E-Health Technologies and Improving Patient Safety:

Exploring Organizational Factors

Anastasius Moumtzoglou
P. & A. Kyriakou Children's Hospital, Greece

Anastasia Kastania
Athens University of Economics and Business, Greece

Medical Information Science Reference

Managing Director: Lindsay Johnston
Editorial Director: Joel Gamon
Book Production Manager: Jennifer Yoder
Publishing Systems Analyst: Adrienne Freeland
Development Editor: Myla Merkel
Assistant Acquisitions Editor: Kayla Wolfe
Typesetter: Deanna Jo Zombro
Cover Design: Nick Newcomer

Published in the United States of America by
Medical Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

E-health technologies and improving patient safety : exploring organizational factors / Anastasius Moumtzoglou and Anastasia Kastania, editors.
pages cm
Summary: “This book presents an overview of information and communication technologies and their impact on the field of both patient safety and e-health and highlights a better understanding of their interaction”-- Provided by publisher.
Includes bibliographical references and index.
ISBN 978-1-4666-2657-7 (hardcover) -- ISBN 978-1-4666-2688-1 (ebook) (print) -- ISBN 978-1-4666-2719-2 (print & perpetual access) (print) 1. Medical informatics. 2. Patients--Safety measures. I. Moumtzoglou, Anastasius, 1959- editor of compilation. II. Kastania, Anastasia, 1965- editor of compilation.
R858.E34 2013
610.285--dc23

2012029409

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 16
An Ontology-Driven Approach to Clinical Evidence Modelling Implementing Clinical Prediction Rules

Derek Corrigan
Royal College of Surgeons in Ireland, Ireland

Adel Taweel
Kings College London, UK

Lucy Hederman
Trinity College Dublin, Ireland

Olga Kostopoulou
Kings College London, UK

Haseeb Khan
Trinity College Dublin, Ireland

Brendan Delaney
Kings College London, UK

ABSTRACT

Diagnostic error is a major threat to patient safety in the context of the primary care setting. Evidence-based medicine has been advocated as one part of a solution. The ability to effectively apply evidence-based medicine implies the use of information systems by providing efficient access to the latest peer-reviewed evidence-based information sources. A fundamental challenge in applying information technology to a diagnostic clinical domain is how to formally represent known clinical knowledge as part of an underlying evidence repository. Clinical prediction rules (CPRs) can provide the basis for a formal representation of knowledge. The TRANSFoRm project defines the architectural components required to deliver a solution by providing an ontology driven clinical evidence service to support provision of diagnostic tools, designed to be maintained and updated from electronic sources of research data, to assist primary care clinicians during the patient consultation through delivery of up to date evidence based diagnostic rules.

INTRODUCTION

Diagnostic error has been shown to be a major threat to patient safety, particularly within the context of the primary care setting (Elder & Dovey, 2002). Where diagnostic errors occur there is the potential for an adverse event to take place that may undermine the safety of the patient, sometimes with disastrous results (Fisseni, Pentzek, & Abholz, 2008). Proponents of evidence-based medicine have analysed the diagnostic process itself and described the flaws they see as existing in traditional diagnostic approaches that can contribute to the possibility of adverse events

DOI: 10.4018/978-1-4666-2657-7.ch016

taking place (Kostopoulou, Delaney, & Munro, 2008). Steps have been proposed that describe the diagnostic process and are seen as requirements for implementing a more rigorous diagnostic process (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996a). The formal steps describe a methodology which advocates obtaining, evaluating and utilising the best available clinical evidence that has been previously demonstrated to be effective and pertains to each clinical case as it arises (Knottnerus & Buntinx, 2002).

If we are to practice evidence-based medicine effectively there is a need to efficiently and easily access the latest validated clinical guidelines that have been shown to be applicable to the diagnostic context and circumstances of the particular case being considered. The use of information technology and the internet for this purpose has greatly assisted in this task and its use has been formalised as part of the steps of the evidence-based medicine philosophy (Greenhalgh, 2010). Whilst this approach has been useful it can still be manually intensive and not without its problems.

There are a number of widely proposed models of clinical knowledge but acceptance of these models has not been without question and more generic ontology driven approaches advocated as possible improvements (Smitha & Ceustersc, 2006). Dissemination of accepted clinical guidelines is often done in the form of peer-reviewed hard-copy texts, soft-copy textual documents or textual web pages. It is difficult for consistent interpretation and representation of those guidelines that allow for their subsequent incorporation into integrated clinical information systems providing more sophisticated functionality such as clinical decision support. This problem can therefore by rephrased as follows: how can we more formally define and represent what constitutes recognised 'clinical knowledge' with a view to disseminating that knowledge more effectively for use by both clinical practitioners and information systems designed to support their use? The use of clinical prediction rules (CPRs) has long been accepted as a valuable means of deriving and disseminating clinical guidelines in respected evidence based journals such as the JAMA series (Ebell, Smith, Barry, Ives, & Carey, 2000). Because of their well defined structure they can be more easily incorporated and utilised in clinical decision support systems. CPRs can also be derived using well understood statistical methods such as logistic regression that could be applied to large sources of epidemiological data that are currently available in the form of both public and privately owned electronic research repositories.

Another problem is the time lag and delay that exists between clinical research outputs and their translation into frontline primary care practice (Lenfant, 2003). The ability of primary care practitioners to effectively practice evidence-based medicine is undermined by the delays that exist in disseminating accepted clinical best practice that has been derived from epidemiological based clinical research. This second question can be stated as: how do we expose our formal models of clinical knowledge with a view to easily integrating them with other systems that are utilised to bridge the gaps that exist between the research and primary care processes.

The work presented here is currently in development as part of a wider project called TRANSFoRm. The TRANSFoRm project is an EU funded FP7 project consisting of academic and industry partners who are developing an electronic infrastructure to support the vision of a rapid-learning healthcare system that integrates the research domain and the frontline primary care domain. A key development output of this project is the development of a decision support service that provides diagnostic decision support tools through a clinical evidence service utilised by primary care practitioners as part of the patient consultation. Integration with electronic health record systems will be achieved through the development of an ontology driven model of clinical evidence. This model for representation of clinical evidence allows for the utilisation of

clinical prediction rules as a mechanism to provide clinical decision support tools in the form of diagnostic advice. Underpinning this model is an analysis of the formal representation of electronic CPRs to allow for their derivation from currently available electronic sources of epidemiological data and their subsequent dissemination to clinical practitioners in the frontline using clinical decision support tools. TRANSFoRm aims in a real sense to provide an electronic infrastructure that supports the vision of bridging research and clinical practice through the use of translational medicine.

In this chapter we will examine the following areas as part of the aims and objectives:

- describe diagnostic factors which negatively impact on patient safety in the primary care setting
- present an overview of the architectural components required to deliver the complete TRANSFoRm solution and how it will address these diagnostic issues
- specifically discuss the provision of a TRANSFoRm clinical evidence service which provides an ontology-driven solution to support the utilisation of CPR based decision support that is designed to be maintained and updated from electronic sources of research data
- demonstrate how we can use the underlying clinical evidence model and ontology to answer diagnostic questions and describe its development to support the provision of our clinical evidence service

BACKGROUND

What is Meant by the Term Primary Care?

It is necessary to define at the outset what exactly we mean when we use the term 'primary care'. A number of definitions exist which focus on provision of primary care from subtly different perspectives that focus on the process or the clinician concerned, and may interchangeably use the term 'general practice' instead (Olesen, Dickinson, & Hjortdahl, 2000). For the purposes of this work the following definition from the Institute of Medicine captures the important context:

Primary care is the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practising in the context of family and community. (Donaldson, 1996, p.31)

We are therefore concerned with problems identified as part of the diagnostic process carried out by clinicians who typically provide first point-of-contact health services and advice to families in a frontline community setting.

Patient Safety and Diagnostic Error in Primary Care

Incidents that occur as part of day-to-day clinical treatment that ultimately result in a negative impact on the clinical outcome of any particular patient are commonly referred to as 'adverse events'. If we are to effectively propose solutions that contribute positively to patient safety it is necessary to identify the most problematic areas of patient treatment that could potentially result in an adverse event taking place. Where an adverse event occurs there is an associated impact on the patient resulting in a decrease from optimum levels of patient safety to a varying degree depending on the seriousness of that adverse event. The following definition provides a working description that interestingly also includes incidents which *may* have the potential for harm to a patient to take place: "An unintended event, no matter how seemingly trivial or commonplace that could have harmed, or did harm a patient" (Bhasale, Miller, Reid, & Britt, 1998, p.73)

Previous systematic reviews of the nature of adverse events have focussed on analysis of clinician reported adverse events in the primary care context and have proposed useful classifications of the major underlying problem areas that have been identified (Elder & Dovey, 2002). The classification proposed by Elder and Dovey attempts to address the two key questions of *where* adverse events occur most often in primary care and *why* those adverse events may have occurred.

Elder and Dovey identified and classified three main problem areas in relation to the question of *where* adverse events take place within the primary care process. These are:

- **Diagnosis:** errors relating to misdiagnosis through either missing a diagnosis entirely or mistakes which result in a delay in reaching a final correct diagnosis
- **Treatment:** errors relating to incorrect provision of treatment to a patient broadly broken into two areas of drug and non-drug treatment
- **Preventive Services:** errors relating to incorrect provision of preventative services to patients

With a clearer understanding of the potential causes of adverse events we can begin to develop solutions which aim to improve patient safety in the context of provision of primary care. The next step is to analyse these in more detail to establish why adverse events may be happening within each of these major causative areas.

In relation to the question of *why* adverse events take place within the clinical process, four potential factors for further investigation are described by Elder and Dovey. These are:

- **Clinician Factors:** factors directly attributable to decisions made based on the judgement and evaluation of the clinician regarding a particular case
- **Communication Factors:** factors relating to communication failures or misunderstanding between clinical entities collaborating as part of the overall patient treatment process
- **Administration Factors:** factors attributable to administrative errors in patient case management either by the clinician, other associated staff or third parties collaborating with them
- **External Factors:** impacting factors not directly in the control of the clinical consultation itself e.g. regulations, social and family factors.

