Patient Information Model to Support Population-level Workload Analysis

Manjula Devananda  
*University of Otago*, manjula.devananda@postgrad.otago.ac.nz

Stephen Cranefield  
*University of Otago*, stephen.cranefield@otago.ac.nz

Hywel Llyod  
*University of Otago*, hywel@bpac.org.nz

Michael Winikoff  
*University of Otago*, michael.winikoff@otago.ac.nz

Follow this and additional works at: https://aisel.aisnet.org/acis2017

**Recommended Citation**
https://aisel.aisnet.org/acis2017/15

This material is brought to you by the Australasian (ACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ACIS 2017 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Patient Information Model to Support Population-level Workload Analysis

Manjula Devananda
Department of Information Science
University of Otago
Dunedin, New Zealand
Email: manjula.devananda@postgrad.otago.ac.nz

Stephen Cranefield
Department of Information Science
University of Otago
Dunedin, New Zealand
Email: stephen.cranefield@otago.ac.nz

Hywel Lloyd¹
Department of General Practice and Rural Health
University of Otago
Dunedin, New Zealand
Email: hywel@bpac.org.nz

Michael Winikoff
Department of Information Science
University of Otago
Dunedin, New Zealand
Email: michael.winikoff@otago.ac.nz

Abstract

Current patient information models are able to support individual patient-level analysis. However, population-focused care is gaining importance due to an increase in the number of people with long-term conditions. A population-focused care addresses the care needs of a group of patients who share a common trait. Primary health care (PHC) being the first point of contact with a health system, our previous work aims to predict and analyse this population-level workload at a PHC centre. As a part of this work, we identified that current patient information models lack the ability to support population-level analysis. In this paper, we discuss an extended ontology of patient information models to support population-level workload analysis. We describe the three cycles of Design Science Research (DSR) applied to develop our ontology. Then, we discuss the existing health data models. Thus, this paper makes a domain-specific application of DSR to develop a patient information model that supports population-level analysis.

Keywords Population-level analysis, patient information model, long-term conditions, ontology, design science research.

¹ Hywel Lloyd is also Director of Informatics at Best Practice Advocacy Centre (BPAC) New Zealand.
1 Introduction

An aging population and the increasing prevalence of long-term conditions (LTCs) are major challenges for health care systems globally. A key property of a high performing health care system is the use of a population-based approach to manage its LTC patients (Ham 2010). We aim at better primary health IT systems for decision support at a population-level. Our work therefore focuses on population-level management of LTC patients. Furthermore, since a primary health care (PHC) setting is the first point of contact with the health system in most developed countries (Montague 2014), we therefore focus on the PHC context.

There has been a range of work on using IT to manage patients’ health from a population-wide perspective (Dale 2015; Devananda et al. 2017; Hefford 2006; The European Society of Cardiology 2002; WHO 2005). However, despite information technology advances, public health care systems need to continue to reorient themselves to meet the challenge of supporting and maintaining the health of a growing number of people with long-term conditions (Mays 2013). In doing so, it is important to understand the efficiency of PHC from a population perspective.

In this paper, we address the information requirements for a population-level assessment of efficiency in a PHC context. We focus specifically on extending a patient information model by developing an ontology to support population-level workload analysis. This extended primary health care information model has been applied to construct a rule-based workload prediction simulation system (Devananda et al. 2017). In section 2, following Hevner’s (2007) “three cycle view” of a design science research, we discuss why and how we extended the existing ontology and how it can support population-level analysis in the context of PHC centre\(^2\). Later, we present the existing health data models. We conclude in section 4. Thus, this paper makes a domain-specific application of Design Science Research (DSR) to develop a patient information model.

2 Our Design Science Approach

Design Science is an iterative problem-solving paradigm that is applied in the fields of engineering and sciences (Hevner et al. 2004). In this section, we describe our ontology and how it can be used for a population-level analysis. Figure 1 below shows our DSR approach to develop the ontology in a “Three Cycle View” (Hevner 2007). The Environment is the application domain, here the PHC context. The environment describes the people (here general practitioners (GPs), nurses and other care providers at a PHC), organisational systems, technical systems, and the problems or shortcomings in the application domain. The design cycle in DSR involves building the design artefacts and processes and evaluating them in an iterative fashion. The relevance cycle maps these artefacts and processes with respect to the application domain. The rigour cycle contributes the results of design science research to the knowledge base. The knowledge base has existing scientific theories, models or methods, experience and expertise and meta-artefacts, to which design science research contributes.

