use, despite ongoing attempts to standardise data collection and a variety of clinical coding systems implemented by EMR vendors [3]. Terminologies and terminological systems play an important

role in many medical information processing environments, giving rise to the “big knowledge” challenge, when terminological content comprises of tens of thousands to millions of concepts arranged in a non-relational manner [4]. The science of bioinformatics can provide essential tools and a framework for population scientists to manage this potentially overwhelming amount of data [1].

The challenge for Health Information Technology is to design systems that are powerful enough not only to handle the volume and complexity of medical data, but also to support both patients and professionals, resulting in improved health care, reduced costs and health outcomes for patients [2]. Data standardization and harmonization can address some of the obstacles to data sharing and pooling [3] [5] [6] [7]. Data harmonization is used when data standardization is not possible to achieve interoperability across systems. This work is laborious and entails a trans-disciplinary approach wherein informaticists, measurement and topical experts, biostatisticians, and ethicists combine their knowledge to ensure the integrity and security of harmonized data [5] [6]. One method applied to harmonise terminology is the abstraction network.

An abstraction network overlays a terminology’s underlying network structure at a higher level of abstraction. In particular, it provides a more compact view of the terminology’s content, avoiding the display of minutiae [4]. The notion of an abstraction network is presented as a means of facilitating the usability, comprehensibility, visualization, and quality assurance of terminologies [8]. One very important feature of an abstraction network is that it is typically multiple orders of magnitude smaller in size than its underlying terminology. This compact structure makes abstraction networks much more manageable from visualization and comprehension perspectives [9]. The reduction in size of an abstraction network is obtained by structurally dividing a large terminology into smaller parts, each of which is represented by one constituent entity (node) of the abstraction network [4].

There are two main ways to define the set of nodes of an abstraction network, intrinsic and extrinsic. Intrinsic abstraction networks derive them from the concepts and relationships of the underlying terminology itself. That is, some terminology concepts are of a general nature and can be used to properly categorize other elements [4]. Another alternative for choosing nodes is by a domain expert gleaning broad categories, befitting the terminology’s subject matter, from external sources [4]. For example, the categories might be taken from the general body of literature in a subject area or from a standard reference work. An abstraction network derived in this way from sources external to the terminology itself is called extrinsic. For example, for a terminology in the medical field, broad categories could be disease, laboratory test, and procedure. Extrinsic networks do put a burden on the designer in terms of determining the level of refinement [4].

The aim of this study was to map medication data retrieved from individual patient health records for population health researcher’s use through the use of an abstraction network to create a compact and more easily understandable view for public health research.

## Methods

The modelling of medication data within PHReDMS was based on the Monthly Index of Medical Specialties (MIMS) [10]. In Australia MIMS is commonly used to define independent medicine information used by Australian healthcare professionals. Medications within MIMS are classified by:

- The body system for which the medicine is used, e.g. cardiovascular system
- Medication class, e.g. beta-adrenergic blockers
- Medication subcategory, e.g. rapid acting (under insulins)

- Route, e.g. inhaled, topical, anorectal
- Medication brand & generic name

Medications within patient records are often recorded with both Brand and Generic names being used interchangeably during the prescribing process. Whilst a category of medication may have utility across multiple medical conditions, MIMS medication subcategories are made up of drugs that are similar in terms of both function and formulation. The formulation of any given medication, in terms of its constituent compounds and their function, remains similar regardless of pharmaceutical producer, brand or generic naming. For example, beta blockers are a subcategory of anti-hypertensive drugs. Atenolol (brand names include Noten & Tenormin) and propranolol (commonly branded as Apo-Propranolol) are both beta blockers. The active ingredient in atenolol is always a benzacetamide compound named atenolol, and in propranolol it is propranolol hydrochloride. Both compounds/ active ingredients exert a therapeutic effect by blocking beta adrenergic receptors. For research purposes, we might want to analyse data at the level of beta receptor blockers, or at anti-hypertensives, or even at a higher level of all drugs that exert an action upon the cardiovascular system.