This classification is useful but, as stated by the authors, does not attempt to quantify the individual contribution of each of these factors to potentially negatively impact on patient safety. Other studies have taken the logical next step and focussed on a specific identified category of error, such as clinician diagnostic error, and attempted to assess the contribution of that specific category of error as potential threats to patient safety (Kostopoulou et al., 2008).

Of all the previously identified error categories, adverse events in the form of clinician diagnostic errors were identified as having the potential to result in the most serious consequences for the patient (Kostopoulou & Delaney, 2007). This systematic review looked at the properties of cases that were characterised as being potentially problematic for clinicians in forming a correct final diagnosis. These were found to be:

- **Atypical Presentation:** presenting features of illness are different to classic features normally associated with that illness
- **Non-Specific Presentation:** no presenting features unique in distinguishing the illness resulting in similar presentation to many other differential diagnoses
- **Very Low Prevalence:** uncommon illness in the general population where clinician is

not used to seeing characteristics of such a rare illness on a regular basis
- **Co-Morbidity:** diagnoses that may be accompanied with other complicating illness
- **Perceptual Features:** illness where diagnosis is strongly indicated by visual or auditory cues rather than measureable, testable or symptomatic features

Clinical pattern recognition based on previous clinical case history built up through clinician experience has long been considered as valid and useful in identifying a potential diagnosis based on the presenting characteristics of a patient. Studies have yet to demonstrate a conclusive positive relation relationship between diagnostic accuracy and clinician experience (Kostopoulou et al., 2008). Diagnosis based on previous case history may not be as successful where case presentations occur that are outside the more commonly presented cases for any particular clinical environment. In these situations the potential for a diagnostic adverse event may increase.

We therefore conclude that adverse events related to the primary care clinician of a diagnostic nature may be more likely to take place where more unusual clinical cases present that do not have classic associated features that easily distinguish the illness in question. The challenge therefore is to develop decision support tools that can assist primary care clinicians as part of the diagnostic process to consider diagnoses that might otherwise potentially be missed where they present in these types of problematic cases. In order to identify where best to propose potential solutions to this problem, we need to gain a better understanding of the diagnostic process itself.

Understanding the Diagnostic Process in Primary Care

Having highlighted diagnostic error as a threat to patient safety in primary care, and having identified the characteristics of certain diagnoses that are particularly problematic, it is useful to explore in more detail the underlying cognitive context of diagnostic reasoning as traditionally employed in a clinical consultation. By doing this we can consider where potential interventions may usefully be deployed.

The development by the clinician of an initial set of working diagnostic hypotheses has been identified as a crucial first step at the outset of the patient consultation (Kostopoulou, 2009;Heneghan et al., 2009). The formulation and subsequent investigation of differential diagnoses in response to any given clinical problem has the potential to introduce areas of cognitive bias that may result in errors later in the diagnostic process (Kostopoulou, 2009). The selection of differential diagnoses will drive the investigations done to rule in or rule out each of the available options. Potential issues highlighted by Kostopoulou relating to formulation of diagnostic hypotheses include:

- Not considering a suitable set of differential diagnoses to begin with may result in unnecessary investigations that do not help address the underlying cause of the clinical problem in question and results in lost time reformulating more suitable hypotheses
- The perceived likelihood of the chosen focal hypothesis may vary depending on the number of other differential diagnoses considered at the outset

It is argued that the focal diagnostic hypothesis in a small set of differentials will be perceived to be more likely than the same hypothesis when considered with a larger number of differentials.

We conclude that relying solely on previous experience based on pattern recognition of case histories may not be sufficient for cases presenting with the problematic characteristics discussed previously. In these cases diagnostic decision support that is based on the latest available peer-reviewed clinical evidence that assists in identifying potential differential diagnoses to consider at

the outset may ensure that a correct final diagnosis is not ultimately missed resulting in harm to the patient. The goal is not to attempt to replace clinician judgement but to facilitate consideration of all potential diagnoses at the outset that might otherwise be missed in certain difficult cases. A possible solution is the development of decision support tools that provide for the formulation of appropriate differential diagnoses informed by effective practice of evidence-based medicine.

Evidence-Based Medicine

The most widely cited definition of evidence-based medicine was proposed by Sackett and describes it as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients" (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996b, p. 71). Further development of this definition has focussed on being more explicit about the use of mathematical based models to quantify benefit and harm when making a clinical judgement: "evidence-based medicine is the use of mathematical estimates of the risk of benefit and harm, derived from high-quality research on population samples, to inform clinical decision-making in the diagnosis, investigation or management of individual patients" (Greenhalgh, 2010, p.1). It is clear from these definitions therefore that developing tools to assist clinicians with the practice of evidence-based medicine will also involve the development of underlying data models for representation of that clinical evidence.

Translational Research: Integrating Research with Clinical Evidence

Whilst the concept of practicing evidence-based medicine tends to focus around providing clinicians with tools and formal processes that facilitate access to the latest available evidence on which to base their decisions, there is another critical element that is problematic. Where will the latest evidence-based data upon which clinical decisions will be made come from and how do we ensure that it is up to date? Currently evidence is generated from clinical research trials. Studies have demonstrated that as clinical guidelines are refined through research efforts, there is a considerable time delay between when that new research is made available and when frontline clinicians begin to actually apply it in practice (Haines & Jones, 1994; Bero et al., 1998; Grol & Grimshaw, 2003).

The currently accepted gold standard for generation of evidence is the randomised control trial (RCT). There may be a considerable period of time when practicing clinicians are actually out of step with the latest available research as generated through RCTs relating to a particular topic. The study of the pathways and mechanisms that enable clinical research knowledge to be translated into actual clinical practice is known as 'translational research'. The term 'translational research' can be defined as follows:

Research that seeks to characterize the sequence of events through which a scientific discovery moves between basic scientists, clinical researchers, practitioners, and consumers, and to find more effective ways to facilitate this process. (Mold et al., 2008, p. 571)

We conclude that any proposed solution that wishes to provide for effective practice of evidence based medicine should be strongly linked with translational research by investigating more efficient mechanisms to quickly translate research findings into actual clinical knowledge available to front line clinical practitioners. The tools being developed as part of the TRANSFoRm project specifically aim to address these issues through utilisation of electronic models of both research and clinical evidence data to support more efficient ways to disseminate this knowledge between research and frontline clinical environments.

The TRANSFoRm Project: Delivering the Benefits of Evidence Based Medicine through Translational Research

The underlying concept of TRANSFoRm is to develop a 'rapid learning healthcare system' driven by advanced computational infrastructure that can improve both patient safety in primary care consultation and the conduct and volume of clinical research in Europe.

The major aim of TRANSFoRm, as shown in Figure 1, is to bridge the gap that exists between research and primary care practice. TRANSFoRm will achieve this by:

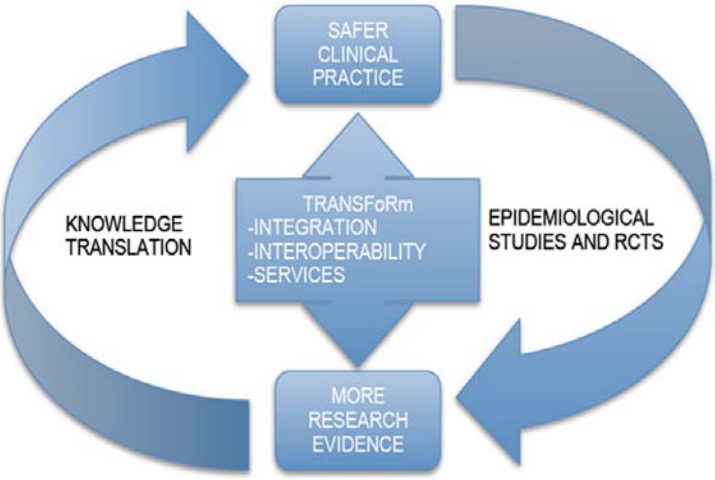
- Allowing for the derivation and update of evidence based guidelines in the form of CPRs through identification of epidemiological patterns from the latest electronic research trial data created from repositories of electronic primary care patient data that are currently exist throughout Europe
- Delivering a remote clinical evidence service that is based on an ontology driven model that utilises CPRs as the primary knowledge representation format
- Providing up to date dissemination of these guidelines through diagnostic decision support tools that utilise the clinical evidence service and are integrated with the electronic health record, assisting primary care practitioners in formulating and considering appropriate differential diagnoses as part of the clinical consultation.

MODELS OF CLINICAL KNOWLEDGE

The Need for Formal Models of Clinical Knowledge

We have stated previously that TRANSFoRm will deliver an updateable clinical evidence service to be utilised in a diagnostic decision support role by primary care practitioners as part of the consultation process. This aims to reduce the potential for adverse events occurring during formulation of potential differential diagnoses as part of the diagnostic process we previously described. In order to implement this it is necessary that we develop formal models of our understanding of clinical evidence upon which the formulation and

Figure 1. Vision for TRANSFoRm



investigation of potential differential diagnoses can be informed. These models of evidence can be considered to be an agreed 'contract' with other electronic systems to allow for sharing and utilisation of clinical evidence data using unambiguous representations of the underlying data. Using this model we can then interpret information to form knowledge that is framed within the known limitations of our developed model. A model is in essence an abstraction of some real world process with known limitations and estimations of that process that we deem are acceptable for our stated purposes (Coiera, 2003).

Our goals can therefore be stated as follows:

- To develop a formal model that will allow the capture of clinical evidence to support the construction of a diagnostic decision support system that models aspects of the real world diagnostic reasoning process
- To develop a formal model that specifically supports the application and use of clinical prediction rules as a mechanism to represent clinical knowledge to provide diagnostic decision support

These goals specifically state that our model will support the use and application of clinical prediction rules as part of the wider model. A clinical prediction rule can be considered a specific model of clinical knowledge and needs further explanation.

What is a Clinical Prediction Rule?