2.1 The Environment and Relevance Cycle.

Even though efforts are in place to plan care delivery and effectiveness at a population level (Ministry of Health 2017a) there has been little work on understanding the increasing workload of primary care in New Zealand. A paper published last year in the Lancet (Lovibond et al. 2011) demonstrated a 12-36% increase in the standardised consultation rate for British general practice for a seven-year period to 2014. Part of the health system solution to the growing demand for care across the system is to reduce costs and shift care from secondary to primary-based services (Hall et al. 2002). Handling the increased workload without a significant increase in funding relies on efficiency gains achievable by practices and primary health care organisations (Ernst & Young 2017).

2.1.1 A PHC Context

Traditionally, patient initiated appointments drive primary care workload (Green Cross Health 2017). In broad terms, these appointments can be LTC-related or acute care appointments. The acute care episodes can be for people with long-term conditions and may reflect an acute exacerbation of a long-term condition or an acute care episode for someone without a long-term condition. However, current systems do not give much information in an analysable form to understand if a specific appointment slot is for a long-term condition treatment review or a new problem not linked to an LTC.

\(^2\) This was not discussed in our previous paper.
PHC centres have patients registered with one of their GPs. Other roles that may be involved in care delivery are practice nurses, enrolled nurses, community workers and other carers. Traditionally individual patient outcomes are seen as a measure of effectiveness of care; however managing care of a population as a whole has been an area of improvement since the 1990s (Weiss 1998). A population-level approach views the individual patient within the context of the larger community. For example, primary care providers use similar treatment plans for populations with specific LTCs or similar health conditions (Loyola University 2016).

2.1.2 LTC Management

LTC patients require timely interventions, which include clinical reviews and laboratory tests, and may span across various care provider roles (Oldroyd et al. 2003; Burt et al. 2014). These LTC management needs specific to a particular LTC are formulated as its care pathway (Burt et al. 2014; Reeves et al. 2014). Usually care pathways are represented as flow charts. Each path on the flow chart specifies the actions to manage a particular severity of the condition and is implemented through care plans for each individual patient with the condition (Amir et al. 2015; Sox and Stewart 2015). A plan-of-care, otherwise called a treatment plan, is the agreed care delivery plan for each individual patient. This must consider all the care plans applicable to the patient considering all their conditions (Burt et al. 2014).

Currently, PHC systems have a capability for predicting what interventions need to happen for an individual within the primary prevention domain e.g. which laboratory tests and when should they be done for individual patients for early detection and prevention of a disease. However, the recall rules, i.e. specifying when the next patient visit should be scheduled, are condition or task-specific, such as immunisation and cervical smear taking. There is a need to change this from condition-specific recalls, to merge all the conditions then present in the patient, and drive the care planning using a single plan of care. This care planning should also consider various roles in PHC, like GPs, practice nurses and receptionists, and their competencies in order to schedule care.

2.1.3 Shortcomings of Existing Primary Health Care Information Model

Health Care models continue to focus on acute care. The structure of health systems reinforces the acute model with fixed 8, 10 and 15-minute appointment slots (Iacobucci 2016). This reinforces a reflex response to acute crisis interventions. A more systematic proactive response for the LTC population is required (Reilly 2013). However, primary health care information systems and their database architectures are structured to support acute ‘siloed’ unplanned care (Marshall et al. 2016; Salgado et al. 2016). In particular, while working with an anonymised database from a PHC\(^3\), we have identified some information and relationships missing from that database that were required to facilitate population-level analysis, including the following:

---

\(^3\) The University of Otago Human Ethics Committee (Health) has given approval for this research.
There was no explicit representation of whether a patient is an LTC patient.
There was no explicit representation of whether appointments and prescribed medications were related to acute conditions or to management of LTCs in a patient.
It is difficult to determine which treatment path of a “care pathway” (generic treatment plans for given conditions) was followed for a patient.
Care pathways have different options (e.g., how often the patient should be seen) depending on the severity of the condition. However, condition severity was not explicitly represented and may require an analysis of the results of lab tests over a period of time.
There was no explicit representation of the individualised plan of care for the patient, which merges and resolves incompatibilities between the recommendations of different care pathways.

We had to infer knowledge to fill the gaps due to the missing data in the dataset. For instance, in order to find how many diabetic patients are treated with more than two LTC medications, we had to infer which are the LTC medications and then find which patients are classified by Read codes that belong to the hierarchy of Diabetes. The discussion of how this inference was made is out of scope for this paper. In summary, the existing information model underlying our patient data needed to be extended to meet these shortcomings listed above and support a population-level analysis.