A working group comprising of a database administrator, a three person database development team, two researchers and a principal investigator met on an as needed basis to develop the medications mapping model. The researchers were qualified general medical practitioners, who have a number of years of experience with patient medical record information. An electronic extract of medications from a clinical information system was compared to the MIMS database by the researchers. After primary classification of brand and generic medications by the researchers, the working group utilised an iterative process over multiple meetings to clarify the mapping. Medications within the PHReDMS were mapped based on active ingredients/constituent compounds (See Table 1 for an example). By basing categorisation on the common constituent compounds (often listed in the generic medication name) instead of nomenclature (brand or generic medication names), the working group were able to map medication into subcategories and categories to enable analysis on multiple levels from a single data collection, as well as aid categorisation of data from varied sources.

Medications within the PHReDMS are mapped based on active ingredients/constituent compounds (Table 1). The MIMS database was used to link generic and brand names to active ingredients, with generic names often mirroring the active ingredient present. PHReDMS does not store the name of branded drugs, it only stores the ingredients. A custom built sql script extracting as many ingredients as are listed against the branded drug in Communicare. This was usually no more than three active ingredients. The PHReDMS attaches an active ingredient to any number of subcategories. Subcategory descriptors relate either to function or route of administration. Categories can be constituted of multiple subcategories and multiple categories can be used to treat a particular health condition. An exception column exists to detail any medications that deviate in function from the rest of their class.

Medication categories recorded in PHReDMS were examined, with medications focused to the major medication classes and of interest to the research group. However this model can be scaled to include more categories of medication as required. Each medication category contains an Other subcategory, to classify any new medications that do not fall into pre-existing subcategories. There is also an overall Other category, to classify any medications extracted that do not fall into the limited categories as defined in PHReDMS. New medications which fall into existing subcategories or categories can be added by the researcher forwarding their details to the database administrator.

**Table 1: Medication Mapping Model in PHReDMS**

Brand name	Generic name/ active compound	Function	Medication subcategory	Medication category
Apo carvedilol	carvedilol	Non selective beta blocker	Beta blocker	Anti-hypertensive medication
Dilatrend	carvedilol	Non selective beta blocker	Beta blocker	Anti-hypertensive medication
Coveram	amlodipine		Calcium channel blocker	Anti-hypertensive medication
	Perindopril arginine		ACE inhibitor	

## Results

Initially there were 7416 medications extracted from the Communicare from which were derived a list of 601 compounds. These compounds were then mapped into 63 sub-categories and 19 categories. Nutritional supplements have been collapsed to just their generic compounds as they are primary data for the researchers using the PHReDMS (Table 2).

**Table 2: Medication Categories Recorded in The Public Health Research Data Management System.**

Category	Subcategory	
Full name	Full name	Abbreviation
Insulins		
	Short acting insulins	Short
	Long acting insulins	Long
	Intermediate acting insulins	Intermediate
	Rapid acting insulins	Rapid
	Mixed insulins	Mixed
Oral hypoglycaemic agents		
	Biguanides	
	Sulphonylureas	
	Thiazolidinediones	
	Alpha-glucosidase inhibitors	
	Meglitinides	
	Dipeptidyl peptidase-4 inhibitors	DPPIs
Antihypertensives		

	Beta blockers	B-blockers
	Angiotensin Converting Enzyme inhibitors	ACE Inhibitors
	Calcium channel blockers	Ca channel blockers
	Centrally acting antihypertensive agents	Centrally acting
	Sympatholytics	
	Other Antihypertensives	
	Angiotensin 2 receptor antagonists	A2 antagonists
<b>Vasodilators</b>		
	Nitrates	
	Other Vasodilators	
<b>Diuretics</b>		
	Thiazides	
	Potassium sparing diuretics	K+
	Other diuretics	
	Loop diuretics	Loop
<b>Antiarrhythmics</b>		
	Class I	
	Class IA	
	Class IB	
	Class IC	
	Class II	
	Class III	
	Class IV	
	Class V	
<b>Antibiotics</b>		
	Penicillins	
	Macrolides	
	Quinolones	
	Rheumatic Heart disease	RHD
	Other antibiotics	
<b>Nutritional Supplements</b>		
	Fatty acids	