A CPR is a user-friendly clinical guideline derived from statistical analysis of epidemiological patterns of patient data that quantifies the contribution of clinical variables to a target clinical outcome. The following definitions of a clinical prediction rule capture the key points from a diagnostic evidence point of view:

- "Clinical prediction rules (CPRs) are tools that quantify the contributions of symptoms, clinical signs and available diagnostic tests, and in doing so stratify patients according to the probability of having a target disorder." (McGinn et al., 2000, p.79)
- "Clinical prediction rules use clinical findings (history, physical examination, and test results) to make a diagnosis or predict an outcome. They quantify the relative importance of particular findings when evaluating an individual patient." (Reilly & Evans, 2006, p. 201)

Based on these definitions we can explicitly identify the key individual parts of CPR that need to be modelled to use them as part of a broader clinical evidence service. These are:

- A collection of diagnostic cues indicative of the target disorder focussing on:
- Associated diagnostic symptoms as reported by the patient
- Associated diagnostic signs as measured or observed by the clinician
- Associated diagnostic tests as carried out by the clinician
- A collection of criteria to be applied to each of the individual diagnostic cues that, where found applicable, indicates the presence of the target disorder e.g. temperature > 38° C
- A scoring scheme to be applied that quantifies the contribution of each individual diagnostic cue and its associated criteria, relative to the other diagnostic cues, to the clinical outcome under investigation where that cue is found to hold true in a particular patient case
- A stratification of scores to indicate categories of risk or confidence relating to the presence of the target disorder

- Optionally a decision outcome based on the risk stratification that indicates a suitable clinical course of action to be taken for managing that risk category

Derivation of CPRs is typically done using logistic regression as the underlying statistical model. Correlation coefficients are calculated that quantify the contribution of each individual diagnostic variable on the clinical outcome variable under investigation. The goal of logistic regression modelling, and by extension useful CPRs, is to develop a parsimonious model that contains as few diagnostic variables as possible but which explains as much variance in the target binary outcome diagnostic variable as possible. A CPR can therefore be considered to be a normalised logistic regression model where correlation coefficients have been rounded to nearest whole numbers to give a more easily readable and interpreted score system that can be used as a user friendly clinical guideline.

An Example of a CPR

It is useful to look at a specific example of a CPR for illustrative purposes to understand how each of the constituent parts is used. The Centor score is a CPR for predicting the risk of strep throat in patients presenting with sore throats and demonstrates an example of the key constituent parts that collectively constitute a CPR (Centor, Witherspoon, Dalton, Brody, & Link, 1981). It is summarised in Table 1 and consists of four signs and symptoms: tonsillar exudate, tender cervical anterior adenopathy, history of fever and the absence of cough. Points are assigned for the presence of any of these signs and symptoms, with one point allocated per sign or symptom to give a possible total of 0-4. Those scoring 0-1 are deemed to have a low risk of strep throat and the American College of Physicians guidelines recommend no treatment. Those scoring 2-3 are deemed to have an intermediate risk of strep throat and guidelines recommend further testing with a rapid antigen detection test. Those scoring 4 are deemed to have a high risk of strep throat and guidelines recommend empirical treatment with antibiotics.

Problems with Traditional CPR Development and Deployment

Traditional development of CPRs typically involves carrying out some form of clinical trial based on a population of subjects which are confirmed by a gold standard test to have the underlying clinical problem for which a rule is being developed. Typically epidemiological data is captured from all of the underlying subjects in the trial and that data is collated and statistically analysed to detect the strongest diagnostic indicators that provide a correlation to the clinical problem being investigated. The development of logistic regression models to detect the strongest correlations of diagnostic cues and criteria to be applied is the favoured method for creating CPRs.

There are some specific problems inherent in this traditional manual approach:

Table 1. The Centor score CPR

Diagnostic Cues and Criteria	Any of the following signs and symptoms: • Tonsillar exudates present • Tender cervical anterior adenopathy • Fever (temperature >38.0°C) • Absence of cough
Scoring Scheme	1 point is scored for each of the 4 criteria that are found to be true in the patient
Scoring Stratification	0-1 = Low risk 2-3= Intermediate risk 4 = High risk
Decision Outcome - Treatment	Low = No treatment Intermediate = Further testing rapid antigen detection test (RADT) High = Empirical treatment with antibiotics

- Large population samples are required to build statistically valid logistic regression models
- Creation of sufficiently sized manual based clinical trials is costly and time consuming
- Identification and recruitment of suitable candidates to take part in such trials is difficult
- Lack of formal electronic models of CPRs means dissemination is typically done through paper based or e-journal articles rather than electronically
- There are no established standard electronic models of CPRs that encourage easy integration of such rules as part of CDSS systems that can then be used to carry out validation of rules in larger patient populations

The models discussed here will provide a clinical evidence service that allows for dissemination and generation of CPRs from electronic sources of research patient data that can directly address these problems found in traditional CPR development by allowing utilisation of large patient data sets from electronic sources of data to drive CPR derivation.

SOLUTIONS AND RECOMMENDATIONS

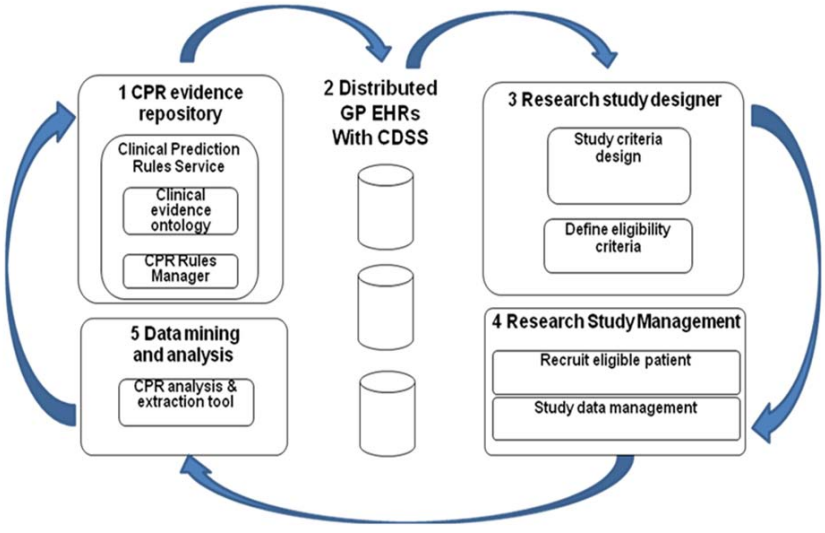
TRANSFoRm: Architecture of the Broader Solution

In order to deliver the overall TRANSFoRm vision that was presented in Figure 1, a number of key components of a technical architecture solution are being developed by other TRANSFoRm collaborators as shown in Figure 2. These components define an iterative translational approach to knowledge generation.

The architectural components identified as being required to support an overall solution are:

1. A repository of diagnostic clinical evidence in the form of derived electronic clinical prediction rules and associated rule management tools, to support GP diagnostic decision making that is provided as a web based clinical evidence service accessible from remote GP electronic health record systems
2. (i) a clinical decision support interface that extracts patient diagnostic cues from a GP

Figure 2. TRANSFoRm architectural components



electronic health record and calls the clinical evidence service to compare the rules contained in the clinical evidence repository with the patient data and returns suggested ranked differential diagnoses that should be considered based on statistical likelihood (ii) the identification and use of suitable third party electronic sources of primary care data for research in the form of consolidated summarised repositories of anonymous primary care data (examples of such as the General Practice Research Database in the UK (http://www.gprd.com) and Nivel in the Netherlands (http://www.nivel.eu).

3. Research study designer tools that allow the creation of electronic clinical research studies that define the research question to be answered along with clinical eligibility criteria required to participate in the study
4. Research study management tools that allow identification and recruitment of eligible patients electronically from third party primary care data sources into a defined research study along with electronic case report forms to specify what clinical information is to be recorded as part of the operation of that study
5. Data mining and analysis tools to provide for analysis of gathered research study data using statistical techniques such as logistic regression modelling to identify key diagnostic indicators that can be used as the basis for the creation or modification of electronic clinical prediction rules contained in a clinical evidence repository.

The development of clinical evidence models is crucial to implementing computable representations of clinical knowledge that will allow for the development of the clinical evidence repository (1) and allow automatic update through data mining and analysis techniques (5). The other component parts will not be described in detail in this work but are provided to set the context for the development and use of the clinical evidence models described and the solution as part of which they will ultimately be deployed. The focus of the subsequent sections will be on describing the specific development work done to date by the authors as part of this overall solution to create formal models of clinical evidence used to develop a clinical evidence service and repository of evidence to support the decision support elements of TRANSFoRm.