2.2 Design Cycle

In this section, we introduce our ontology, and a few basic concepts and relationships of our ontology. We borrow a few basic concepts and relationships from openEHR, HL7 RIM, SNOMED-CT and a work by Mabotuwana & Warren (2009). Figure 2 represents those concepts and relationships needed for a population-level workload analysis as an ontology modelled in OWL4 (OWL 2012). The additions to the existing patient information model are highlighted using rectangles.

The isLTC attribute shows the LTC status of the entity. Box A shows those entities relating to LTC (e.g. Appointment, Classification). The patientId is an attribute of a patient, plan of care, appointments, classifications, care actions and population. Box B shows this relationship.

![Figure 2: Our ontology, with our additions indicated by boxes.](image)

A health practitioner may hold various roles within the practice. These roles, such as practice nurse and chief medical officer, in a PHC context are shown as instances of a role. A health practitioner can be competent to manage a specific health condition, denoted by isCompetentIn a ProblemClass. A patient isRegisteredWith a health practitioner and may have an encounter with a general

---

4 We developed this ontology using Protégé and created the diagram using the ontology-viewing tool, OWLGrEd.
practice via an appointment with a health practitioner, a re-issue of medications or a lab test. Hence, an appointment isBookedBy a patient and isBookedFor a health practitioner.

A care action is a medical activity that isToManage a health condition in a patient. A care action involves a health practitioner; for instance, a GP prescribes the medication, and a nurse does lab tests. During a consultation, a patient may be diagnosedWith a classification or addresses a classification within a patient. Every classification of a patient will have a diagnosis with severity and date of classification, which help to mine the medical history of a patient. The classification will have isLTC, ProblemClass, severity, and the ProblemCode of the condition along with date of classification. In our ontology, if a patient does not have a medical condition, the patient is classified as a normal patient. So normal is shown as an instance of ProblemClass. A normal patient will follow the generic care pathway for prevention and early detection of LTCs.

To determine the enacted care plans in a patient, we broadly classify the patients using ProblemClass. A ProblemClass (has a genericName) addresses a medical condition like diabetes, and thus refers to various care plans for its care pathway. We analysed the care pathways, derived the care plans based on the interventions required and the period within which the intervention must be made in a patient, captured as a recallPeriod within a care plan. For example, a newly diagnosed patient with a classification C109 (a Read code that classifies as Diabetes) with a reading value greater than 55 for 44TB (Read code for HBA1C blood test), may be monitored on 3-monthly basis while patient whose diabetes is stable with same HBA1C and is on medication, will be monitored every 6 months. This intervention based on severity is captured as different care plans. Thus, a care pathway isComposedOf care plans. The ProblemClass and the severity of the condition in a patient identifies the care plan for the patient. A ProblemClass will have ProblemCodes associated with it. These ProblemCodes can be Read codes, SNOMED-CT or ICD-10 codes. We also emphasise that ProblemClass is not a replacement for clinical codes like read codes, but is additional information and hence it provides more generic information required by both clinical (GPs and nurses) and non-clinical (clinical decision support developers) users of the clinical data.

Every LTC patient may be on one or more care plans depending on the comorbidities in a patient, and so may refer to more than one recallPeriod. However, every patient will have only one plan of care, which considers all the LTCs present in the patient. Every patient, thus, will have only one recall period, leastRecallPeriod. The most severe condition will require interventions that are more frequent and hence the plan of care considers that most severe condition drives the recalls (Amir et al. 2015; Burt et al. 2014; Dennis et al. 2008). The plan of care hasActions from care actions. Aggregation of various care plans into a patient specific plan of care ensures that medical activities, like medications and lab tests, are not duplicated for a patient unnecessarily (Burt et al. 2014). Thus, during the consultation, a health practitioner considers all the medical conditions present in the patient, and prescribes medication or orders lab tests. At times, this would also enable a patient to book for the follow-up appointment. In other words, the various care plans mergesInTo a plan of care and the care actions depend on this plan of care for a patient. We highlight that this information required to draw conclusion on a plan of care for a patient is not addressed in the existing health data models. We have also considered to represent if a prescription for a drug isActive and validFor a certain period. For instance, a prescription for paracetamol may be active for a week i.e patient is on medication for a week but the validity of the prescription could be for 3 days, i.e. the patient will require a new prescription, if they need to collect the medication after three days.

Based on some criteria, a patient belongsTo a population. Though we discussed this ontology from the perspective of LTCs, this ontology can support acute or non-LTC specifications too. Care action could relate to non-LTC events like screening for pre-diabetes, annual health check-up or even encounters due to accidents. We emphasise that these additions to the existing ontology are highlighted with rectangular boxes in Figure 2 above.