	Vitamins	
	Minerals	
	Other supplements	
	Folates & derivatives	Folates
Vaccines		
Urinary Acidifiers		
Urinary Alkalinisers		
Hypolipidaemics		
	Serum cholesterol reductors	
	3-hydroxy-3-methylglutaryl-coenzyme (HMG-CoA) reductase inhibitors	HMG CoA inhibitors
	Fibrates	
	Intestinal absorption inhibitors	
	Bile acid sequestrants	
	Other hypolipidaemics	
Anticoagulants		
	Anti Factor Xa	
	Antiplatelet aggregation	
	Antithrombotic	
	Thrombin Inhibitors	
	Coumarins	
	Superficial thromboses	
	Other anticoagulants	
Antifungals		
Corticosteroids		
	Glucocorticoids	
	Mineralcorticoids	
	Other corticosteroids	
Alpha blockers		
Anti-asthmatics		
	Beta2 agonists	
	Anticholinergics	
	Inhaled corticosteroids	
	Other bronchodilators	

Anaesthetics		
Other		

## Discussion

One way to confront the “big knowledge” challenge is to provide auxiliary support structures to aid in terminology use and maintenance [4]. Highly interactive partially-automated mapping tools that are directed by human input to automate parts of the mapping process, with specific expert input, are a promising alternative to fully automated methods [11]. In this study we were able to successfully use an abstraction network to map medication data from patient electronic medical records. This enables patient information to be easily utilised by population health researchers.

Patient health information is largely stored in EMRs, which are used by most health care providers for the regular documentation of care provided. Administrators also use EMRs for purposes of reporting upon and monitoring activities in health facilities. Whilst some patient data is still stored in paper format, researchers are increasingly asking to access EMR data for the purposes of population level research, including the development of clinical decision support tools. The vast amounts and complexity of information within EMRs complicates information management for each of these groups and increases the risk of incorrect decisions being made due to the difficulties of sifting through so much information. Likewise with paper records, the amount, variability and complexity of information stored within them makes it difficult to collect and analyse this data in a reliable manner. Managing medications data is a particular challenge; as the burden of chronic disease increases in the population [12] so do issues related to polypharmacy and its management [13] [14] [15]. Clinicians and researchers need new tools to manage medication- related information. Although there can be an overwhelming number of medications and medicine classes prescribed per patient, pharmacology is relatively structured when it comes to classification of the active ingredients or constituents of medication. Additionally, various therapeutic guidelines are available that are already in use in practice to thoroughly document and categorise medication, ensuring relatively easy access to the expert opinions required to inform any medications mapping.

For research purposes, data is sourced from various EMRs and paper patient files, which capture medications data related to MIMS in different formats. A research database like PHReDMS needs to be flexible enough to contain medications data;

1. From various sources, which may or may not be inter-relatable
2. Across various topics, which are externally set by researchers accessing the database
3. That reflects changes over time

## Limitations

The development of the abstraction network to manage medication records for individual patient’s for chronic disease management research has been extremely labour intensive. The abstraction network requires updating before new patient records are added to an individual research project to capture new medications that have been recently released into the population.

## Conclusion

Mapping in the manner described above has enabled data to be categorised to allow researchers to query and analyse it at multiple levels (disease grouping, category, subcategory, and compound) relevant to their particular research question, without compromising the richness of individual level medications data. It allows all data collected by various researchers from various sources to be

aggregated into one database, thus providing a valuable source of population level data. The mapping process is geared towards population level chronic diseases medication data, and will need to be expanded for uses outside of this scope (e.g oncology) however the system is flexible enough to enable this as may be required in the future.

## Acknowledgements

The research reported in this paper is a project of the Australian Primary Health Care Research Institute, which is supported by a grant from the Commonwealth of Australia as represented by the Department of Health. The information and opinions contained in it do not necessarily reflect the views or policy of the Australian Primary Health Care Research Institute or the Australian Government Department of Health.