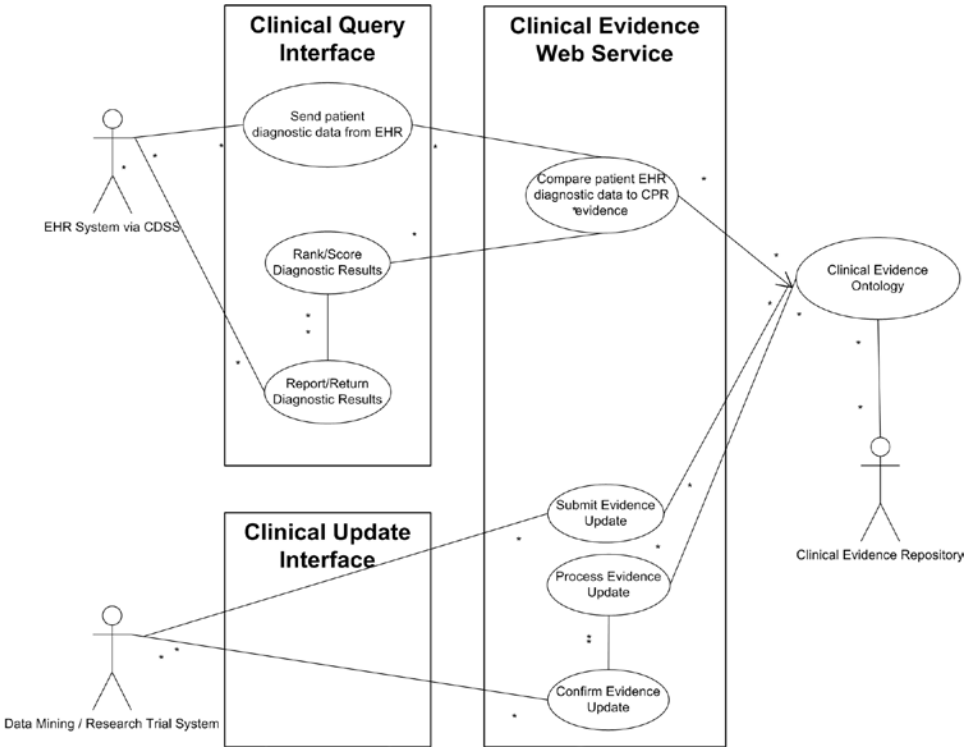
The Clinical Evidence Web Service

The models described later are deployed as part of a complete clinical evidence web service. It will be utilised to provide real-time diagnostic advice to primary care practitioners via remote communication over the internet from a desktop computer. This will be done by extracting patient cue data from their EHR system and passing this to the clinical evidence web service for comparison against the clinical evidence repository of CPRs. This will provide an interactive decision engine to aid the decision making during diagnosis at the 'workbench' of the clinician. Suggested differential diagnoses to consider are dynamically generated from a continuously updated clinical evidence store made available and served to primary care practitioners through a decision support interface linked to their normal working system or environment. The clinical evidence service will support two primary users in the form of primary care practitioners utilising decision support via the evidence service, and clinical researchers who maintain and update the evidence derived from patient data as shown in Figure 3.

The Clinical Query Interface

The web service will provide defined web service methods that will be used to initiate activation of CPRs by making use of queries to the clinical evidence ontology to identify suitable rules to apply. The service will allow for application of

Figure 3. TRANSFoRm web service use case



a diagnostic comparison algorithm which will be applied to the set of diagnostic cues gathered from the input patient record. Each patient record will be associated with evidence consisting of the signs, symptoms, clinical tests and clinical history to be provided to the evidence service. The application of the algorithm can be applied at a number of different levels of evidence granularity as required, namely a comparison of a patients diagnostic cues against:

- All defined scenarios in the entire evidence database
- Evidence relating to a specific diagnostic scenario and all the associated diagnoses e.g. 'abdominal pain'
- Evidence relating a specific diagnosis e.g. 'Urinary Tract Infection'
- Evidence relating to a specific diagnostic cue (or collection of cues) e.g. 'fever with temperature > 38° C'

Comparison against the full evidence database, a clinical scenario or a specific diagnosis is a comparison that will be done in a 'top-down' fashion by activation of CPRs associated within each associated diagnosis. Comparison against a specific diagnostic cue will be done in a 'bottom-up' fashion by comparing the evidence against a specific diagnostic cue and working upwards to find diagnoses and activating CPRs associated with that diagnostic cue.

The primary output will be an evidence report returned for formatting and display in the users' clinical evidence client. This will consist of:

- A list of potential differential diagnoses to consider for investigation
- Key diagnostic cues to consider for each diagnosis
- A list of CPRs triggered with each potential diagnosis along with the score and stratification levels associated with it

- A ranked list, ordered by probability, of potential differential diagnoses to consider based on the application of the CPRs to the particular patient diagnostic cues under investigation

Clinical Evidence Update Interface

In order to support timely dissemination of new or updated CPRs, the web service will provide methods to allow the TRANSFoRm data mining components to insert and update the results of logistic regression models that have been derived from analysis of electronic research trial data. A clinical prediction rule manager tool will allow for the creation or update of CPRs in the clinical evidence repository based on the underlying derived logistic regression models provided by the research data mining components.

Clinical Evidence Model Development Methodology

The provision of the previously described evidence service will be based on a computable model of evidence that underpins the content provided by the service. The following methodological steps were used to drive the development of the clinical evidence models and ontology for TRANSFoRm:

- A number of diagnostic clinical scenarios were chosen by clinicians to provide the basis of case studies for the development of the decision support elements of TRANSFoRm. These clinical scenarios were chosen to demonstrate the diagnostic characteristics found through the research on diagnostic error that was described previously.
- A literature review of evidence supporting the chosen diagnostic scenarios was conducted on Pubmed, Embase and established sources of evidence based medicine (Cochrane reviews, SIGN guidelines, NHS clinical knowledge summaries). These reviews documented the key diagnostic cues, criteria, diagnostic algorithms and any available clinical prediction rules available in clinical literature to support diagnosis for the chosen clinical case studies.
- The structure of the evidence documented in these reviews was used to develop unified modelling language (UML) (http://www.uml.org) models that define the clinical evidence structure with a view to being generic enough to represent any other selected condition above and beyond those chosen for TRANSFoRm.
- The key concepts identified in the UML models were then used for development of an ontology of evidence to support our chosen diagnostic scenarios by adding more detail to describe the semantic characteristics of the relationships that exist between the core clinical evidence concepts with a view to support clinical evidence queries.
- Validation and evaluation of ontology structure was carried out using a definition of 'clinical evidence competency questions' which establish what questions we need our ontology to be able to answer with a view to providing a diagnostic clinical evidence service.

A Presenting Clinical Scenario: Female Abdominal Pain

In order to develop our models we have considered a number of clinical scenarios or use cases falling under three broad categories: 'female abdominal pain', 'male chest pain' and 'dsypnoea'. For the purposes of this discussion we will focus on 'female abdominal pain' and within that on the diagnosis of 'urinary tract infection'. A large number of potential diagnoses may need to be considered by a clinician when considering 'female abdominal pain'. For our scenario there are a number of differential diagnoses that might

potentially be considered and only a particular subset have been specifically chosen for our purposes that collectively demonstrate the diagnostic characteristics that were previously discussed as being potentially problematic for clinicians (Kostopoulou et al., 2008). Examples we will consider are: ectopic pregnancy, pyelonephritis, urinary tract infection, chrons disease, appendicitis, bowel cancer, irritable bowel syndrome and bacterial enteritis. We will look at a specific CPR related to diagnosing urinary tract infection which we have called 'The Little symptom rule' (Little et al., 2006).

Based on our previous description of the diagnostic process, the challenge is to model a diagnostic process that is capable of suggesting potential differential diagnoses to consider for a particular clinical problem which then uses evidence based CPRs that are triggered to rule in or rule out the most likely ones to consider. This is done based on the specific patient case data that is presented as part of the clinical consultation with the primary care practitioner as extracted from a patient EHR and passed to our clinical evidence service.

The starting point for model development is to consider what sort of diagnostic questions we might wish our decision support tool to be able to answer relating to this clinical scenario. The following is a subset of all basic requirements for this clinical case:

- For a patient presenting with abdominal pain provide a list of differential diagnoses that should be considered for investigation for that patient
- For urinary tract infection present a list of the key clinical diagnostic cues that should be considered as clinical evidence to support the inclusion or exclusion of that diagnosis
- Provide a list of the clinical prediction rules that may assist with diagnosing urinary tract infection
- Describe all the constituent parts of any available CPR applicable for diagnosing urinary tract infection
- Describe the population context and characteristics from which the clinical criteria to be applied to diagnostic cues for a particular CPRs were derived

Description of Underlying Clinical Evidence Data Model to Support Clinical Diagnosis

Based on these initial diagnostic questions we can identify three clinical areas that will need to be modelled:

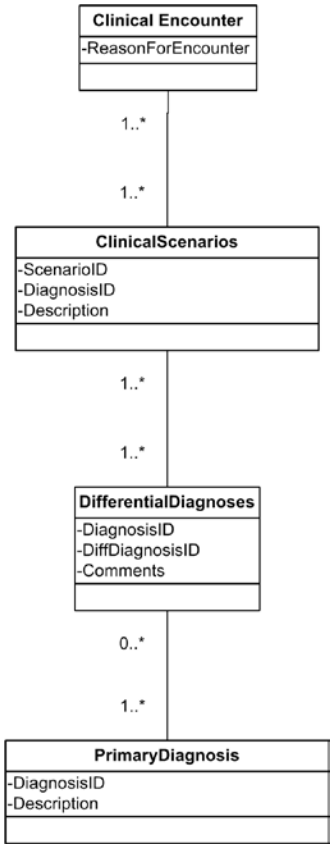
- A diagnostic trigger process that initiates the selection of appropriate differential diagnoses for investigation
- A model of clinical prediction rules and associated underlying clinical evidence diagnostic cues
- A model of clinical context identifying the epidemiological context of the population from which the CPR was derived

Based on our previous analysis of the diagnostic process and the underlying problems that are present as traditionally practiced, we have developed a model of clinical evidence addressing these three areas that can support provision of clinical decision support through use of clinical prediction rules derived from electronic research data. These models are described using UML diagrams which identify the key concepts and the relationships that exist between them. The diagrams shown subsequently identify the key concepts and some important characteristics describing those concepts along with relationships that exist between concepts (only a subset of concept characteristics is shown for clarity).

The Diagnostic Trigger Process

The entry point for the overall model is the clinical encounter. This results in the triggering of the diagnostic process. A patient will present to a primary care practitioner with one or perhaps multiple clinical complaints. Each of these complaints can be captured by what is referred to as a 'reason for encounter' (RFE) that will be recorded as a coded value in the EHR system and trigger the associated diagnostic decision support process. The RFE will be a textual description of the nature of the complaint associated with a selected code from a clinical coding scheme (such as the International Classification of Primary Care 2nd Edition, ICPC-2 for example (http://www.who.int/classifications/icd/adaptations/icpc2/en/index.html)). The model is described in Figure 4 and Table 2.