### 2.3 The Knowledge Base and Rigour Cycle

Existing health information models support reasoning over details at an individual patient level (see section 3 below). We extended this patient information model to support a population-level view. Our information model has been applied in the construction of a rule-based simulation system which,

---

5 Population, in this paper, refers to a set of patients registered with a practice or a set of patients with specific LTCs.
given patient coding laboratory results and health care measurements such as blood pressure, predicts the workload for the next year from the LTC patients registered at a PHC centre (Devananda et al. 2017).

Here we discuss a few examples of how our extended model supported population level view. These have been implemented, and use real data.

1. Simulated workload.
Figure 3 shows the workload\(^6\) predicted for 2014 given patients’ historic visit dates in 2013 and compares this to the workload from these patients was historically in 2014\(^7\), considering all their appointments and those appointments attended by a GP or a nurse at the practice (“C_NC”). The gap between the predicted workload and the actual appointments are of interest but discussion of that is beyond the scope of this paper.

![Figure 3: Workload comparison](image)

2. LTC Workload per role.
Figure 4 shows for each health practitioner at the PHC, out of their total workload in 2014, how many appointments were related to LTC appointments. This can be used to plan resourcing required to meet the predicted workload.

These population level analyses communicated to domain experts\(^8\) resulted in feedback that the tool is useful for a general practice, especially to schedule care delivery knowing the capacity of the practice. They responded that this tool would be very useful to address the changes needed if the practice amends its organisational structure or policies such as shift of care among roles, involving multidisciplinary team, at a population-level.

3 Existing Health Data Models

In this section, we briefly review existing work on health data models. Prior work falls broadly into four areas: standard classification schemes for clinical terminology, standard data models for electronic patient information, models for LTC management needs (e.g. generic treatment plans for given conditions), and “ontologies” that allow reasoning with health information from multiple sources.

The importance of standardising the representation of clinical terminology dates back to the 16th Century (Chute C.G 2000). In modern times, much effort has been put into developing national standards for clinical terminology coding systems such as Read codes (NHS Digital n.d.), standardised in the UK, the SNOMED Reference Terminology (Spackman et al. 1997), developed in the US, and the result of merging and expanding these two coding systems: the SNOMED Clinical Terms (Ministry of Health 2017b). The World Health Organisation has developed an international standard coding system, ICD-10 (WHO 2016), for “the universe of diseases, disorders, injuries and other related health conditions” which is designed to support the reporting of mortality and morbidity statistics. These

---

\(^6\) We simulated future visits for the next year based on a steady-state view of the population. We do not simulate change in health of an individual patient.

\(^7\) For actual use we would be predicting workload for a future period, and would not have actual visit data.

\(^8\) These are various domain experts including GPs, health operation officers, and chief medical advisers for a PHO.
terminologies all define a hierarchical classification system with codes associated with each entry. These codes are commonly used within health records as a standard and unambiguous representation of a condition, body part, medical test, etc. As such, they do not attempt to define where this information might appear in a database and what context relationships might exist in a database between that condition and other entities (such as medications, consultations and procedures). However, SNOMED CT is specified as an ontology that defines concept hierarchies, attributes of concepts, and relationships between them. Its concepts cover clinical findings, symptoms, diagnoses, procedures, body structures, and other factors relevant to documenting specific patient conditions, rather than patient encounters, prescriptions, etc.

The second area of work is concerned with standard data models for electronic patient information, in order to allow for data exchange between health IT systems (Tsiknakis et al. 2002). For example, the standards body Health Level Seven International (HL7) has developed a family of standards including the HL7 Reference Information Model (HL7 2016, p. 7). This includes the key concepts of entities (e.g. people, organisations and places) acting in roles, while participating in acts (e.g. patient encounters, observations and procedures), which may have relationships with other acts. This model is very general at the high level, while also providing expressiveness through the existence of many subclasses of the high-level concepts (e.g., there are 17 different subclasses of Act). Similarly, the openEHR Electronic Health Record (EHR) Information Model (openEHR 2016) is a high-level model that makes no commitments about the relationships between information items. Figure 5 shows the part of the EHR Information Model that models an entry in an EHR. This provides a classification of entry types, but does not define any possible relationships between them. Both the HL7 and openEHR models can be specialised with domain-specific profiles (for HL7) or archetypes (for openEHR). However, such specialisations do not seem to exist for the area of primary health care.