## References

1. Forman MR, Greene SM, Avis NE, Taplin SH, Courtney P, et al. 2010. Bioinformatics Tools to Accelerate Population Science and Disease Control Research. *Am J Prev Med.* 38(6), 646-51. [PubMedhttp://dx.doi.org/10.1016/j.amepre.2010.03.002](http://dx.doi.org/10.1016/j.amepre.2010.03.002).
2. Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, Quesenbery W, Shneiderman B (2011) Usability and Accessibility in consumer Health Informatics Current Trends and Future Challenges, *American Journal of Preventive Medicine* 2011; 40(5S2): S187-S197
3. de Lusignan S, Teasdale S. 2007. Achieving benefit for patients in primary care informatics: the report of a international consensus workshop at Medinfo 2007. *Inform Prim Care.* 15(4), 255-61. [PubMed](#)
4. Halper M, Gu H, Perl Y, Ochs C. 2015. Abstraction Networks for Terminologies: Supporting Management of "Big Knowledge". *Artif Intell Med.* 64, 1-16. [PubMedhttp://dx.doi.org/10.1016/j.artmed.2015.03.005](http://dx.doi.org/10.1016/j.artmed.2015.03.005).
5. McBride SJ, Lawley MJ, Leroux H, Gibson S. 2012. Using Australian medicines terminology (AMT) and SNOMED CT-AU to better support clinical research. *Stud Health Technol Inform.* 178, 144-49. [PubMed](#)
6. Nachimuthu SK, Lau LM. 2007. Practical issues in using SNOMED CT as a reference terminology. *Stud Health Technol Inform.* 129(Pt 1), 640-44. [PubMed](#)
7. Schnipper JL, Gandhi TK, Wald JS, Grant RW, Poon EG, et al. 2012. Effects of an online personal health record on medication accuracy and safety: a cluster-randomized trial [doi.]. *Journal of the American Medical Informatics Association: JAMIA.* 19(5), 728-34. <http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/178/CN-00854178/frame.html> and <http://jamia.bmj.com/content/19/5/728.full.pdf>. [PubMedhttp://dx.doi.org/10.1136/amiajnl-2011-000723](http://dx.doi.org/10.1136/amiajnl-2011-000723)
8. Ochs C, Agrawal A, Perl Y, Halper M, Tu SW, et al. 2012. Deriving an abstraction network to support quality assurance in OCRE. *AMIA Annu Symp Proc.* 2012, 681-89. [PubMed](#)
9. Ochs C, Perl Y, Geller J, Haendel M, Brush M, et al. 2015. Summarizing and visualizing structural changes during the evolution of biomedical ontologies using a Diff Abstraction Network [doi.]. *J Biomed Inform.* 56, 127-44. [PubMedhttp://dx.doi.org/10.1016/j.jbi.2015.05.018](http://dx.doi.org/10.1016/j.jbi.2015.05.018)

10. Specialties MM. (MIMS), accessed: 26th November 2012, <http://www.mims.com.au>
11. Saitwal H, Qing D, Jones S, Bernstam EV, Chute CG, et al. 2012. Cross-terminology Mapping Challenges: A Demonstration Using Medication Terminology Systems. *J Biomed Inform.* 45, 613-25. [PubMedhttp://dx.doi.org/10.1016/j.jbi.2012.06.005](http://dx.doi.org/10.1016/j.jbi.2012.06.005).
12. Haagsma JA, Graetz N, Bolliger I, Naghavi M, Higashi H, et al. Murray CJL, Vos T. (2015). The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study 2013. *Inj Prev.* doi:10.1136/injuryprev-2015-041616.
13. Hubbard RE, Peel NM, Scott IA, Martin JH, Smith A, et al. 2015. Polypharmacy among inpatients aged 70 years or older in Australia. *Med J Aust.* 202(7), 373-77. [PubMedhttp://dx.doi.org/10.5694/mja13.00172](http://dx.doi.org/10.5694/mja13.00172).
14. Martin JH, Henry D, Gray J, Day R, Bochner F, et al. 2015. Achieving the world health organization's vision for clinical pharmacology. *Br J Clin Pharmacol.* doi:10.1111/bcp.12803.
15. Wimmer BC, Bell JS, Fastbom J, Wiese MD, Johnell K. 2015. Medication Regimen Complexity and Polypharmacy as Factors Associated With All-Cause Mortality in Older People: A Population-Based Cohort Study. *Ann Pharmacother.* doi:10.1177/1060028015621071.