Figure 4. Diagnostic trigger model



In our example a female patient may present with abdominal pain which might be recorded as the ICPC-2 codes D01 'Abdominal pain/cramps general' or D02 'Abdominal pain epigastric'. The 'female abdominal pain' scenario may be associated with more than one coded RFE to allow for support of multiple clinical coding schemes. The scenario is then associated with a number of potential differential diagnoses that should be considered for investigation when triggered by an associated RFE. In the case of female abdominal pain we have selected to consider the differential diagnoses 'urinary tract infection', 'ectopic pregnancy', 'Chrons disease', 'bacterial enteritis', 'appendicitis', 'irritable bowel syndrome', 'ovarian cancer'.

Each differential diagnosis is then treated in turn as the primary diagnosis for investigation (the order of investigation can be set as a recorded investigation priority by the evidence administrator). Each primary diagnosis will have a collection of diagnostic evidence supporting the inclusion or exclusion of that diagnosis.

CPR Interpretation Model

The second model described in Figure 5 and Table 3 represents the structure of clinical prediction rules and links possibly many CPRs with each diagnosis. It describes the main CPR constructs that are associated with a clinical prediction rule along with the diagnostic evidence and statistical model that underpin it.

The Clinical Context Model

The third model described in Figure 6 and Table 4 provides additional epidemiological context to support appropriate identification of populations supported by a selected CPR. It provides for querying of CPRs based on appropriate epidemiological population sets for the patient case in question.

Table 2. Diagnostic trigger model description

Item No	Item Name	Description
1	ClinicalEncounter	The trigger object for the entire diagnostic process. This will be based on a coded "reason for encounter" as presented by a patient at consultation. This could be a clinical complaint or a request by the patient for a clinical intervention. Each coded RFE will be associated with a clinical scenario for investigation.
4	ClinicalScenarios	A broad category used to associate clinical diagnoses with a coded clinical encounter as part of a scenario investigation e.g. 'female abdominal pain', 'male chest pain', 'dsypnoea'.
2	DifferentialDiagnoses	A list of potential differential diagnoses that will be considered for investigation as part of the investigation of any particular clinical scenario
3	PrimaryDiagnosis	The current primary diagnosis under consideration for investigation

Table 3. CPR interpretation model description

Item Name	Description
ClinicalPredictionRule	A clinical prediction rule that has references to the core structural elements that define a clinical prediction rule
CPRScoringScheme	A scoring scheme used to evaluate the evidence criteria that are applicable for a particular patient case in the context of application of a clinical prediction rule
CPRDecisionOutcomes	A set of decision outcomes or conclusions that are used to interpret the scoring scheme being applied.
CPRPublicationEvidence	A collection of supporting information describing research publications upon which the clinical prediction rule was derived and/or applied in practice
ResearchEvidenceTrials	A link to a supporting electronic TRANSFORM research trial reference which has been used to derive information used to derive or validate the clinical prediction rule
EvidenceCueSet	A named collection that references the diagnostic cues indicative of a clinical prediction rule, which are independent of patient data. These are signs, symptoms, clinical tests and risk factors indicative of a primary diagnosis which the CPR is related to.
SignCollection	A named collection of clinical signs observed by the clinician that are indicative of a primary diagnosis
SymptomCollection	A named collection of clinical symptoms reported by the patient and are indicative of a primary diagnosis
TestCollection	A named collection of clinical tests that are indicative of a primary diagnosis
RiskFactors	A named collection of clinical risk factors that are indicative of a primary diagnosis
Sign	A single individual clinical sign that is indicative of a primary diagnosis
Symptom	A single individual clinical symptom that is indicative of a primary diagnosis
Test	A single individual clinical test that is indicative of a primary diagnosis
Risk	A single individual clinical risk factor that is indicative of a primary diagnosis
DiagnosticCue	A generic diagnostic cue object that signs, symptoms, tests and risks are derived from
EvdienceCriteriaCollection	A named collection of diagnostic criteria that are to be applied to the diagnostic cues relating to a primary diagnosis that is indicative of that primary diagnosis. E.g. a diagnostic cue and criteria combination could be "fever" and ">38C"
EvidenceContraint	An individual constraint that can be applied to a diagnostic cue. This could be a test for a particular threshold or range of values using a particular operator and measurement to test.
LogisticRegressionModel	A logistic regression model from which the CPR has been derived which is inserted or updated from the data mining and analysis component
RegressionParameters	The individual diagnostic variables contained in a logistic regression model with calculated correlation coefficients

Figure 5. CPR interpretation model

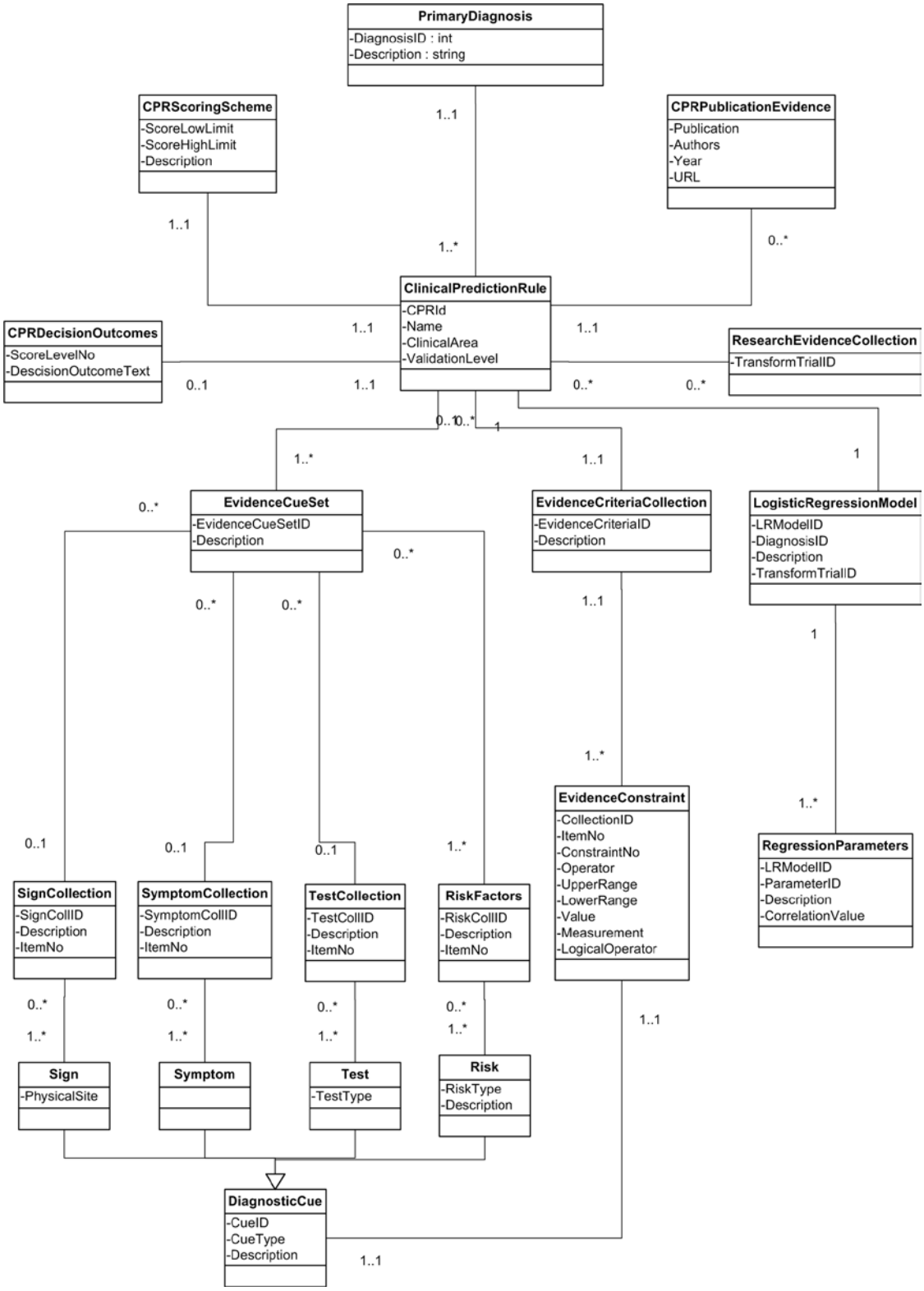
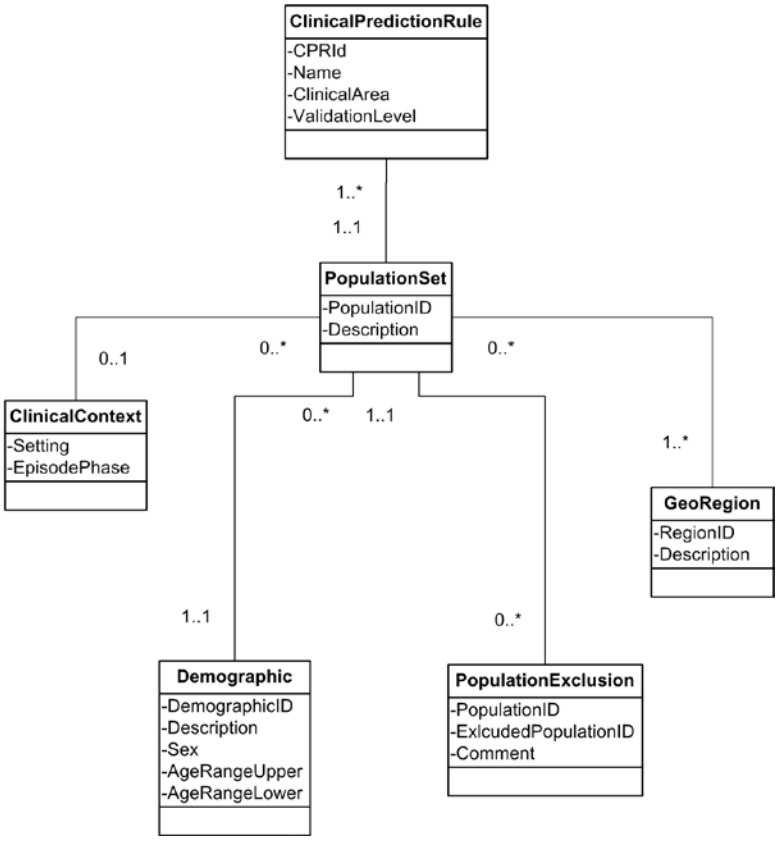


Figure 6. Clinical context model



Definition and Role of an Ontology

The UML models as described are a starting point in modelling what the key concepts are that we wish to capture. UML models present a view of concepts that is traditionally focussed on application development and implementation using object oriented programming languages such as java, but they lack the ability to capture the semantic meaning of the relationships that exist between those concepts. In addition UML models provide no mechanism for actually querying our model to ask specific questions of it. We therefore need to take the key concepts we have previously identified in our UML models and add an additional semantic layer of meaning to the relationships that exist between them which allows us to develop formal queries of that information that answer the types of questions we defined previously in our clinical scenario definition.