The US Office of the National Coordinator for Health Information Technology has developed the Quality Data Model (QDM) (ONC 2016). This defines the structure for several types of data items related to the care of a patient, as well as a language for defining expressions expressing measures of the patient’s care, e.g. that certain medication is prescribed within a specified period after a given lab test. The models intend to support clinical quality improvement processes. The data items modelled cover encounters (ordered and performed), diagnoses, medication that is currently prescribed, as well as medication dispensing and administering events, and laboratory tests, amongst others. The defined attributes include some relationships such as the diagnosis addressed during an encounter and the severity of a diagnosis or a symptom. These relationships are important in our problem domain and are included in our model.
Kawamoto et al. (2012) report the results of a needs analysis for the development of an HL7 common information model for clinical decision support, known as a virtual medical record (vMR). Of interest to our work is the presence of both a problem code and problem class(es) amongst the suggested data elements. The report lacks definitions of these terms, but the examples suggest that the difference is in the degree of specificity, with the problem code being a more general description, e.g. “cardiovascular disease”. This distinction does not appear in the vMR specification, in which only a problem code appears, defined as “the code that identifies the problem or condition with as much specificity as available, or as required by a template” (HL7 2016, p. 7). While a hierarchical coding system allows the use of codes at any level of a classification hierarchy, we believe that for the needs of population level analysis there is a need to explicitly represent a problem class selected from some pre-defined set of high-level classes, which plays a different role than specific problem codes. The vMR specification includes a severity attribute for a problem, which is a feature that we believe is useful in our context of population-level reasoning.

The third area of related work addresses LTC management needs. The LTC management needs for specific conditions are captured in respective care pathways (Best Practice Advisory Centre New Zealand 2012). Clinical guidelines are viewed as the textual representation of care pathways, which can then inform individual care plans. Clinical guidelines aim to reduce variation in clinical decision-making across various care providers. Computer-interpretable guideline languages (CIGLs) use computer-based reasoning to apply clinical guidelines at a medical practice (Peleg 2013). Adopting CIGLs ensures uniformity in clinical decision-making. There are major clinical guideline repositories such as the National Institute for Health and Care Excellence (NICE, UK), the National Guideline Clearinghouse (NGC, USA) and the Guideline International Network, G-I-N n.d.; NICE n.d.; NGC n.d.). Notations for representing care pathways include Proforma (Fox et al. 1997), GLIF and SDA (Peleg 2013). O’Leary, Noll and Richardson (2013) use resource flow to model care pathways. Mulyar et al. (2007) and Peleg et al. (2003) compare several notations for care pathways. Data models for care pathways tend to focus on generic concepts such as actions and events, and lack key medical concepts such as patient classification and medications. Thus, they do not consider the information and relationships we seek to express in our data model.

The fourth class of work that we consider is the application of semantic web or linked data technology to integrate health information from different sources. This work involves using an ontological modelling language such as the Web Ontology Language (OWL Working Group 2012) to define an ontology (a high level model of the concepts and relationships that are of interest in a problem domain), which allows an integrated dataset to be built by importing information from multiple sources of health data. The representation of health information in terms of an ontology language also supports the use of automated reasoning processes, such as the execution of rule-based decision support systems (Riaño et al. 2012). This work tends to focus on specific problems that can benefit from the use of integrated data. For example, Hong et al. (2016) define an ontology to support the authorship of rule-based clinical diagnostic criteria, by integrating the WHO’s International Classification of Disease (ICD-11) and the Quality Data Model (QDM). Riaño et al. (2012) have defined a detailed ontology to support the generation of personalised patient information records (including health and social factors) and formal intervention plans for the coordinated home care of chronically ill patients by multiple health-care professionals and caregivers. Their ontology integrates concepts from ICD and a chemical classification scheme. The importance to temporal relationship between prescribing a medication and the recall for a patient has been considered by Mabotuwana & Warren (2009) to develop an ontology to manage a population of hypertensive patients. This work focuses on a specific LTC while we are considering multiple conditions at the same time. Moreover, we are looking at the activities to manage various LTCs present in a patient rather than just the medications for managing the condition as a whole.

4 Conclusion

Improving efficiency of PHCs is a challenge given aging populations, increases in LTCs and fewer care providers (Ernst & Young 2017). We approach this global problem starting from a PHC in New Zealand and understanding what a primary care workload is like. We observed that even though the data models in health information systems successfully recorded all information related to patients and health conditions, population-level analysis was impossible, without multiple interventions. This paper’s contribution is an ontology that contains the elements required to support population-level workload analysis at a PHC.
5 References


Copyright: © 2017 Manjula Devananda, Stephen Cranefield, Hywel Lloyd & Michael Winikoff. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 Australia License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and ACIS are credited.