An ontology provides such a formal definition of a model that describes the core concepts and semantic relations that exist between those concepts as necessary to describe a selected knowledge domain of which questions will be asked. The ontology describes a formal contract for meaning that can be used to formulate queries of data from the selected knowledge. The application of what are termed 'semantic web' technologies in the form of ontologies has shown itself to be highly successful in achieving this in real world applications in areas as diverse as biomedicine, social networking and online retailing by providing for a common definition of the core concepts contained in a knowledge domain that can easily be shared and used by other applications outside

Table 4. Clinical context model description

Item Name	Description
PopulationSet	The CPR may be applicable only in the context of a particular population set or context. This is a collection object that references a number of aspects of the context such as the clinical context, the demographic that it applies to or the geographical region that is relevant.
ClinicalContext	A description of the clinical context within which the constraint weighting is applicable. This will indicate items such as the clinical setting (primary, secondary) and the episode of care that it relates to.
GeoRegion	A description of the geographical region which the constraint weighting is applicable to. E.g. "UK", "France", "England", "Northern Europe"
PopulationExclusion	A reference to population sets that are specifically excluded as being not applicable for the evidence weighting for a constraint.
Demographic	A demographic description of the population subset to which the evidence weighting is applicable e.g. Caucasian males aged over 40 years of age.

of the original development application (Allemang & Hendler, 2008). The use of an ontology provides for a clear abstraction of the definition of knowledge from the underlying operational applications that will make use of them. As an example, the clinical evidence ontology being developed is distinct from the actual diagnostic matching algorithms that will make use of the concepts in that ontology of clinical evidence. Because of this abstraction and the fact that the technologies used to describe ontologies are based on recognised standards (defined by the world wide web consortium W3C), the ontologies that are developed as part of TRANSFoRm could then be re-used by other applications outside of TRANSFoRm by providing mappings to their own repositories of application data using a single consistent interpretation of that clinical knowledge. This common definition of clinical knowledge can further support interoperability in the form of knowledge sharing and combination of diverse data sources.

A primary reason that ontologies have been successful in real world application is the simplicity of the underlying data representation used. The primary data representation mechanism underpinning ontology development is a standard known as 'resource description framework' (RDF). The focus in RDF is to reduce data representation to an extremely simple representation using statements in the form of 'triples'. These define statements in the form of three component parts: subject-predicate-object. Looking at an example of triples in the clinical evidence domain, we may define generic concepts such as 'Diagnosis' and 'DiagnosticCue'. We may define two inversely related relationships between these two concepts called 'hasCue' and 'isCueOf'. Within each generic concept we can define concept instances that are specific cases of those generic concepts that we wish to define for our application. We can then establish directional relationships between those instances that define triple statements of a form such as:

- 'Dysuria' - 'isCueOf'
  - 'UrinaryTractInfection'
- 'UrinaryTractInfection' - 'hasCue'
  - "Dysuria'

By building a complete set of these triple statements we define both the generic concepts and relationships that exist and also the specific instances of those concepts that we wish to work with for our particular application. Collectively these define what we know about the knowledge domain we are modelling.

At first glance this definition seems quite similar to object-oriented modelling techniques such as UML that we saw earlier. The focus on

ontology definition however is more concerned with describing the semantic characteristics of the relationships that exist between the defined ontology concepts in a more complete way than UML does. In UML relationships between concepts are bi-directional and generally restricted to defining cardinality. In an ontology, relationships are defined in a single direction along with a definition of characteristics that describe the specific semantic nature of those relationships and how they can be applied to the generic concepts defined in the ontology. This allows data to be queried in a directional manner such as starting from a diagnosis and working down to diagnostic cues, or starting with a diagnostic cue and working up to a diagnosis. The ability to achieve this semantic expression is achieved through use of additional meta-data technologies, such as 'RDFS' (RDF-schema language) and 'web ontology language' (OWL), that are layered upon the basic data representations of RDF specifically with a view to defining generic formal models that fully describe a particular knowledge domain in the form of ontologies.

Clinical Evidence Ontology Derivation and Validation Based on Competency Questions

The use of what are known as 'competency questions' has been proposed as a useful starting point for the design, implementation and validation of ontologies (Gruninger et al., 1995). This approach advocates the following:

- Draw up a list of informal questions that your proposed ontology should be able to provide answers to
- Deconstruct those questions by identifying more formal concepts in the form of knowledge domain objects, relationships and instances that are required to answer them
- Construct the ontology based on the formal concepts you have identified
- Express the informal competency questions using formal ontology query languages and your constructed ontology
- Validate your ontology through clinical use cases to demonstrate that the results generated from formal ontology questions correctly answer your clinical use cases

The first step in this process has already been completed through our previous definition of a presenting clinical scenario and definition of questions to be answered related to it. The second step has been accomplished through development of our UML models as previously described. The remaining tasks are to create the ontology in terms of generic concepts and specific data instances of them and add named directional relationships which allow us to formulate queries that answer the diagnostic questions we require of our ontology, and validate it.

Construct the Ontology

We have developed a working clinical evidence ontology using a graphical tool called Protégé (http://protege.stanford.edu) which simplifies the process of building ontology structures and generates the underlying models in the form of formatted text files. The text file format produced by Protégé is not sufficient on its own to support a multi-user web service as required by TRANSFoRm. This requires more application friendly database structures known as 'triple stores' that provide functionality for persistent storage and querying of an ontology. We have constructed our triple store using a product called Sesame (http://www.openrdf.org) that is loaded with the ontology developed using Protégé.

Protégé allows for definition of the generic model of evidence by constructing the key generic concepts in the ontology along with named directional relationships that are allowed to exist

between them as shown in Figure 7. Directional relationships in one direction are only shown for clarity but equivalent inverse relationships are also defined e.g. 'hasCPR' has an inverse relationship 'isCPROf'.

Instances of these generic concepts which represent specific occurrences of the generic concepts can be defined that represent the actual data we wish to capture for our application. Table 5 shows a narrative description and Figure 8 shows an ontological representation of the 'Little symptom rule'.

Table 5. The little symptom rule

Diagnostic Cues and Criteria	• Urine cloudiness present • Urine smell present • Dysuria present • Nocturia present
Scoring Scheme	1 point is scored for each of the 4 criteria that are found to be true in the patient
Scoring Stratification	0-2 = Low risk 3-4= High risk
Decision Outcome - Treatment	Low = No treatment High = Empirical treatment with antibiotics

Validating the Ontology Structure through Clinical Competency Questions

In addition to our model definition we require a mechanism for querying that model to answer required questions. The core language used for querying ontologies is called 'Simple Protocol and RDF Query Language (SPARQL) (http://www.w3.org/TR/rdf-sparql-query). It is similar to relational database query languages such as structured query language (SQL) but focuses on use of combinations of triple patterns to form the query. These triple patterns can either be generic concepts or specific concept instances in the ontology or wildcard items (indicated with a ? prefix). A number of triple patterns can be concatenated together to allow for very complex querying in a very compact format.

Figure 7. Generic ontology model

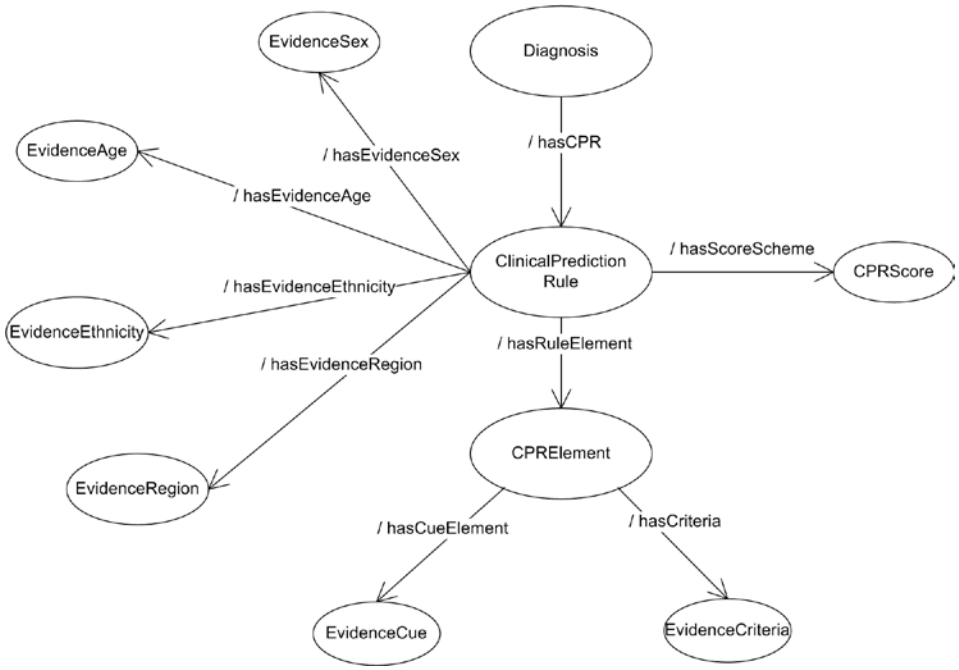
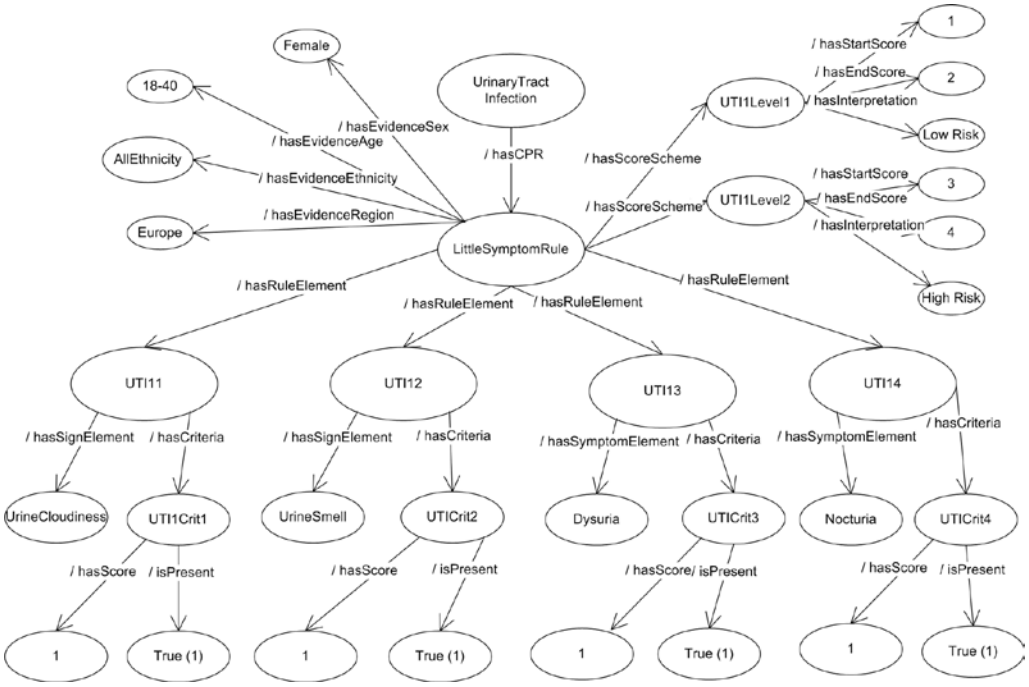


Figure 8. A specific ontology instance



If our ontology is conceptually complete and accurately defined with respect to our clinical competency questions, then it should be possible to easily translate those informal questions that we previously introduced during clinical scenario development (Table 6) into the equivalent formal questions (Table 7). The results generated from those questions should be consistent with the expected data we have expressed as instances of our general ontology concepts (in our case the clinical evidence populated from our clinical scenario to support diagnosis of female abdominal pain). Table 7 details our formal clinical evidence competency questions and associated results from the SPARQL commands that were successfully run against the Protégé ontology that we hosted and queried using the Sesame infrastructure.

Evaluating the Ontology

The work being carried out on TRANSFoRm is currently a work in progress at this time of writing and a substantial amount of work is yet to be completed. To date the initial focus has been on

Table 6. Informal competency questions

Q1	For a patient presenting with abdominal pain provide a list of differential diagnoses that should be considered for investigation for that patient
Q2	For urinary tract infection present a list of the key clinical diagnostic cues that should be considered as clinical evidence to support the inclusion or exclusion of that diagnosis
Q3	Provide a list of the clinical prediction rules that may assist with diagnosing urinary tract infection
Q4	Describe all the constituent parts of any available CPR applicable for diagnosing urinary tract infection
Q5	Describe the population context and characteristics from which the clinical criteria to be applied to diagnostic cues for a particular CPRs were derived

Table 7. Formal competency questions as SPARQL queries

Q1 - SPARQL Query	**Results**
SELECT ?anyDifferentialDiagnosis WHERE {?anyRFE hasICPC2Code “D01”^^xsd:string . ?anyDifferentialDiagnosis isDifferentialDiagnosisOf ?anyRFE .}	anyDifferentialDiagnosis EctopicPregnancy Pyelonephritis UrinaryTractInfection ChronsDisease Appendicitis BowelCancer IrritableBowelSyndrome BacterialEnteritis
Q2 - SPARQL Query	**Results**
SELECT ?anyEvidenceCue WHERE {?anyEvidenceCue isCueOf UrinaryTractInfection.}	anyEvidenceCue Fever UrineCloudiness Dysuria NewSexualPartner Nocturia DipstickUrinalysis PregnancyTest PreviousUTI SexualActivity UrineSmell
Q3 - SPARQL Query	**Results**
SELECT ?anyCPR WHERE {?anyCPR isCprOf UrinaryTractInfection .}	anyCPR LittleSymptomRule LittleDipstickRule
Q4 - SPARQL Query	**Results**
SELECT ?anyCriteriaElement ?anyCueElement ?anyProperty ?anyValue WHERE {?anyRuleElement isRuleElementOf LittleSymptomRule . ?anyCriteriaElement isCriteriaOf ?anyRuleElement. ?anyCueElement isCueElementOf ?anyRuleElement. ?anyCriteriaElement ?anyProperty ?anyValue. ?anyProperty rdf:type owl:DatatypeProperty. }	anyCriteriaElement anyCueElement anyProperty anyValue UTI1Crit1 UrineCloudiness isPresent 1 UTI1Crit2 UrineSmell isPresent 1 UTI1Crit3 Dysuria isPresent 1 UTI1Crit4 Nocturia isPresent 1
Q5 -SPARQL Query	**Results**
SELECT ?anyEvidenceContext WHERE {?anyEvidenceContext isEvidenceContextOf LittleSymptomRule.}	anyEvidenceContext 18-40 AllEthnicity Europe Female PrimaryCare

identifying the component parts of what the end solution will ultimately look like and on development of models of clinical evidence that will provide for implementation of that overall solution. As such it is too early to accurately evaluate how successful the end products of this research will be. We can however evaluate the development of the ontological models by populating them with clinical data that represents real clinical scenarios and assess if those models are descriptive enough to allow us describe that data in a format that allows us to answer competency questions as defined by our decision support requirements. This has been successfully completed using a Sesame triple store implementation containing our developed ontology which we have programmatically queried to answer our defined diagnostic questions. Using this approach we can say that the models described in this document are sufficiently descriptive to allow for representation of the clinical evidence knowledge domain specifically with a view to supporting representations of clinical prediction rules from electronic sources of clinical data. The Sesame triple store we have created and the results generated from our ontology queries are consistent with what we expected based on the data populated in the ontology as part of the formulation of our clinical use cases. Whilst the models developed in this work have been presented in the context of our specific clinical cases, the ontological representation captures the general clinical concepts and relationships required to being customisable to support applicability to other clinical cases above and beyond the clinical scenarios detailed here.

Discussion

The use of semantic web technologies to define consistent ontologies as a means for integrating and reasoning on diverse sources of data has proved itself to be extremely successful on commercial enterprises found on the world wide web (Allemang & Hendler, 2008). These technologies offer a very different approach to addressing some interoperability issues than have been found in traditional health informatics that utilise complex data models. The power of these semantic web approaches is based on taking a diametrically opposed approach which is to reduce the underlying data representation to its simplest form through use of simple triples of data using RDF. The required formal definitions of knowledge domain structures and models are then provided as additional meta-data layers built on top of this data representation through use of ontology definition tools that implement technologies such as OWL and RDFS. It is this underling simplicity of data representation in combination with formal shared definitions of knowledge structures that makes the use of ontologies potentially so powerful and can be leveraged to address some of the traditional interoperability challenges faced in development of models to support health systems and processes.

CONCLUSION

The effective practice of evidence based medicine will never be made reality unless better ways can be found to disseminate clinical findings generated from research environments as clinical guidelines practised in front line clinical care, by bridging knowledge gaps between the two environments. As primary care practitioners are a first point of contact for many patients that enter our health systems it is essential that we research ways of helping them make clinical decisions that are informed by the latest available clinical evidence. The use of IT infrastructures to bridge that gap, as proposed by the TRANSFoRm project, provides a blueprint for what the 'rapid learning healthcare system' may ultimately look like. In order to achieve this we need flexible models of how we represent clinical facts to support that decision making and design them in such a way that this data can be easily updated and disseminated in a more efficient fashion than is currently provided for with traditional research networks. This is in the

interests of both patients and clinical practitioners by more tightly integrating the research process with frontline care to inform decisions that reduce the possibility of diagnostic adverse events, and thereby increase levels of patient safety.

REFERENCES

Allemang, D., & Hendler, J. A. (2008). *Semantic web for the working ontologist: Modeling in RDF, RDFS and OWL*. Morgan Kaufmann.

Bero, L. A., Grilli, R., Grimshaw, J. M., Harvey, E., Oxman, A. D., & Thomson, M. A. (1998). Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *British Medical Journal*, *317*(7156), 465–468. doi:10.1136/bmj.317.7156.465

Bhasale, A. L., Miller, G. C., Reid, S. E., & Britt, H. C. (1998). Analysing potential harm in Australian general practice: an incident-monitoring study. *The Medical Journal of Australia*, *169*(2), 73–76.

Centor, R. M., Witherspoon, J. M., Dalton, H. P., Brody, C. E., & Link, K. (1981). The diagnosis of strep throat in adults in the emergency room. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *1*(3), 239. doi:10.1177/0272989X8100100304

Coiera, E. (2003). *Guide to health informatics*. London, UK: Hodder Arnold.

Donaldson, M. S. (1996). *Primary care: America's health in a new era*. National Academies Press.

Ebell, M. H., Smith, M. A., Barry, H. C., Ives, K., & Carey, M. (2000). Does this patient have strep throat? *Journal of the American Medical Association*, *284*(22), 2912–2918. doi:10.1001/jama.284.22.2912

Elder, N. C., & Dovey, S. M. (2002). Classification of medical errors and preventable adverse events in primary care: A synthesis of the literature. *The Journal of Family Practice*, *51*(11), 927–932.

Fisseni, G., Pentzek, M., & Abholz, H.-H. (2008). Responding to serious medical error in general practice—Consequences for the GPs involved: Analysis of 75 cases from Germany. *Family Practice*, *25*(1), 9–13. doi:10.1093/fampra/cmm071

Greenhalgh, T. (2010). *How to read a paper: The basics of evidence-based medicine* (*Vol. 8*). BMJ books.

Grol, R., & Grimshaw, J. (2003). From best evidence to best practice: Effective implementation of change in patients' care. *Lancet*, *362*(9391), 1225–1230. doi:10.1016/S0140-6736(03)14546-1

Gruninger, M., Fox, M. S., et al. (1995). Methodology for the design and evaluation of ontologies. *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI* (Vol. 95).

Haines, A., & Jones, R. (1994). Implementing findings of research. *British Medical Journal*, *308*(6942), 1488–1492. doi:10.1136/bmj.308.6942.1488

Heneghan, C., Glasziou, P., Thompson, M., Rose, P., Balla, J., Lasserson, D., Scott, C., et al. (2009). Diagnostic strategies used in primary care. *Medical Journal, 338*.

Knottnerus, J. A., & Buntinx, F. (2002). *The evidence base of clinical diagnosis*. Wiley Online Library.

Kostopoulou, O. (2009). Diagnostic errors: Psychological theories and research implications. In Hurwitz, B., & Sheikh, A. (Eds.), *Health care errors and patient safety* (pp. 95–111). doi:10.1002/9781444308150.ch8

Kostopoulou, O., & Delaney, B. (2007). Confidential reporting of patient safety events in primary care: Results from a multilevel classification of cognitive and system factors. *Quality & Safety in Health Care*, *16*(2), 95. doi:10.1136/qshc.2006.020909

Kostopoulou, O., Delaney, B. C., & Munro, C. W. (2008). Diagnostic difficulty and error in primary care—A systematic review. *Family Practice*, *25*(6), 400. doi:10.1093/fampra/cmn071

Kostopoulou, O., Oudhoff, J., Nath, R., Delaney, B. C., Munro, C. W., Harries, C., & Holder, R. (2008). Predictors of diagnostic accuracy and safe management in difficult diagnostic problems in family medicine. *Medical Decision Making*, *28*(5), 668. doi:10.1177/0272989X08319958

Lenfant, C. (2003). Clinical research to clinical practice—Lost in translation? *The New England Journal of Medicine*, *349*(9), 868–874. doi:10.1056/NEJMsa035507

Little, P., Turner, S., Rumsby, K., Warner, G., Moore, M., & Lowes, J. A. (2006). Developing clinical rules to predict urinary tract infection in primary care settings: sensitIvity and specificity of near patient tests (dipsticks) and clinical scores. *The British Journal of General Practice*, *56*(529), 606–612.

McGinn, T. G., Guyatt, G. H., Wyer, P. C., Naylor, C. D., Stiell, I. G., & Richardson, W. S. (2000). Users' guides to the medical literature. *Journal of the American Medical Association*, *284*(1), 79. doi:10.1001/jama.284.1.79

Mold, J. W., Pasternak, A., McCaulay, A., Manca, D., Rubin, G., & Westfall, J. (2008). Definitions of common terms relevant to primary care research. *Annals of Family Medicine*, *6*(6), 570. doi:10.1370/afm.932

Olesen, F., Dickinson, J., & Hjortdahl, P. (2000). General practice—Time for a new definition. *British Medical Journal*, *320*(7231), 354–357. doi:10.1136/bmj.320.7231.354

Reilly, B. M., & Evans, A. T. (2006). Translating clinical research into clinical practice: Impact of using prediction rules to make decisions. *Annals of Internal Medicine*, *144*(3), 201.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. It's about integrating individual clinical expertise and the best external evidence. *British Medical Journal*, *312*(7023), 71–72. doi:10.1136/bmj.312.7023.71

Smitha, B., & Ceustersc, W. (2006). HL7 RIM: An incoherent standard. *Studies in Health Technology and Informatics*, *124*, 133–138.

ADDITIONAL READING

Ahmadian, L., van Engen-Verheul, M., Bakhshi-Raiez, F., Peek, N., Cornet, R., & de Keizer, N. F. (2011). The role of standardized data and terminological systems in computerized clinical decision support systems: Literature review and survey. *International Journal of Medical Informatics*, *80*(2), 81–93. doi:10.1016/j.ijmedinf.2010.11.006

Choudhry, N. K., Fletcher, R. H., & Soumerai, S. B. (2005). Systematic review: The relationship between clinical experience and quality of health care. *Annals of Internal Medicine*, *142*(4), 260–273.

Digby, A. (1999). *The evolution of British general practice 1850-1948*. New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780198205135.001.0001

Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., & Laleci, G. (2006). *Electronic health record standards–A brief overview.* 4th International Conference on Information and Communications Technology, Cairo, Egypt.

Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., & Laleci, G. B. (2005). A survey and analysis of electronic healthcare record standards. *ACM Computing Surveys*, *37*(4), 277–315. doi:10.1145/1118890.1118891

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.

Greenhalgh, T., & Russell, J. (2009). Evidence-based policymaking: A critique. *Perspectives in Biology and Medicine*, *52*(2), 304–318. doi:10.1353/pbm.0.0085

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs*, *24*(5), 1103. doi:10.1377/hlthaff.24.5.1103

Jaspers, M. W. M., Smeulers, M., Vermeulen, H., & Peute, L. W. (2011). Effects of clinical decision-support systems on practitioner performance and patient outcomes: A synthesis of high-quality systematic review findings. *Journal of the American Medical Informatics Association*, *18*(3), 327. doi:10.1136/amiajnl-2011-000094

Kalra, D. (2006). Electronic health record standards. *IMIA yearbook of medical informatics, 2006*, 136–144.

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *British Medical Journal*, *330*(7494), 765. doi:10.1136/bmj.38398.500764.8F

Knottnerus, J. A., & Buntinx, F. (Eds.). (2002). *The evidence base of clinical diagnosis*. Wiley Online Library.

Kostopoulou, O., Oudhoff, J., Nath, R., Delaney, B. C., Munro, C. W., Harries, C., & Holder, R. (2008). Predictors of diagnostic accuracy and safe management in difficult diagnostic problems in family medicine. *Medical Decision Making*, *28*(5), 668. doi:10.1177/0272989X08319958

Laupacis, A., & Sekar, N. (1997). Clinical prediction rules. *Journal of the American Medical Association*, *277*(6), 488. doi:10.1001/jama.1997.03540300056034

Lee, T. B., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 34–43. doi:10.1038/scientificamerican0501-34

Leslie, H., Heard, S., Garde, S., & McNicoll, I. (2009). Engaging clinicians in clinical content: Herding cats or piece of cake. *Studies in Health Technology and Informatics*, *150*, 125–129.

Miles, J., & Shevlin, M. (2001). *Applying regression & correlation: A guide for students and researchers*. Sage Publications Ltd.

Mowatt, G., Grimshaw, J. M., Davis, D. A., & Mazmanian, P. E. (2001). Getting evidence into practice: The work of the cochrane effective practice and organization of care group (EPOC). *The Journal of Continuing Education in the Health Professions*, *21*(1), 55–60. doi:10.1002/chp.1340210109

Norman, G. R., & Brooks, L. R. (1997). The non-analytical basis of clinical reasoning. *Advances in Health Sciences Education : Theory and Practice*, *2*(2), 173–184. doi:10.1023/A:1009784330364

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Citeseer.

Pearce, C., Arnold, M., Phillips, C., Trumble, S., & Dwan, K. (2011). The patient and the computer in the primary care consultation. *Journal of the American Medical Informatics Association*, *18*(2), 138. doi:10.1136/jamia.2010.006486

Reilly, B. M., & Evans, A. T. (2006). Translating clinical research into clinical practice: Impact of using prediction rules to make decisions. *Annals of Internal Medicine*, *144*(3), 201.

Reynolds, T. (2011). Clinical trials: Can technology solve the problem of low recruitment? *British Medical Journal*, 342.

Richardson, W. S. (2007). We should overcome the barriers to evidence-based clinical diagnosis! *Journal of Clinical Epidemiology*, *60*(3), 217–227. doi:10.1016/j.jclinepi.2006.06.002

Segaran, T., Evans, C., & Taylor, J. (2009). *Programming the Semantic Web*. O'Reilly Media.

Steyerberg, E., Eijkemans, M., Van Houwelingen, J., Lee, K., & Habbema, J. (2000). Prognostic models based on literature and individual patient data in logistic regression analysis. *Statistics in Medicine*, *19*(2), 141–160. doi:10.1002/(SICI)1097-0258(20000130)19:2<141::AID-SIM334>3.0.CO;2-O

Stumme, G., Hotho, A., & Berendt, B. (2006). Semantic Web mining: State of the art and future directions. *Web Semantics: Science.Services and Agents on the World Wide Web*, *4*(2), 124–143. doi:10.1016/j.websem.2006.02.001

Treweek, S., Mitchell, E., Pitkethly, M., Cook, J., Kjeldstrom, M., & Taskila, T. (2010). Strategies to improve recruitment to randomised controlled trials. *Cochrane Database of Systematic Reviews*, (4): MR000013.

Treweek, S., Pearson, E., Smith, N., Neville, R., Sargeant, P., Boswell, B., & Sullivan, F. (2010). Desktop software to identify patients eligible for recruitment into a clinical trial: Using SARMA to recruit to the ROAD feasibility trial. *Informatics in Primary Care*, *18*(1), 51–58.

Weed, L. L. (1997). New connections between medical knowledge and patient care. *British Medical Journal*, *315*(7102), 231. doi:10.1136/bmj.315.7102.231

Weed, L. L., & Zimny, N. J. (1989). The problem-oriented system, problem-knowledge coupling, and clinical decision making. *Physical Therapy*, *69*(7), 565.

Wright, A., Sittig, D. F., Ash, J. S., Bates, D. W., Feblowitz, J., & Fraser, G. (2011). Governance for clinical decision support: Case studies and recommended practices from leading institutions. *Journal of the American Medical Informatics Association*, *18*(2), 187. doi:10.1136/jamia.2009.002030